

IEEE Signal Processing MAGAZINE

[VOLUME 31 NUMBER 1 JANUARY 2014]



SIGNAL AND IMAGE PROCESSING IN HYPERSPECTRAL REMOTE SENSING

ADVANCES, CHALLENGES, AND OPPORTUNITIES

IMAGE INPAINTING

SECURING DIGITAL REPUTATION

THE SIGNAL PROCESSING DISCIPLINE: PART 2





This publication offers open access options for authors

IEEE Open Access Publishing

What does IEEE Open Access mean to an author?

- Top quality publishing with established impact factors
- Increased exposure and recognition as a thought leader
- A consistent IEEE peer-review standard of excellence
- Unrestricted access for readers to discover your publications
- Great way to fulfill a requirement to publish open access

Learn more about IEEE Open Access Publishing:
www.ieee.org/open-access



CONTENTS

[VOLUME 31 NUMBER 1]

SPECIAL SECTION—SIGNAL AND IMAGE PROCESSING IN HYPERSPECTRAL REMOTE SENSING

22 FROM THE GUEST EDITORS

Wing-Kin Ma, José M. Bioucas-Dias, Jocelyn Chanussot, and Paul Gader

24 DETECTION ALGORITHMS IN HYPERSPECTRAL IMAGING SYSTEMS

Dimitris Manolakis, Eric Truslow, Michael Pieper, Thomas Cooley, and Michael Brueggeman

34 HYPERSPECTRAL TARGET DETECTION

Nasser M. Nasrabadi

45 ADVANCES IN HYPERSPECTRAL IMAGE CLASSIFICATION

Gustavo Camps-Valls, Devis Tuia, Lorenzo Bruzzone, and Jón Atli Benediktsson

55 MANIFOLD-LEARNING-BASED FEATURE EXTRACTION FOR CLASSIFICATION OF HYPERSPECTRAL DATA

Dalton Lunga, Saurabh Prasad, Melba M. Crawford, and Okan Ersoy

67 A SIGNAL PROCESSING PERSPECTIVE ON HYPERSPECTRAL UNMIXING

Wing-Kin Ma, José M. Bioucas-Dias, Tsung-Han Chan, Nicolas Gillis, Paul Gader, Antonio J. Plaza, ArulMurugan Ambikapathi, and Chong-Yung Chi

82 NONLINEAR UNMIXING OF HYPERSPECTRAL IMAGES

Nicolas Dobigeon, Jean-Yves Tourneret, Cédric Richard, José Carlos M. Bermudez, Stephen McLaughlin, and Alfred O. Hero

95 ENDMEMBER VARIABILITY IN HYPERSPECTRAL ANALYSIS

Alina Zare and K.C. Ho

105 COMPRESSIVE CODED APERTURE SPECTRAL IMAGING

Gonzalo R. Arce, David J. Brady, Lawrence Carin, Henry Arguello, and David S. Kittle

116 SPARSITY AND STRUCTURE IN HYPERSPECTRAL IMAGING

Rebecca M. Willett, Marco F. Duarte, Mark A. Davenport, and Richard G. Baraniuk

FEATURE

127 IMAGE INPAINTING

Christine Guillemot and Olivier Le Meur

COLUMNS

4 FROM THE EDITORRisky Long-Term Research: Two Success Stories in Audio Signal Processing
Abdelhak Zoubir**6 PRESIDENT'S MESSAGE**A Lasting Journey
K.J. Ray Liu**8 READER'S CHOICE**Top Downloads in IEEE *Xplore***15 SPECIAL REPORTS**Fresh Approaches Promise Wireless Quality and Reliability Improvements
John EdwardsAutomotive Industry Is a Key Component to the Success of the DSP Sector
Ron Schneiderman**145 STANDARDS IN A NUTSHELL**Royalty-Free Video Coding Standards in MPEG
Kiho Choi and Euee S. Jang**149 APPLICATIONS CORNER**Securing Digital Reputation in Online Social Media
Yuhong Liu and Yan (Lindsay) Sun**157 REFLECTIONS**The Discipline of Signal Processing: Part 2
Mos Kaveh and Li Deng

DEPARTMENTS

10 SOCIETY NEWS

New Members-at-Large, Directors-at-Large, and Class of Distinguished Lecturers for 2014

156 DATES AHEAD

IEEE SIGNAL PROCESSING magazine

IEEE SIGNAL PROCESSING MAGAZINE

Abdelhak Zoubir—*Editor-in-Chief*
Technische Universität Darmstadt, Germany

AREA EDITORS

Feature Articles

Marc Moonen—KU Leuven, Belgium

Columns and Forum

Andrea Cavallaro—Queen Mary, University
of London, UK

Andres Kwasinski—Rochester Institute
of Technology

Special Issues

Fulvio Gini—University of Pisa, Italy

e-Newsletter

Z. Jane Wang, —University of British Columbia

EDITORIAL BOARD

Moeness G. Amin—Villanova University

Sergio Barbarossa—University of Rome
“La Sapienza,” Italy

Mauro Barni—Università di Siena, Italy

Helmut Bölcskei—ETH Zürich, Switzerland

A. Enis Cetin—Bilkent University, Turkey

Yen-Kuang Chen—Intel Corporation

Bernd Girod—Stanford University

Michael Jordan—University of California,
Berkeley

Bastiaan Kleijn—Victoria University
of Wellington, New Zealand

Visa Koivunen—Aalto University, Finland

Ying-Chang Liang—Institute for Infocomm
Research, Singapore

Mark Liao—National Chiao-Tung University,
Taiwan

Hongwei Liu—Xidian University, China

V. John Mathews—University of Utah

Satoshi Nakamura—Nara Institute
of Science and Technology, Japan

Kuldip Paliwal—Griffith University, Australia

Béatrice Pesquet-Popescu—Télécom ParisTech,
France

H. Vincent Poor—Princeton University

Eli Saber—Rochester Institute of Technology

Ali Sayed—University of California, Los Angeles

Victor Solo—University of New South Wales,
Australia

Sergios Theodoridis—University of Athens,
Greece

Michail K. Tsatsanis—Entropic Communications

Min Wu—University of Maryland

ASSOCIATE EDITORS—
COLUMNS AND FORUM

Rodrigo Capobianco Guido —
University of Sao Paulo, Brazil

Aleksandra Mojsilovic —
IBM T.J. Watson Research Center

Douglas O'Shaughnessy — INRS, Canada

Gene Cheung — National Institute
of Informatics

Alessandro Vinciarelli — IDIAP-EPFL

Michael Gormish — Ricoh Innovations, Inc.

Xiaodong He — Microsoft Research

Fatih Porikli — MERL

Stefan Winkler — UIUC/ADSC, Singapore

Carlos Óscar Sánchez Sorzano — UAM, Spain

Azadeh Vosoughi — University of Central Florida

Danilo Mandic — Imperial College, UK

Roberto Togneri — The University of Western
Australia

Gail Rosen — Drexel University

ASSOCIATE EDITORS—
E-NEWSLETTER

Gwenael Doerr—Technicolor, France

Vitor Nascimento—University of São Paulo, Brazil

Shantanu Rane—MERL

Yan Lindsay Sun—University of Rhode Island

IEEE SIGNAL PROCESSING SOCIETY

Alex Acero—*President*

Rabab Ward—*President-Elect*

Konstantinos (Kostas) N. Plataniotis—*Vice
President, Awards and Membership*

Wan-Chi Siu—*Vice President, Conferences*

Alex Kot—*Vice President, Finance*

Mari Ostendorf—*Vice President, Publications*

Charles Bouman—*Vice President,
Technical Directions*

COVER

PLANE: ISTOCKPHOTO.COM/CHRISTOPHERDODGE.
FARM: ISTOCKPHOTO.COM/FLUTTER

IEEE PERIODICALS
MAGAZINES DEPARTMENT

Jessica Barragué
Managing Editor

Geraldine Krolin-Taylor
Senior Managing Editor

Susan Schneiderman
Business Development Manager
+1 732 562 3946 Fax: +1 732 981 1855

Felicia Spagnoli
Advertising Production Manager

Janet Dudar
Senior Art Director

Gail A. Schnitzer
Assistant Art Director

Theresa L. Smith
Production Coordinator

Dawn M. Melley
Editorial Director

Peter M. Tuohy
Production Director

Fran Zappulla
Staff Director, Publishing Operations

IEEE prohibits discrimination, harassment, and bullying.
For more information, visit
<http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html>.

IEEE SIGNAL PROCESSING MAGAZINE (ISSN 1053-5888) (ISPREG) is published bimonthly by the Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, 17th Floor, New York, NY 10016-5997 USA (+1 212 419 7900). Responsibility for the contents rests upon the authors and not the IEEE, the Society, or its members. Annual member subscriptions included in Society fee. Nonmember subscriptions available upon request. Individual copies: IEEE Members US\$20.00 (first copy only), nonmembers US\$201.00 per copy. Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of U.S. Copyright Law for private use of patrons: 1) those post-1977 articles that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 USA; 2) pre-1978 articles without fee. Instructors are permitted to photocopy isolated articles for noncommercial classroom use without fee. For all other copying, reprint, or republication permission, write to IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854 USA. Copyright ©2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals postage paid at New York, NY, and at additional mailing offices. Postmaster: Send address changes to IEEE Signal Processing Magazine, IEEE, 445 Hoes Lane, Piscataway, NJ 08854 USA. Canadian GST #125634188 Printed in the U.S.A.

Digital Object Identifier 10.1109/MSP.2013.2282426



EE Live!

Join us at the Industry's Leading Embedded Systems Event – EE Live!

With cutting edge product demonstrations, visionary keynotes, 200+ exhibitors and hundreds of essential training classes, EE Live! is the annual event for the embedded design community to learn, collaborate, and recognize excellence.

esc

More to see, do and learn than ever before.

EMBEDDED SYSTEMS
CONFERENCE

MARCH 31 – APRIL 3, 2014 | SAN JOSE, CA | EELIVESHOW.COM

Engineering Summits



INTERNET OF THINGS
ENGINEERING SUMMIT



HARDWARE STARTUP
ENGINEERING SUMMIT



BLACK HAT EMBEDDED
SECURITY SUMMIT

Engineering Boot Camps & Tutorials



ANDROID ENGINEERING
BOOT CAMP



SUPER C++ TUTORIAL
STROUSTRUP & SUTTER



FPGA ENGINEERING
BOOT CAMP

REGISTER NOW AND SAVE \$100 OFF AN ALL ACCESS
PASS WITH PROMO CODE: IEEEAD

SPONSORS

Diamond



Silver



[from the **EDITOR**]

Abdelhak Zoubir
Editor-in-Chief
zoubir@spg.tu-darmstadt.de
<http://signalprocessingsociety.org/publications/periodicals/spm>

Risky Long-Term Research: Two Success Stories in Audio Signal Processing

The year 2013 marks the loss of two pioneers in audio signal processing who were, at some point, members of the IEEE Signal Processing Society; Amar Gopal Bose died 12 July 2013 at the age of 83. His name is now identified with the high-quality audio equipment, such as speakers and noise-canceling headphones, that his company has been manufacturing for nearly 50 years. Ray Milton Dolby, whose name is synonymous with home sound systems and cinema, died 12 September 2013 at the age of 80.

Dr. Bose studied electrical engineering at the Massachusetts Institute of Technology (MIT) under Norbert Wiener, along with Dr. Y.W. Lee. He returned to MIT as a professor in 1956 and was a member of its faculty until 2001. Primarily known for his work on audio signal processing and acoustics and for the successful company he set up to pursue long-term research in acoustics, Dr. Bose also performed research on nonlinear control theory, which led to an electromagnetic active control suspension for automobiles, a product in today's market. He received, respectively, in 1972 and 2010, his IEEE Fellowship and the IEEE/RSE Wolfson James Clerk Maxwell Award.

Dr. Dolby graduated in electrical engineering from Stanford University in 1957 and completed his Ph.D. in physics at Cambridge University in 1965. Dolby was best known for his work in noise reduction of audio recordings and surround sound. He received his IEEE Fellowship in 2010 and, in the same year, the IEEE Edison Medal.

Bose and Dolby are prime examples of successful research engineers whose

innovations have spawned successful industrial products and have had an impact on our daily lives. The question one wishes to answer is "What made their success possible?" Is it the conduct of high-risk research that led to a big return? Or did other factors also play a key role in their successes?

There exist many success stories in engineering and science; a wealth of literature on the greatest ideas, the greatest feuds; and on discovery, creativity, and innovation, written by engineers or scientists, but also by sociologists. Many of us have read some of these books and the fascinating stories they recount. I recall a discussion on the history of science I had at a dinner table with colleagues and friends during the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) in Atlanta. I don't recall how the discussion started, but I very well remember that at some point we discussed the discovery of the number and numerical digit "zero" (0). The discussion was both entertaining and exciting. The number that plays a central role in mathematics and seems obvious today was not easy to conceptualize, and now several books exist on the history of its discovery.

About two years ago, I read an interesting book by Simonton [1] who examined scientific creativity from four principal perspectives. These are *logic*, *genius*, *chance*, and *Zeitgeist*. Simonton poses the question whether creativity is assured once a scientist masters the logic of science and the substance of a particular discipline. Clearly, this would contradict the idolizing praise, bestowed on scientific genius. Joseph Louis Lagrange called Isaac Newton "the greatest genius that had ever existed" while Albert Einstein reported "to these elementary laws there leads no logical path, but only intuition." Not all of us

would accept the limitations of the logic stance. Arguably, great scientists don't follow logic. Then, there is chance. Many known episodes of serendipity, which is defined as the faculty of making fortunate discoveries by accident, exist. Louis Pasteur said, "chance favors only the prepared mind." According to a dictionary, *Zeitgeist* is the defining spirit (*Geist*) or mood of a particular period of history as shown by the ideas and beliefs of the time (*Zeit*). The above four viewpoints are incoherent except for the logic and the *Zeitgeist*. They are more independent than mutually supportive, which is not surprising [1].

Coming back to Bose and Dolby, it is reported that Dolby worked at an audio-recording company as a teenager while Dr. Bose began repairing radio sets for pocket money for repair shops in Philadelphia at the age of 13. This is an indication of their passion to engineering practice from an early age. This passion is what drives many of us to do research and strive for the advancement of technology with novelties, even when it means taking risks. In a 2004 interview in *Popular Science* magazine, Bose said "I would have been fired a hundred times at a company run by M.B.A.s. But I never went into business to make money. I went into business so that I could do interesting things that hadn't been done before."

Conducting interesting research and being passionate about what we do is what keeps us motivated and is likely to lead to innovation with a high impact to society.

REFERENCE

[1] D. K. Simonton, *Creativity in Science: Chance, Logic, Genius, and Zeitgeist*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Digital Object Identifier 10.1109/MSP.2013.2285947

Date of publication: 5 December 2013



Who develops green technologies, video games, rescue robots and more?

Explore the amazing world of engineers—
all in one web site...

TryEngineering.org

- **See** the exciting work that engineers do
- **Learn** how engineers make a difference
- **Start** now to prepare to be an engineer
- **Play** online games and challenges
- **Download** free engineering lesson plans
- **Explore** the fascinating FAQ
- **Search** for accredited engineering programs
- **Find** competitions and summer camps



Visit www.tryengineering.org today!

Brought to you by:



[president's MESSAGE]

K.J. Ray Liu
2012–2013 SPS President
kjrliu@umd.edu



A Lasting Journey

A few years ago, I was asked by my university's president to join a reception honoring some dignitaries from another country. A woman who appeared to be of nobility asked me what type of work I do.

"Signal processing," I said. Her eyes lit up with excitement and said that the traffic signals near her home were operating strangely and asked if I could perhaps offer my expertise. Not knowing how to respond, I pretended to be busy eating my soup...

After the 2008 International Conference on Acoustics, Speech, and Signal Processing ended in Las Vegas, I was one of the last to leave since I stayed for the Board of Governors meeting immediately after. On my way to the airport, the taxi driver asked me if I was in town for the game.

I responded simply, "No, for a conference." The driver replied, "For the signal processing conference?" I was quite impressed that he was aware of our field! When asked how he knew, the driver replied that he had taken quite a few attendees to the airport, and they told him what signal processing is about.

So the world is not that hopeless. Since I raised the "signal processing inside" phenomenon [1] in 2004, however, it seems that signal processing is still quite inside. Over the past decade there has been an overwhelming feeling of powerlessness in terms of how/where to address this visibility issue that is unique to our field.

When I was a member of the IEEE Technical Activities Board, more than 40

presidents of Societies and/or councils would gather and often remark how signal processing is ubiquitous when we were working together. Our field of interest as officially approved by that body is defined by this opening statement: "Signal processing is the enabling technology for the generation, transformation, extraction, and interpretation of information."

If you were to refer to this statement among your colleagues in engineering, or even science, they would most likely know what you are talking about.

Nevertheless, when we go beyond our comfort zone, signal processing becomes an unknown. So who is to speak for us? Is it important for "the public" to know more about us?

Well, of course! Otherwise the field may not receive the recognition and influence it deserves, as well as the research funding to accompany. But the visibility of our field is certainly about more than sheer pride and ego.

The taxi driver knew about signal processing because people had repeatedly talked to him about it. I would bet that if the lady of nobility had been exposed to or heard of signal processing a couple of times, she would have instead asked me to fix her TV or cell phone, both better propositions than the traffic signals...

As a result, the IEEE Signal Processing Society (SPS) Executive Committee has decided to face this daunting issue again by establishing an ad hoc committee that will work with John Cozzens of the U.S. National Science Foundation. I hope that we can at least have a starting point for us to work toward developing a concrete plan in the near future.

Over the past few years, the SPS has changed significantly and grown a lot. It

has been my honor and privilege to work with so many dedicated colleagues who offer their unselfish volunteer time and effort to make our Society better. There are so many people to thank, and I would definitely miss some names if I tried to name them all. Please allow me to say "Thank you so much!" to all of you.

It has been my pleasure and good fortune to have had the opportunity to work with the incoming SPS President Alex Acero over many years. As an outstanding researcher and experienced industry leader, Alex is ready to move the Society to a higher level. I wish him and the Society all the best!

I would like to thank all of the SPS staff, especially Rich Baseil and Theresa Argiropoulos. The last two years were perhaps the most difficult time for our office due to a severe shortage of manpower. But under their amazing leadership, not only did the staff manage to provide outstanding services but they also made the quality of member services a new priority.

I would also like to thank Tracy Chung for helping to proofread every draft of my message in an effort to make the reading more enjoyable.

And so a chapter is over, yet another is about to begin. It is time for an old soldier to fade away. But the journey continues...

REFERENCE

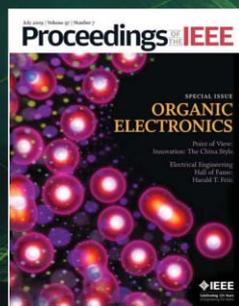
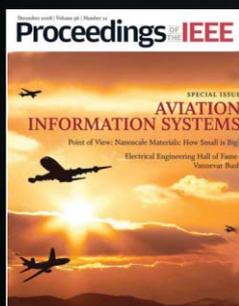
[1] K. J. R. Liu, "Signal processing inside?" *IEEE Signal Process. Mag.*, vol. 21, no. 5, p. 2, Sept. 2004.

Digital Object Identifier 10.1109/MSP.2013.2285263
Date of publication: 5 December 2013

SP

Think you know about the latest technology? You haven't even scratched the surface.

See all the layers of technology with *Proceedings of the IEEE*. Every issue brings comprehensive, in-depth coverage on technology breakthroughs. From outlining new uses for existing technology to detailing cutting-edge innovations in a variety of disciplines, you'll find the breadth of content and depth of knowledge that only IEEE can provide.



Go beyond the surface—subscribe today.
www.ieee.org/proceedings



[reader's **CHOICE**]

Top Downloads in IEEE *Xplore*

The “Reader’s Choice” column in *IEEE Signal Processing Magazine* contains a list of articles published by the IEEE Signal Processing Society (SPS) that ranked

among the top 100 most downloaded IEEE *Xplore* articles. This issue is based on download data through August 2013. The table below contains the citation information for each article and the rank obtained in IEEE

Xplore. The highest rank obtained by an article in this time frame is indicated in bold. Your suggestions and comments are welcome and should be sent to Associate Editor Michael Gormish (gormish@ieee.org).

TITLE, AUTHOR, PUBLICATION YEAR IEEE SPS PUBLICATIONS	ABSTRACT	RANK IN IEEE TOP 100 (2013)						N TIMES IN TOP 100 (SINCE JAN 2011)
		AUG	JUL	JUN	MAY	APR	MAR	
A TUTORIAL ON PARTICLE FILTERS FOR ONLINE NONLINEAR/NON-GAUSSIAN BAYESIAN TRACKING Arulampalam, M.S.; Maskell, S.; Gordon, N.; Clapp, T. <i>IEEE Transactions on Signal Processing</i> vol. 50, no. 2, 2002, pp. 174–188	This paper reviews optimal and suboptimal Bayesian algorithms for nonlinear/non-Gaussian tracking problems, with a focus on particle filters. Variants of the particle filter are introduced within a framework of the sequential importance sampling (SIS) algorithm and compared with the standard EKF.	4	4	2	4	11	81	29
IMAGE PROCESSING USING SMOOTH ORDERING OF ITS PATCHES Ram, I.; Elad, M.; Cohen, I. <i>IEEE Transactions on Image Processing</i> vol. 22, no. 7, 2013, pp. 2764–2774	This paper extracts overlapping image patches, orders these patches and applies one dimensional filtering to the reordered set of pixels. These techniques are applied to denoising and inpainting.	9	15	99				3
AN INTRODUCTION TO COMPRESSIVE SAMPLING Candes, E.J.; Wakin, M.B. <i>IEEE Signal Processing Magazine</i> vol. 25, no. 2, Mar. 2008, pp. 21–30	This article surveys the theory of compressive sampling, also known as compressed sensing or CS, a novel sensing/sampling paradigm that goes against the common wisdom in data acquisition.	13	19		5	6	7	31
PRIVACY PRESERVING DATA SHARING WITH ANONYMOUS ID ASSIGNMENT Dunning, L.A.; Kresman, R. <i>IEEE Transactions on Information Forensics and Security</i> vol. 8, no. 2, 2013, pp. 402–413	Existing and new algorithms for assigning anonymous IDs are examined with respect to communication and computational requirements. The new algorithms are built on top of a secure sum data mining operation using Newton’s identities and Sturm’s theorem.	23					81	2
IMAGE QUALITY ASSESSMENT: FROM ERROR VISIBILITY TO STRUCTURAL SIMILARITY Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. <i>IEEE Transactions on Image Processing</i> vol. 13, no. 4, 2004, pp. 600–612	This paper introduces a framework for quality assessment based on the degradation of structural information. Within this framework a structure similarity index is developed and evaluated. MATLAB code available.	38	30	50	38	33	69	11

Digital Object Identifier 10.1109/MSP.2013.2282791

Date of publication: 5 December 2013

TITLE, AUTHOR, PUBLICATION YEAR IEEE SPS PUBLICATIONS	ABSTRACT	RANK IN IEEE TOP 100 (2013)						N TIMES IN TOP 100 (SINCE JAN 2011)
		AUG	JUL	JUN	MAY	APR	MAR	
LEARNING-BASED, AUTOMATIC 2D-TO-3D IMAGE AND VIDEO CONVERSION Konrad, J.; Wang, M.; Ishwar, P.; Wu, C.; Mukherjee, D. <i>IEEE Transactions on Image Processing</i> vol. 22, no. 9, 2013, pp. 3485–3496	This paper uses repositories of true three-dimensional (3-D) data to learn conversion of two-dimensional (2-D) to 3-D data. One method learns a point mapping based on image/video attributes including color, position, and motion using regression. A second method globally estimates the depth map using a nearest neighbor technique.	64						1
COMPRESSIVE SENSING [LECTURE NOTES] Baraniuk, R.G. <i>IEEE Signal Processing Magazine</i> vol. 24, no. 4, 2007, pp. 118–121	This lecture note presents a new method to capture and represent compressible signals at a rate significantly below the Nyquist rate. This method, called compressive sensing, employs nonadaptive linear projections that preserve the structure of the signal; the signal is then reconstructed from these projections using an optimization process.	71			94			5
IMAGE SUPER-RESOLUTION VIA SPARSE REPRESENTATION Yang, J.; Wright, J.; Huang, T.S.; Ma, Y. <i>IEEE Transactions on Image Processing</i> vol. 19, no. 11, 2010, pp. 2861–2873	This paper presents an approach to single-image superresolution, based upon sparse signal representation of low- and high-resolution patches.	82	68					3
ADAPTIVE MOBILE CLOUD COMPUTING TO ENABLE RICH MOBILE MULTIMEDIA APPLICATIONS Wang, S.; Dey, S. <i>IEEE Transactions on Multimedia</i> vol. 15, no. 4, 2013, pp. 870–883	Mobile cloud computing can help bridge the gap between PCs and mobile devices, providing mobile applications the capabilities of cloud servers and storage, possibly enabling a new generation of truly ubiquitous multimedia applications on mobile devices: cloud mobile media (CMM) applications.		48					1
MULTI-VIEW VIDEO REPRESENTATION BASED ON FAST MONTE CARLO SURFACE RECONSTRUCTION Salvador, J.; Casas, J.R. <i>IEEE Transactions on Image Processing</i> vol. 22, no. 9, 2013, pp. 3342–3352	A Monte Carlo discrete surface reconstruction method for foreground objects with static background is presented, which outperforms volumetric techniques and is suitable for GPU environments. Then, a fast meshing algorithm is applied, which allows interpolating a continuous surface from the discrete reconstructed points.		53					1

SP

moving?

You don't want to miss any issue of this magazine!

change your address

BY E-MAIL: address-change@ieee.org

BY PHONE: +1 800 678 IEEE (4333) in the U.S.A.
or +1 732 981 0060 outside the U.S.A.

ONLINE: www.ieee.org, click on quick links, change contact info

BY FAX: +1 732 562 5445

Be sure to have your member number available.

[society NEWS]

New Members-at-Large, Directors-at-Large, and Class of Distinguished Lecturers for 2014

The Board of Governors (BoG) is the governing body that oversees the activities of the IEEE Signal Processing Society (SPS). The SPS BoG has the responsibility of establishing and implementing policy and receiving reports from its standing boards and committees. Members-at-large represent the member view point in the Board decision making. They typically review, discuss, and act upon a wide range of items affecting the actions, activities, and health of the Society.

MEMBERS-AT-LARGE ELECTION

Three new members-at-large will take their seats on the SPS BoG beginning 1 January 2014 and will serve until 31 December 2016. Eight candidates competed for the three member-at-large positions. The successful candidates represent a broad spectrum of the SPS. The BoG will confirm the election results at their 6 December 2013 Board meeting. The successful candidates are Geert Leus (Delft University of Technology), Helen Meng (Chinese University of Hong Kong), and Fernando Pereira (Instituto de Telecomunicacoes).

Completing their terms as members-at-large on 31 December 2013 are Petar M. Djuric (Stony Brook University), William Clem Karl (Boston University), and Sergios Theodoridis (University of Athens).

REGIONAL DIRECTORS-AT-LARGE ELECTION RESULTS

Two new regional directors-at-large will take their seats on the SPS BoG and Membership Board beginning 1 January 2014 and will serve until 31 December

2015. Regional directors-at-large promote and foster local activities (such as conferences, meetings, and social networking) and encourage new Chapter development; represent their Regions to the core of SPS; offer advice to improve membership relations, recruiting and service to their regions; guide and work with their corresponding Chapters to serve their members; and assist the vice president, Awards and Membership, in conducting Chapter reviews. The new regional directors-at-large are as follows:

- *Regions 7 and 9*: Douglas O'Shaughnessy, INRS

- *Region 10*: Hong-Yuan Mark Liao, Institute of Information Science.

Completing their terms as regional directors-at-Large on 31 December 2013 are

- *Regions 7 and 9*: Roxana Saint-Nom, Instituto Tecnológico de Buenos Aires

- *Region 10*: Ta-Sung Lee, National Chiao Tung University.

2014 MEMBERS-AT-LARGE

GEERT LEUS

Geert Leus received the electrical engineering degree and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Belgium, in 1996 and 2000, respectively. Currently, he is a full professor at the Faculty of Electrical Engineering, Mathematics, and Computer Science of the Delft University of Technology, The Netherlands. In 2001 and 2002, he was a visiting researcher and lecturer at the University of Minnesota. His research interests are in the broad area of signal processing for communications, with a special focus on underwater communications, localization, and sensor networks. In this area, he has approximately

80 journal and 190 conference publications, most of them in IEEE flagship journals and conferences. For his outstanding work, he received a 2002 IEEE SPS Young Author Best Paper Award and a 2005 IEEE SPS Best Paper Award.

Prof. Leus is an IEEE Fellow. His IEEE and Society activities include: member, vice chair, chair, and past chair, SPS Signal Processing for Communications and Networking Technical Committee (2002–2011); member, SPS Sensor Array and Multichannel Technical Committee (2012–present); general and technical cochair, 2009 IEEE Workshop on Signal Processing Advances in Wireless Communications; technical cochair, 2012 IEEE Workshop on Signal Processing Advances in Wireless Communications; technical program committee member of over 30 IEEE-sponsored conferences; organizer of special sessions at IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Asilomar, IEEE Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP), and IEEE International Conference on Ultra-Wideband (ICUWB); associate editor, *IEEE Transactions on Signal Processing* (2006–2010), *IEEE Transactions on Wireless Communications* (2002–2006) and *IEEE Signal Processing Letters* (2001–2005); and guest editor, “Soft Detection for Wireless Transmission,” *IEEE Journal of Selected Topics in Signal Processing* (2011). Other volunteer activities include: editor-in-chief, *EURASIP Journal on Advances in Signal Processing* (2013–present); associate editor, *EURASIP Journal on Advances in Signal Processing* (2004–2012); guest editor, “Object Tracking and Monitoring Using Advanced Signal Processing Techniques,” *EURASIP Journal on Advances in Signal Processing*

Digital Object Identifier 10.1109/MSP.2013.2282089

Date of publication: 5 December 2013

(2011); lead guest editor, “Reliable Communications over Rapidly Time-Varying Channels,” *EURASIP Journal on Advances in Signal Processing* (2006); lead guest editor, “Improved CDMA Detection Techniques for Future Wireless Systems,” *EURASIP Journal on Advances in Signal Processing* (2005); guest editor, “Interference Management in Wireless Communication Systems: Theory and Applications,” *EURASIP Journal on Wireless Communications & Networking* (2011); and guest editor, “Compressive Sensing in Communications,” *Elsevier Physical Communications* (2011) and advisory board member Campus do Mar (2010–present).

HELEN MENG

Helen Meng received the S.B., S.M., and Ph.D. degrees, all in electrical engineering, from the Massachusetts Institute of Technology, Cambridge. She joined The Chinese University of Hong Kong (CUHK) in 1998, where she is currently a professor and chair of the Department of Systems Engineering and Engineering Management. In 1999, she established the Human–Computer Communications Laboratory and serves as its director. In 2005, she established the Microsoft–CUHK Joint Laboratory for Human–Centric Computing and Interface Technologies and serves as its founding director. This laboratory was conferred the national status of the Ministry of Education of China Key Laboratory in 2008. She also served as associate dean (research) of the Faculty of Engineering from 2006 to 2010. She is a Fellow of the IEEE.

FERNANDO PEREIRA

Fernando Pereira graduated with a degree in electrical and computer engineering from Instituto Superior Técnico (IST), Universidade Técnica de Lisboa, Portugal, in 1985. He received the M.Sc. and Ph.D. degrees in electrical and computer engineering from IST in 1988 and 1991, respectively.

He is currently a professor in the Electrical and Computer Engineering Department at IST. He is responsible for IST’s participation in many national and international research projects. He acts

as project evaluator and auditor for various organizations.

Prof. Pereira is an IEEE and EURASIP Fellow. He has been involved in the following professional activities: area editor, *Signal Processing: Image Communication Journal*; editorial board member, *IEEE Signal Processing Magazine* (2007–present); associate editor, *IEEE Transactions of Circuits and Systems for Video Technology*, *IEEE Transactions on Image Processing* (2002–2007), *IEEE Transactions on Multimedia* and *IEEE Signal Processing Magazine*; member, SPS Image, Video, and Multidimensional Signal Processing Technical Committee (2004–2009), Multimedia Signal Processing Technical Committee (2006–present), CAS Visual Signal Processing and Communications Technical Committee, and Multimedia Systems and Applications Technical Committee; and SPS Distinguished Lecturer (2005). Since January 2013, he has been the editor-in-chief of *IEEE Journal of Selected Topics in Signal Processing*.

He is and has been a member of the scientific and program committees of many international conferences. Among others, he has been the general chair, 2007 Picture Coding Symposium (PCS), and Technical Program cochair, 2010 International Conference on Image Processing (ICIP). He participates in the MPEG standardization activities, notably as the head of the Portuguese delegation; chair of the MPEG Requirements Group; and chair of many ad hoc groups related to the MPEG-4 and MPEG-7 standards. He is a coeditor of *The MPEG-4 Book* and *The MPEG-21 Book*.

Prof. Pereira won the first Portuguese IBM Scientific Award in 1990, an ISO Award for Outstanding Technical Contribution for his contributions to the MPEG-4 Visual Standard in 1998, and an Honorable Mention of the UTL/Santander Totta Award for Electrotechnical Engineering in 2009, 2010, 2011, and 2012. He won the “Excellent Professor” Award from the Electrical and Computer Engineering Department of IST in 2010, 2011, and 2012. He has contributed more than 200 papers in international journals, conferences, and workshops and gave several

invited talks at conferences and workshops. His areas of interest are video analysis, coding, description and adaptation, and advanced multimedia services.

REGIONAL DIRECTORS-AT-LARGE

DOUGLAS O'SHAUGHNESSY

Douglas O’Shaughnessy received the B.Sc. and M.Sc. degrees in 1972 and the Ph.D. degree in 1976 from the Massachusetts Institute of Technology, Cambridge. He has been a professor at INRS-Telecommunications (University of Quebec) in Montreal, Canada, since 1977 (the institute was renamed INRS-EMT in 2002). For this same period, he also taught as an adjunct professor at McGill University in the Department of Electrical Engineering. From 1991 to 1997 and 2002–present, he has been the program director at INRS-EMT. Dr. O’Shaughnessy has been a teacher and researcher in the speech communication field for 36 years. His interests include automatic speech analysis, enhancement, and recognition. His research team is currently working to improve various aspects of automatic voice dialogues.

He is a Fellow of the IEEE (2006) and the Acoustical Society of America (1992). He is a current member and chair, SPS Speech and Language Processing Technical Committee (1984–1985, 2007–2009, and 2011–2014); associate editor, *IEEE Signal Processing Magazine* (2009–2013); member, Technical Activities Board (TAB) Periodicals Committee (2011–2013); Acoustical Society of America: associate editor of *Journal of the Acoustical Society of America Express Letters* (2009–2013); International Speech Communication Association (ISCA): vice president and elected member of the ISCA board (2009–2013); associate editor, *EURASIP Journal on Advances in Signal Processing* (2008–2013); editor-in-chief, *EURASIP Journal on Audio, Speech, and Music Processing* (2005–2013); member, Technical Committee on Speech of the Acoustical Society of America (1995–1997); associate editor, *IEEE Transactions on Speech and Audio Processing* (1995–1999), *Journal of Acoustical Society of America* (1998–2010); member, SPS conference

board (2000–2005); member-at-large, SPS BoG (2002–2004); and general chair, 2004 ICASSP.

He is the author of *Speech Communications: Human and Machine* (IEEE Press, 2000). He is coauthor, with Li Deng, of *Speech Processing: A Dynamic and Optimization-Oriented Approach* (Marcel Dekker, 2003). He presented tutorials on speech recognition at ICASSP in 1996, 2001, and 2009, as well as the 2003 IEEE International Conference on Communications (ICC). He has been a regular presenter at the major speech conferences of Eurospeech and ICSLP (now Interspeech) and has had papers at virtually every ICASSP since 1986. In *IEEE Xplore*, he has 128 listings, on Google Scholar, 25 publications of more than 25 citations.

HONG-YUAN MARK LIAO

Hong-Yuan Mark Liao received his B.S. degree in physics from National Tsing-Hua University, Hsin-Chu, Taiwan, in 1981 and his M.S. and Ph.D. degrees in electrical engineering from Northwestern University in 1985 and 1990, respectively. In 1991, he joined the Institute of Information Science, Academia Sinica, Taiwan, where he is currently a distinguished research fellow. During 2009–2011, he was the division chair of the Computer Science and Information Engineering Division II, National Science Council of Taiwan. He is jointly appointed as a professor of the Computer Science and Information Engineering Department of National Chiao-Tung University and the Department of Electrical Engineering of National Cheng Kung University. During 2009–2012, he was jointly appointed as the multimedia information chair professor of National Chung Hsing University. Since 2010, he has been an adjunct chair professor of Chung Yuan Christian University. His current research interests include multimedia signal processing, video-based surveillance systems, video forensics, and multimedia protection.

Dr. Liao is a Fellow of the IEEE. He is the recipient of the Young Investigators' Award from Academia Sinica in 1998, Distinguished Research Award from the National Science Council of Taiwan in 2003 and 2010, National Invention Award

of Taiwan in 2004, Distinguished Scholar Research Project Award from National Science Council of Taiwan in 2008, and Academia Sinica Investigator Award in 2010. His professional activities include: cochair, 2004 International Conference on Multimedia and Exposition (ICME); technical cochair, 2007 ICME; general cochair, 17th International Conference on Multimedia Modeling; President, Image Processing, and Pattern Recognition Society of Taiwan (2006–2008); editorial board member, *IEEE Signal Processing Magazine*; associate editor, *IEEE Transactions on Image Processing*, *IEEE Transactions on Information Forensics and Security* (2009–2012) and *IEEE Transactions on Multimedia* (1998–2001).

2014 CLASS OF DISTINGUISHED LECTURERS

The IEEE SPS's Distinguished Lecturer Program provides the means for Chapters to have access to well-known educators and authors in the fields of signal processing to lecture at Chapter meetings. While many IEEE Societies have similar programs, the SPS provides financial support for the Chapters to take advantage of this service. Chapters interested in arranging lectures by the Distinguished Lecturers can obtain information from the Society's Web page (<http://www.signalprocessingsociety.org/lecturers/distinguished-lecturers/>) or by sending an e-mail to sp.info@ieee.org.

Candidates for the Distinguished Lecturer Program are solicited from the Society technical committees, editorial boards, Chapters, and other boards and committees by the Awards Board. The Awards Board vets the nominations, and the BoG approves the final selection. Distinguished Lecturers are appointed for a term of two calendar years. The following Distinguished Lecturers were named for 2014: Maria S. Greco, Alex C. Kot, Bhaskar D. Rao, Andrew C. Singer, and Akihiko (Ken) Sugiyama.

MARIA S. GRECO

Maria S. Greco graduated in electronic engineering in 1993 and received the Ph.D. degree in telecommunication engineering from the University of Pisa,

Italy, in 1998. From December 1997 to May 1998, she was with the Georgia Tech Research Institute, Atlanta, as a visiting research scholar, where she carried on research activity in the field of radar detection in non-Gaussian background. In 1993, she joined the Department of Information Engineering of the University of Pisa, where she has been an associate professor since December 2011.

Dr. Greco has been a general chair, technical chair, and organizing committee member of the 2006 European Signal Processing Conference, the 2007 International Waveform Diversity and Design Conference, the 2008 IEEE Radar Conference, IEEE Workshop on Computational Advances in Multisensor Adaptive Processing (2009 and 2011), the 2010 International Workshop on Cognitive Information Processing, the 2014 IEEE Sensor Array and Multichannel Signal Processing Workshop, and ICASSP 2014. She was a guest coeditor of the special issue "Adaptive Waveform Design for Agile Sensing and Communication" (June 2007), *IEEE Journal of Selected Topics in Signal Processing*, and lead guest editor of the special issue "Modeling and Processing of Radar Signals for Earth Observation" (August 2008) of *International Journal of Navigation and Observation*.

Dr. Greco has been an associate editor, *IET Proceedings—Sonar, Radar, and Navigation* since 2009; *IEEE Transactions on Signal Processing* (2009–2013); associate editor-in-chief, *IEEE Aerospace and Electronic Systems Magazine* since 2011; editorial board member, *Journal of Advances in Signal Processing* since 2008; member, IEEE SPS Signal Processing Theory and Methods Technical Committee (2009–present); chair, EDU Subcommittee (2009–2014); member, IEEE SPS Sensor Array and Multichannel Technical Committee (2010–2015); and member, IEEE Aerospace and Electronic Systems Society (AESS) Radar Panel and IEEE AESS BoG (2012–2014).

She is an IEEE Fellow. She was co-recipient, IEEE AESS's Barry Carlton Award for Best Paper (2001); and recipient, Fred Nathanson Young Engineer of the Year Award (2008) "for contributions to signal processing, estimation,

and detection theory.” Dr. Greco’s general interests are in the areas of statistical signal processing, estimation, and detection theory. In particular, her research interests include radar clutter models, spectral analysis, coherent and incoherent detection in non-Gaussian clutter, constant false alarm rate techniques, radar waveform diversity and bistatic/multistatic radars. She coauthored three book chapters and more than 130 journal and conference papers.

Dr. Greco’s lecture topics include statistical analysis methods for radar clutter modeling, advanced techniques of radar detection in non-Gaussian background, and sensor selection for multistatic radar networks.

ALEX C. KOT

Alex C. Kot received his B.S. degree in electrical engineering and M.B.A. degrees both from the University of Rochester. He obtained his M.S. and Ph.D. degrees in electrical engineering from the University of Rhode Island. Prof. Kot has been with the Nanyang Technological University, Singapore since 1991. He headed the Division of Information Engineering with more than 40 faculty members at the School of Electrical and Electronic Engineering for eight years and served as associate chair/research and vice dean research for the School of Electrical and Electronic Engineering. He is currently a professor and associate dean in the College of Engineering. He is the director of the Rapid-Rich Object Search (ROSE) Lab, which partners with Peking University.

Dr. Kot was an associate editor, *IEEE Transactions on Signal Processing* (2001–2003); *IEEE Transactions on Image Processing* (2010–2012); *IEEE Transactions on Multimedia* (2008–2012); *IEEE Signal Processing Letters* (2009–2012); *IEEE Signal Processing Magazine* (2008–2010); *IEEE Journal of the Special Topics in Signal Processing* (2010–2012); *IEEE Transactions on Circuits and Systems for Video Technology*; *IEEE Transactions on Circuits and Systems Part III*; *IEEE Transactions on Information, Forensics, and Security* (2011–present); and the *EURASIP Journal of Advanced Signal Processing*. He was also a guest editor for special issues of

IEEE Transactions on Circuits and Systems for Video Technology and *IEEE Journal of Applied Signal Processing*.

He has served the IEEE SPS in various capacities such as general cochair, ICIP 2004 and chair, SPS Chapters Committee. He served as member, IEEE Fellow Evaluation Committee; vice president, finance, IEEE SPS (2013–2014); member, SPS Conference Board (2013–2014); and member, SPS Publications Board (2013–2014). He received the Best Teacher of the Year Award and received Best Paper Awards as coauthor from the International Conference on Pattern Recognition, IEEE International Workshop on Information Forensics and Security, International Conference on Electronic Commerce, and International Workshop on Digital Watermarking. He was the IEEE Circuits and Systems Society Distinguished Lecturer in 2005 and 2006 and is a Fellow of the IEEE, IES, and the Academy of Engineering, Singapore.

Dr. Kot has published extensively in the areas of signal processing for communication, biometrics, data hiding, image forensics, and information security. His new research area is in the domain object search and recognition. His lecture topics include “Is Your Biometric Data Safe?” and “Can Digital Photos Be Trusted?”

BHASKAR D. RAO

Bhaskar D. Rao received the B.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in 1979 and the M.S. and Ph.D. degrees from the University of Southern California, Los Angeles, in 1981 and 1983, respectively. Since 1983, he has been with the University of California at San Diego, La Jolla, where he is currently a professor with the Department of Electrical and Computer Engineering. He holds the Ericsson Endowed Chair in Wireless Access Networks and was the director of the Center for Wireless Communications (2008–2011).

Prof. Rao was elected IEEE Fellow in 2000 “for his contributions to the statistical analysis of subspace algorithms for harmonic retrieval.” His work has received

several paper awards; Best Paper Award (2013) for “Multicell Random Beamforming with CDF-Based Scheduling: Exact Rate and Scaling Laws”; SPS Best Paper Award (2012) for “An Empirical Bayesian Strategy for Solving the Simultaneous Sparse Approximation Problem”; Stephen O. Rice Prize Paper Award in the Field of Communication Systems (2008) for “Network Duality for Multiuser MIMO Beamforming Networks and Applications”; and Best Paper Award (2000) for “PDF Optimized Parametric Vector Quantization of Speech Line Spectral Frequencies.”

His students have received several student paper awards; Best Student Paper Award (2006) for D.P. Wipf for the paper “Analysis of Empirical Bayesian Methods for Neuroelectromagnetic Source Localization,” by D.P. Wipf, R.R. Ramirez, J.A. Palmer, S. Makeig, and B.D. Rao; Student Paper Award (2006) for Jun Zheng for the paper “Capacity Analysis of Multiple Antenna Systems with Mismatched Channel Quantization Schemes,” by J. Zheng and B.D. Rao; Best Student Paper Award (2005) for J. McCall and D. Wipf for the paper “Lane Change Intent Analysis Using Robust Operators and Sparse Bayesian Learning” by J. McCall, D. Wipf, M. Trivedi and B.D. Rao; and Student Paper Award (2005) for Haichang Sui for the paper “RAKE Finger Placement for CDMA Downlink Equalization,” by H. Sui, E. Masry and B.D. Rao.

Prof. Rao has been a member, Statistical Signal and Array Processing Technical Committee; member, Signal Processing Theory and Methods Technical Committee (1999–2004); Signal Processing for Communications and Networking Technical Committee (2005–2007); member, Machine Learning for Signal Processing Technical Committee (2012–present); editorial board member, *EURASIP Signal Processing Journal*; and technical member, several IEEE conferences.

Prof. Rao’s interests are in the areas of digital signal processing, estimation theory, and optimization theory, with applications to digital communications, speech signal processing, and biomedical signal processing.

Prof. Rao’s lecture topics include “Bayesian Methods for Sparse Signal

Recovery and Compressed Sensing and Space-Time Signal Processing” and “Theory, Methods, and Applications.”

ANDREW C. SINGER

Andrew C. Singer received the S.B., S.M., and Ph.D. degrees, all in electrical engineering and computer science, from the Massachusetts Institute of Technology. From 1996 to 1998, he was a research scientist at Sanders, a Lockheed Martin company in Manchester, New Hampshire. Since 1998, he has been on the faculty of the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, where he is currently a professor in the Electrical and Computer Engineering Department and the Coordinated Science Laboratory.

In 2000, Dr. Singer cofounded Intersymbol Communications, Inc., based in Champaign, Illinois, which built the world's fastest signal-processing-enhanced receivers for 10 Gb/s optical communications. In 2007, Intersymbol Communications, Inc. was acquired by Finisar Corporation (FNSR). He continues to work in areas related to signal processing algorithms and their potential to enhance mixed-signal analog and digital circuits both in his research and as assistant director of the SRC-funded Systems on Nanoscale Information Fabrics (SONIC) center. His research interests include signal processing and communication systems.

He received the National Science Foundation CAREER Award (2000), the Xerox Faculty Research Award (2001), and was named a Willett Faculty Scholar (2002). In 2005, he was appointed director, Technology Entrepreneur Center at the University of Illinois. He received the IEEE Journal of Solid State Circuits Best Paper Award (2006) and the IEEE Signal Processing Magazine Award (2008). He was selected for participation in National Academy of Engineering, “Frontiers of Engineering” symposium in 2008, and in 2009, he was elected Fellow of the IEEE “for contributions to signal processing techniques for digital communication.”

Dr. Singer is currently the IEEE SPS liaison to computing in science and engineering. He has been a member of the

Massachusetts Institute of Technology Educational Council since 1994. He has served two terms as associate editor, *IEEE Transactions on Signal Processing* (2002–2006); guest editor-in-chief, *Special Issue on Machine Learning Methods in Signal Processing*, *IEEE Transactions on Signal Processing* (2004); guest associate editor, *Special Issue on Facets of Coding Theory: From Algorithms to Networks*, *IEEE Transactions on Information Theory* (2010); guest editor-in-chief, *Special Issue on Soft Detection for Wireless Transmission*, *IEEE Journal on Special Topics in Signal Processing* (2011); member, Signal Processing Theory and Methods Technical Committee (2004–2011); and member, IEEE Machine Learning for Signal Processing Technical Committee (2007–2012).

Dr. Singer's lecture topics include “Statistical Learning in Signal Processing,” “Hybrid Signal Processing and Communication Systems,” “DSP Applications: Underwater Acoustic Signal Processing and Communication,” “DSP to the Rescue of Moore's Law,” and “Signal Processing-Based Technology Entrepreneurship.”

AKIHIKO (KEN) SUGIYAMA

Akihiko (Ken) Sugiyama is a research fellow of Information and Media Processing Laboratories at NEC Corporation, Tokyo, Japan. He received his B.E., M.E., and Dr. Eng. degrees all in electrical engineering from Tokyo Metropolitan University, Japan, in 1979, 1981, and 1998, respectively. He joined NEC Corporation in 1981 and has been engaged in research and development of signal processing algorithms and systems for transmission terminals and consumer products such as trunk switching systems, subscriber-loop transmission systems, remote conference systems, personal computers, cellphone handsets, speech/audio codecs, voice recorders, digital still cameras, camcorders, and related digital signal processor software.

He was a visiting scientist at the Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada (1987–1988), and served as a doctoral research supervisor at Traitement du Signal et Télécommunications, Université de Rennes 1, Rennes, France (1996–1998 and 2003–2006), and a part-time lecturer at

Tokyo Institute of Technology (2002–2009), Tokyo Metropolitan University (2004, 2006, and 2007), Keio University (2008), Tokyo Agriculture and Technology University (2008–2013), Kyoto University (2009), and Niigata University (2010), all in Japan.

Dr. Sugiyama was elected IEEE Fellow in 2011 “for contributions to speech and audio signal processing” and has served as member, Audio and Acoustic Signal Processing (formerly Audio and Electroacoustics) Technical Committee (1990–2009), vice chair (2010) and chair (2011–2012); conference board as a member (2005–2007) and the secretary and member-at-large (2009–2010); Industrial Relations Committee as a member (2011–present); and the organization committees of the 2001 IEEE ICASSP as the far-East liaison and a technical program chair of ICASSP 2012 as well as the 2012 IEEE International Conference on Emerging Signal Processing Applications (ESPA) as the tutorial chair. He was the vice chair (2007–2008) and the chair (2009–2010) of the Japan Chapter, IEEE SPS, and an associate editor of *IEEE Transactions on Signal Processing* (1994–1996) and is currently the vice chair of the Chapter Operations Committee, IEEE Japan Council (2013–present).

Dr. Sugiyama was elevated to Fellow of the Institute of Electronics, Information, and Communication Engineers (IEICE), a Japan-based society to cover signal processing, in 2011 “for contributions to audio and acoustic signal processing.” He has served the Signal Processing Technology Group (formerly Technical Group on Digital Signal Processing), Engineering Science Society, IEICE, as an associate secretary (1994–1995), a secretary (1995–1997), a vice chair (2007–2008), and the chair (2009) as well as the Profitability Improvement Committee, IEICE Engineering Science Committee as a member (2000–2001). He was the director, Service Activities, IEICE Engineering Sciences Society (2000–2002). He has been the local liaison officer for Japan of the EURASIP Administration Committee (2009–present). He was an associate editor of *Circuits, Systems and Signal Processing*,

(continued on page 21)

John Edwards

special **REPORTS**

Fresh Approaches Promise Wireless Quality and Reliability Improvements

With wireless technology providing the foundation for a rapidly growing number of consumer, business, and government applications, researchers worldwide are focusing on ways of resolving long-standing wireless quality and reliability issues.

Getting wireless applications to work better is not a simple or easy task. Network quality and reliability hinges on many different factors, including transmission technology, user device, signal strength, interference and terrain. Yet none of these concerns mean anything to the vast majority of wireless users, who simply expect their data, voice, and video applications to work efficiently and consistently.

With wireless use projected to skyrocket over the next few years, researchers realize that there's a pressing need for innovations that will ensure high-quality wireless links within an increasingly crowded spectrum. "There's a momentum building for systems that 'just work' and that require no effort from the user to maintain an optimal connection," says Kang Shin, a University of Michigan computer science professor and wireless network traffic management researcher. "Enhancing reliability is essential for ensuring wireless' continued growth."

"For wireless to fulfill its promise in the years ahead, users need to be confident in the quality and reliability of their applications," notes Shivendra S. Panwar, a Polytechnic Institute of New York University electrical and computer engineering professor who investigates video

streaming quality technologies (Figure 1). "This is certainly not the case today."

BETTER MOBILE VIDEO

With mobile video now challenging the dominance of conventional over-the-air and cable television broadcasting, researchers are looking for ways of ensuring that wireless content delivery systems offer quality and reliability that at least equals their predecessors.

A new architecture designed to support higher-quality video over wireless networks developed by Finland's VTT Technical Research Centre utilizes information collected from the video application, network, and mobile phone to optimize content quality. The approach allows video stream transfer speeds to be efficiently adjusted in accordance with available network resources, ensuring that quality will not fall below an acceptable level, even when the network connection's transfer speed suddenly drops due to congestion.

The problem with existing adaptive video streaming systems, according to Tiia Ojanperä, a senior scientist at VTT, is a lack of sufficient control over both video streaming and terminal mobility. "Current mobility protocols, such as mobile IP (MIP), base their handover decisions on limited information that doesn't reflect the actual transmission conditions in the networks," she says. "Such protocols also fail to take into account any application-specific requirements."

Current protocols also hide mobility from the upper layers, making it impossible for quality of service (QoS)-sensitive applications, such as video streaming, to adapt to vertical handovers. Existing application layer video adaptation solutions, including those based on the TCP Friendly Rate Control (TFRC), also have difficulties in operating over network paths with

wireless links due to their reliance on limited end-to-end feedback signaling. "The current solutions and protocols for mobile video streaming can lead to increased congestion and unnecessary data transmissions," Ojanperä says.

A cross-layer design, such as the type developed by VTT, "allows the user, network operator or service provider to influence the delivery of video services on the Internet, and especially in wireless networks, for improved video service availability and quality as well as fair distribution of network resources," Ojanperä says. "For example, linking the video stream bit rate efficiently to the available network capacity helps ensure a better functioning service in comparison to a situation where video data is randomly dropped due to congestion."

When the adaptation is done in a controlled manner, only the perceived video fidelity varies and the user can continue using the service despite network capacity changes, Ojanperä says. "In addition, since the proposed solution enables efficient use of multiple access networks, the delivery of the video service can be improved by the means of optimized handovers and simultaneous usage of multiple access networks." For instance, if the currently used network connection is incapable of supporting the QoS demands of the video service, it is possible to transfer the video stream reception to a better network or one may use multiple network connections simultaneously for added capacity.

There are other benefits as well, Ojanperä says. "For example, if a user is forced to stream a video using a network connection that is expensive to use or has limited capacity, the user can choose to only receive an adapted, poorer-quality version of the stream via that connection. If a more affordable network becomes available

Digital Object Identifier 10.1109/MSP.2013.2282793

Date of publication: 5 December 2013



[FIG1] Shivendra S. Panwar of the Polytechnic Institute of New York University is addressing poor quality wireless video streaming with a new technique called *streamloading*. (Photo courtesy of the Polytechnic Institute of New York University.)

during streaming, the architecture automatically switches to that network to give the user access to a high-quality version of the stream.”

VTT is continuing to refine the architecture, Ojanperä says. “We are currently focusing on video bit rate adaptation, handover management, and access network load balancing actions in the attempt to optimize the video quality and network resource usage.”

SPLITTING VIDEO INTO DUAL STREAMS

Panwar and his fellow researchers at the Polytechnic Institute of New York University have addressed the issue of poor-quality wireless video streaming by developing a new technique they call *streamloading*. The approach relies on a video format that splits the streaming video into a pair of layers—a base layer that contains a rough video representation, and an enhancement layer that supplies the video’s finer details.

While conventional video streaming applications work to ensure service quality with a 30- to 60-s content buffer, streamloading operates in the background, pre-downloading a content offering’s enhancement layer onto the user’s device in any location where a strong, fast wireless connection can be found, such as a home or office. This leaves only the smaller base layer to be streamed at the time of viewing, which may be a location offering only a slow or poor-quality wireless link.

Panwar estimates that the technique could remove as much as 75% of streaming content from increasingly overloaded cellular wireless networks while simultaneously slashing consumers’ high data usage charges. “Streamloading offers a ‘win’ for consumers in both quality and cost,” Panwar says.

Although users will be downloading and saving copyrighted content on their devices, Panwar maintains that streamloading is compatible with existing digital rights management protocols. Intellectual property owners have nothing to fear from the technology, Panwar says, since only a single layer of content—which is impossible to watch without the accompanying layer—is actually being stored on the user’s device. “There’s no way that streamloading could be used to pirate content,” Panwar says.

The researchers are continuing to test and refine the technology. Meanwhile, Panwar says discussions about potential applications are already being discussed with several wireless carriers, Panwar says.

BOOSTING AD HOC NETWORK RELIABILITY

Ad hoc wireless networks, widely used for sensor networks and emergency and military communications, are designed to be highly reliable under both ordinary and



[FIG2] Cal Newport, of Georgetown University, is looking to boost ad hoc network reliability. (Photo courtesy of Georgetown University.)

adverse conditions. Yet this isn’t always the case in actual deployments. Researchers at the Massachusetts Institute of Technology (MIT) and Georgetown University are working to make ad hoc networks more usable and reliable with a new framework and algorithm that are

designed to allow such networks to achieve maximal efficiency.

“When it comes to wireless ad hoc networks, there’s long been a wide gap between theory and reality,” says Cal Newport, an assistant professor of computer science at Georgetown University (Figure 2). “Theoretical algorithms tend to be unrealistic and not suitable for the real world.”

“There’s been a discrepancy between the theory, with its idealized models, and the reality of wireless networks,” says Nancy Lynch, the NEC professor of software science and engineering at MIT and head of the Theory of Distributed Systems Group, in a statement released by the school. “When people start designing theoretical algorithms, they tend to rely too heavily on the specific assumptions of the models. So the algorithms tend to be unrealistic and fragile.”

Newport notes that the models theoreticians use to design algorithms for ad hoc networks have straightforward designs. “Because real networks are complex, the models we use to design algorithms for them need to be more complex,” Newport says.

To bring anticipated ad hoc wireless network performance closer to reality, the researchers set out to create a new model that would better represent real-world scenarios. As they began their work, they decided to tap into the potential of randomness. “In modeling wireless networks we should use more uncertainty,” Newport states.

Newport notes that earlier researchers focused their network link reliability models on random fluctuations. Yet these efforts failed to deliver accurate real world results, Newport says. The team tried another approach: modeling quality fluctuations as the intentional actions of an artificial adversary. Such an entity isn’t able to maintain control over all network links, and some will remain up throughout the execution of the communication algorithm. Yet the adversary is also able to alter the bandwidth of other network connections at will, and the network designer has no way of knowing beforehand which links are likely to be reliable and which will fail.

“An algorithm must work for all potential adversaries, some of which are harmless and some of which are not,” Newport says. “A successful algorithm must be able to cope with all possible strategies for controlling the network.”

As they started their project, the researchers created an adversary that understands how the communications algorithm works and has the ability to try defeating it. To achieve victory, however, the adversary would have to be able to determine a pattern of link manipulations in advance, before the algorithm even starts running. The adversary the researchers developed had the potential to be much more disruptive than the types of interference that real-world wireless networks are likely to encounter.

The researchers next examined two kinds of message dissemination. In the first type—a single network node trying to broadcast a message to all other nodes—they discovered that efficient communication is possible, even in the adversary’s presence. In the second case—with several nodes each transmitting messages, and every one of their immediate neighbors required to receive a message from at least one transmitter—turned out to be more problematic.

In the second scenario, the researchers discovered that the adversary’s presence has the potential to derail efficient communication, yet only if the network has an odd shape in which a central node is connected to many nearby nodes that aren’t connected to each other. Fortunately, this type of ad hoc network topology isn’t likely to exist in the real world since, if two wireless devices are close enough to a third to communicate with it, they’re likely to be able to communicate with each other, too. After the researchers added another assumption—that two devices connected to a third unit will at least sometimes be able to establish links with each other, too—efficient communication again became possible.

With both types of message dissemination, the researchers’ communication algorithms were able to thwart the adversary by using randomness. Newport observes that the biggest challenge facing designers of communications protocols

for ad hoc wireless networks is interference. If two nearby nodes begin transmitting at the same time on the same frequency, they can interfere with each other, preventing either transmission from being received. A protocol, therefore, should assign each node a probability of transmitting during any one round of communication (where a round is defined by the time it takes for a node to send a message to its immediate neighbors).

The researchers’ algorithms follow this rule fundamental to many medium access algorithms. Yet, instead of simply cycling through a prescribed sequence of steadily shrinking probabilities, the algorithms jumble the sequence. Therefore, during a local broadcast, each separate message possesses its own unique sequence of probabilities. Meanwhile, clusters of nodes also temporarily elect local leaders that coordinate the probabilities for different transmitters.

Newport notes that the new approach doesn’t appear to have any meaningful negative impact on computation or network speed. “While reliability improved, we didn’t observe any major effect on performance,” he says.

CURBING SIGNAL CONFLICTS

Software being developed at the University of Michigan aims to improve wireless network traffic management while reducing the interference created by conflicting signals. GapSense, as the software is called, allows Wi-Fi, Bluetooth, or ZigBee devices, which ordinarily can’t communicate with each other, to swap simple stop

and warning messages, potentially reducing the likelihood of signal collisions.

GapSense uses a common language of energy pulses and gaps. The length of each gap creates a “stop” or “warning” message. Devices equipped with the software send the messages at the beginning of a transmission, or in between data packets, to inform other nearby devices about their plans to use a particular slice of spectrum

Bluetooth and ZigBee devices operate in the same spectrum as Wi-Fi, yet speak completely different languages. All of these systems are already equipped with the standard carrier sense multiple access protocol, which enables them to listen for radio silence before initiating their own transmissions. Yet this approach hasn’t proven to be completely reliable. “All of these gadgets are using the same frequencies and fighting for space,” says project leader Kang Shin (Figure 3). “Since they don’t have a direct means of communicating with each other, because they use different and noncompatible protocols, we wanted to give them a way of coordinating their transmissions so that each device can perform its designed functions while minimizing interference to the other systems.”

Current technology design features make some level of interference virtually inevitable. ZigBee, for instance, takes 16 times longer than Wi-Fi to rouse itself from an idle state to transmit information. “Sometimes,” Shin says, “Wi-Fi believes that a channel is clear when a ZigBee packet is already being sent over it.” Also, ZigBee’s transmit power is 100 times lower than Wi-Fi’s output. Shin compares the situation to a loud person talking over someone who speaks softly. “The little guy is saying something, but the big guy doesn’t hear it,” he says. “So the little guy’s information gets shouted down.”

The researchers claim that GapSense, operating in a simulated office environment, can reduce interference by more than 88% on busy networks comprising different types of devices. Running a moderate amount of Wi-Fi traffic, the researchers detected a 45% collision rate between ZigBee and WiFi, which GapSense was able to reduce to 8%.

Shin notes that the software could also address the so-called “hidden terminal”



[FIG3] Kang Shin of the University of Michigan is working to curb signal conflicts. (Photo courtesy of the University of Michigan.)

[special **REPORTS**] continued

problem. Newer Wi-Fi standards support faster data rates on wider bandwidths than the current standard of 20 MHz. Yet devices utilizing different bandwidths can't hear one another's communications to avoid talking over them. GapSense promises to allow devices on different standards to talk in an orderly sequence. At moderate Wi-Fi traffic levels, the researchers detected around a 40% collision rate between

wider- and narrower-bandwidth devices, which GapSense was able to lower to virtually zero.

GapSense could also reduce Wi-Fi device energy consumption by 44%, Shin says. The software would accomplish this feat by allowing the Wi-Fi receiver to operate at low clock rates. Using the new software, a faster-clocked Wi-Fi transmitter could send a wake-up message to the slower-clocked receiver in time for it to

sync and snare an incoming data packet. "The impact of GapSense is potentially huge," Shin states. "It could be the solution that creates order out of chaos within an increasingly diversified world of wireless devices."

AUTHOR

John Edwards (jedwards@johnedwardsmedia.com) is a technology writer based in the Phoenix area.

Ron Schneiderman

Automotive Industry Is a Key Component to the Success of the DSP Sector

The automotive industry is going strong with sales at near record levels in the United States and with technology vendors helping to drive the market. According to a recent Scotiabank Global Auto Report, each car and light truck built in North America (the United States, Canada, and Mexico) now contains US\$3,200 in electronic equipment, and that's expected to increase 7% a year through this decade, providing opportunities for developing innovation bumper-to-bumper.

Significantly, industry analysts are reporting that signal processing in today's vehicles is increasing rapidly, almost from model to new model.

For this new editorial feature for *IEEE Signal Processing Magazine (SPM)*, in July 2013 we talked to three key members of the Texas Instruments (TI) automotive business sector, Robert Tolbert III (Figure 1), product marketing/business development manager of automotive infotainment processors,

Brooke Williams (Figure 2), business and marketing manager for TI's advanced driver assistance systems (ADAS), and Ian Sherlock (Figure 3), Wi-Fi product manager.

IEEE SPM: Advancements in technology seem to have shortened the design cycle for both electronics and automobile manufacturers. How has this impacted TI?

Brooke Williams: TI is certainly seeing a much harder push to put advanced

technology in vehicles in faster cycle time. The carmakers are finding themselves under extreme pressure to advance the cycle and speed up the adoption of the advanced technology that we have in the consumer world in the vehicle.

One good example is infotainment systems. We see infotainment systems catching up and adopting new technology in vehicles. For TI, it's a great trend because it advances the time from the design-in to actually generating revenues.

Robert Tolbert: You will find that original equipment manufacturers (OEMs) will lean more heavily on companies that they know are able to deliver year after year, and that's why TI feels so good about this acceleration of the time line. We have been doing automotive from analog processors to embedded processors for more than 30 years.

IEEE SPM: More specifically, can we expect anything new from you in this area any time soon?

Tolbert: We announced a new "Jacinto 6" automotive infotainment processor [DRA74x] at the International Consumer Electronics Show in Las



[FIG1] Robert Tolbert III is the product marketing/business development manager for automotive infotainment processors at Texas Instruments. (Photo courtesy of Texas Instruments.)

Digital Object Identifier 10.1109/MSP.2013.2282794

Date of publication: 5 December 2013

Vegas in January. It has been much anticipated from a customer perspective.

One of the trends we're seeing in infotainment automotive processors is a bigger need to drive higher-resolution displays. It's the industry's first dual Cortex 815-based processor that's designed for automotive. We'll be talking about this more in the future in terms of availability, as we get closer to sampling the product in the middle of this year. As far as in what model vehicle this will initially appear, we're readying ourselves for calendar year 2015.

IEEE SPM: Two recently published market research studies report that the use of digital signal processors (DSPs) in the automotive industry is increasing rapidly over time, but they also point out that the DSP market is witnessing the increasing use of field-programmable gate array (FPGA)-based DSP equipment. Can you talk about that, and where signal processing continues to fit into the automotive sector?

Tolbert: We see a lot of opportunities for DSP in automobiles. Examples would be infotainment, telematics, and in ADAS applications. In the areas specific to telematics and infotainment, you will see DSPs being leveraged in four areas: one is to do the audio pre- and postprocessing for things like noise suppression, echo cancellation, and equalization. That's an area where TI DSPs are used quite heavily. The second area would be the actual audio encode and decode. The third area that it lends itself very well to is the voice application, particularly the voice recognition application. The fourth area is software-defined radios. This would be for high-definition radio, playback, and digital audio broadcasting.

Williams: ADAS is a very exciting market. One reason is that there are multiple cameras and radar systems in each vehicle, and we continue to see a trend in ADAS moving in luxury cars down to more midline cars and even into some entry-level cars. The pace of that is accelerating. So, this market is truly exploding.

To drive home the point of multiple sockets in each vehicle, you can buy a luxury car today that has up to five TI DSPs in the ADAS alone. So, DSPs are

applicable to the front cameras, which are running anywhere from your entry-level lane departure, traffic sign recognition, all the way up to higher-level applications like pedestrian detection. This requires a tremendous amount of signal processing to run those algorithms simultaneously on a single processor.

A second opportunity is rear cameras. We're starting to see the emergence of smart rear cameras where the camera has ADAS algorithms that will warn the driver if a child or an object is behind the car. It will warn you if you're not looking at the screen, and we're starting to see applications that will actually take control of the vehicle and automatically brake. This is another opportunity for DSP to run those ADAS algorithms.

There are also surround-view cameras. Typically, it's four cameras—one front, one in the rear, and two on the sides. The images are fused together to give the driver a kind of top-down view. We're starting to see more of a three-dimensional, rather than a two-dimensional top-down view where you can literally pan around your vehicle.

Then there's radar. The trend will be for more surround-view radar. And applications are emerging for night vision with pedestrian detection or animal detection. That would require another camera. The next step would be the fusion of radar and cameras. For a central decision unit, that is, taking preprocessed data from a radar system and a camera system, using that as redundant information, and then getting a more intelligent decision on control of the vehicle.

And we're moving on to autonomous driving. We have the technology to do that; Google has a car operating autonomously and others are working on this. It's not yet affordable, but we talked about the acceleration of the technology, and the amount of signal processing horsepower we can deliver. So, we're at the early stages of autonomous driving.

IEEE SPM: Do you see any product differentiation in the technology that is being introduced into automobiles? Is this even important to your automotive manufacturing customers? Are we seeing



[FIG2] Brooke Williams is the ADAS business and marketing manager at Texas Instruments. (Photo courtesy of Texas Instruments.)

the same apps with perhaps a different "look" or interface?

Tolbert: Differentiation is extremely important to our immediate customers, Tier 1s and the OEMs. Tier 1s want to offer unique applications and end-user experiences to the OEMs. The OEMs, in turn, want to enable the driver and passengers to have a driving/riding experience that is unique to that automobile manufacturer. Ultimately, the differentiation leads to familiarity, which holds the promise of repeat buyers.

The vehicle infotainment system also offers lots of potential for differentiation. It can be as simple as instant access to your navigation system, audio/video, or rear-view camera system when you enter the vehicle, and reducing the wait time it takes to start your driving experience to more complex things like sign or street marking recognition to improving your global positioning system experience. The look, feel, and way you interact with your infotainment system can also provide a form of differentiation, as OEMs and drivers are always looking to reduce complexity and make your in-vehicle experience as intuitive as possible.

IEEE SPM: Are DSPs playing any direct role in improving fuel economy and emissions in vehicles? In making vehicles greener?

Sherlock: The vehicular environment poses many electromagnetic compatibility challenges. There are already

special **REPORTS** continued

[FIG3] Ian Sherlock is the Wi-Fi product manager at Texas Instruments. (Photo courtesy of Texas Instruments.)

many sources of RF noise within the vehicle, including intentional RF emitters, as well as motors and electronic subsystems that generate broadband RF noise. Because of this, our wireless connectivity products have been designed for colocation with the other RF subsystems, for example, the integration on a printed circuit board with cellular interfaces.

IEEE SPM: Is there an aftermarket opportunity in automobiles for DSP technology? If so, how and where, and are traditional aftermarket vendors taking advantage of it?

Tolbert: There is an aftermarket opportunity for DSP technology, particularly in the vehicle center console and rear seat entertainment systems. The DSP technology is being leveraged for audio/video playback, voice recognition, digital radio, and the like. China is a very large market for aftermarket infotainment systems, but volume is flattening versus the traditionally manufacturer installed systems. Aftermarket infotainment center consoles are a combination of resale models and dealer optional install models. A couple of players in the China market are ForYou, FlyAudio, HASE, and CarIT.

IEEE SPM: We talk mostly about cars. Do commercial vehicles have any special tech requirements and are those being addressed?

Tolbert: Commercial vehicles are similar in their usage of DSP technology, especially in the areas of speech recognition,

audio processing to handle noise cancellation/reduction, and vision analytics for surround-view camera capture and object recognition to assist the driver.

IEEE SPM: Are there currently any spectrum issues that could impact auto (and ultimately chip) development and design?

Tolbert: I don't see any spectrum issues on the horizon that impact chip development for infotainment processors.

IEEE SPM: As engineers, but also consumers, do you think there's enough attention being given to ease-of-use for drivers?

Tolbert: I think automobile technologists are paying a great deal of attention and spending lots of R&D dollars on improving the in-vehicle experience and ease of use. The more intuitive and enjoyable the in-vehicle experience, the higher the chance of a repeat buyer. And it lowers the driver distraction, which is extremely important to manufacturers and consumers.

As technology providers, we sit at the forefront of automobile manufacturers question to make driving simpler, safer, and more fun. Our role is to provide the horsepower and tools to enable things like autonomous driving, which will require an inordinate amount of complex computations and instantaneous data manipulations, drowsy driver detection, special/visual recognition, and more.

IEEE SPM: What's the status of eCall, the automotive emergency call system for which TI has introduced a reference design? (eCall systems will automatically place calls to an emergency service center in the event of an accident.)

Tolbert: It's going to be legislation driven, completely from a European perspective. That's where the initiative comes from. Our reference design is ultimately going to be leveraged by a lot of our customers in eCall solutions. The reference design itself brings together analog and embedded processing ICs that we have in our portfolio into a single platform to help drive this solution. (Legislation is expected to make eCall mandatory in Europe beginning in 2015.)

IEEE SPM: Do you have any problems dealing with different rules and

regulations, and even design cycles, in working with vehicle manufacturers in different countries?

Tolbert: I can't speak specifically from a design perspective, but we want to cover as much ground as we possibly can. When there are different standards and different requirements country to country, we have to work those out. We want to make sure that we can meet all of the requirements from different regions.

Williams: Most of the different standards are fairly transparent to TI. We just want to make sure we enable our customers to respond to those standards. Typically, they're the ones responsible for addressing the standards. In ADAS, there are no standards today in terms of how you implement an algorithm or a function. We are starting to see the emergence of a safety rating that requires autonomous breaking, and that is starting to form a framework around some standardization. But we're still a long way from having any standard algorithm implementation in ADAS.

IEEE SPM: There is an IEEE standard, IEEE 1616a, for a so-called "black box" or motor vehicle electronic data recorder (MVEDR), that dictates what information has to be captured in the event of an accident. Where do we stand with the development and use of this type of device in motor vehicles?

Williams: We are seeing an increasing level of data recording being done in motor vehicles. We offer solutions in that space and from an ADAS business unit perspective. We are seeing a combination to adding black box recording into some of these ADAS cameras. So, you not only can record vehicle data and information, but you can record the video data as well on a system that's integrated into the vehicle.

People are installing aftermarket cameras in their vehicles to do video black box recording. As cameras become more popular and more pervasive in the market, it doesn't make sense to have two separate systems. Carmakers are starting to spend more time integrating black box recording functionality into the ADAS camera.

IEEE SPM: There's a lot of interest and coverage now of the so-called "connected" car. How connected can it get?

Williams: We see a role for ADAS in vehicle-to-vehicle and vehicle-to-infrastructure. The U.S. Federal Communications Commission has allocated a band at 5.9 GHz for short-range communications. It's currently under test by the U.S. Department of Transportation (DoT) and other countries around the world to see if it's a feasible communications methodology that can work from car-to-car. This would cover things like, I'm in your blind spot, or there's an accident a half mile ahead, or there's fog ahead; so it could be real-time emergency information.

In terms of vehicle-to-infrastructure, that could cover road conditions, or any

general information related to the road for vehicle guidance. We'll hear more about this after the DoT completes its tests this year. It's an interesting new opportunity for TI.

IEEE SPM: There has been some media coverage suggesting concern about how "connected" cars can be hacked. Is this possible and, if so, is it being addressed?

Tolbert: Car hacking is a growing concern for automakers. The introduction of the "connected" car adds yet another means for hackers to gain access to critical and noncritical resources in the vehicle. Cloud and network connectivity can provide a means for hackers to attack the vehicle through malicious or nuisance software that makes its way into

the vehicle via downloadable applications, rooting the open operating system, or over-the-air updates of vehicle software. Automakers are trying to combat and neutralize the threat posed by hackers in various ways, including physically separating the real-time safety critical components of the vehicle network from the nonreal-time safety critical components, like infotainment, into separate ECUs in the vehicle, fire walling the vehicle controller area network bus, or trying to contain the threat in a secure environment on the main system-on-a-chip of the infotainment system.

Editor's Note: This interview was conducted by Ron Schneiderman, a regular contributor to *SPM*.



[society **NEWS**] (continued from page 14)

Birkhäuser (1999–2011), an associate editor of *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Sciences* (2000–2002), and a guest associate editors for six special sections of the same transaction (1994–1999) as well as the guest editor-in-chief for the special section on signal processing of the same transaction (2009).

He was a member of Japanese delegation to ISO/IEC JTC1/SC29/WG11 for MPEG Audio Standardization (1990–1994, 1998, 2002, 2007–2008), and the interim chair, Audio Subgroup, ISO/IEC JTC1/SC29/WG11 at Angra dos Reis Meeting, Brazil (1992). He also served as the secretary (1991–1996) of SC29/WG11 Audio Subcommittee and a member of SC29 Committee (1992–1996), both of the Information Technology Standards Commission of Japan, Information Processing Society of Japan. He has drafted four Japanese industrial standards, which cover MPEG-1, MPEG-2/BC, MPEG-2/AAC, and MPEG-4 as the chair of the Audio Subworking Group (SWG3: 1993–1995, SWG4: 1995–1996, 2001–2003), a member of the MPEG JIS Standardization Working Group (WG8: 1993–1998),

MPEG-1 JIS Standardization Subworking Group (SWG1: 1996–1998), MPEG-2 JIS Standardization Subworking Group (SWG2: 1996–1998), MPEG JIS Standardization Working Group (WG6: 1998–2001), and MPEG JIS Standardization Working Group (WG5: 2001–2003), all of the Image Processing Technology and Standardization and Research Committee, Image Processing Technology Standardization Center, Japanese Standards Association.

Dr. Sugiyama is a recipient of the Shinohara Memorial Academic Encouragement Award from IEICE (1987), Ohm Technology Award from the Promotion Foundation for Electrical Science and Engineering (2001 and 2013), the Best Paper Award from IEICE (2002), the Incentive Award from the Japanese Society of Artificial Intelligence (2005), the Achievement Award from IEICE (2006), the Sankei Newspaper Award of Fuji-Sankei Business I, Advanced Technology Award from Nikkan Kogyo Shinbun Ltd. (2010), and local commendation for invention in the Kanto region from the Japan Institute of Invention and Innovation (2011).

Dr. Sugiyama's current research interests are in signal processing and its applications to commercial products and services. He is an author of 15 book chapters on audio and acoustic signal processing in Japanese and English with one Korean translation. He has published more than 100 technical papers in journals and at international conferences and is the inventor or coinventor of 150 registered patents in Japan, the United States, Canada, Australia, and the European Patent Convention with more pending.

His lecture topics include "Interference and Noise Control for Industry Applications," "25 Years of Audio Coding: Technology, Standardization, Competition, Collaboration, and Compromise," "Multichannel Echo Cancellation: Discovery of the Uniqueness Problem and Development of Solutions," "Efficient and Fruitful Collaboration with Students and Junior Engineers: 28-Year Experiences as a Supervisor of 70+ International/Domestic Internship Students and Master/Ph.D. Candidates," and "What I Wish I Knew When I Was an Entry-Level Engineer."



[from the **GUEST EDITORS**]Wing-Kin Ma, José M. Bioucas-Dias,
Jocelyn Chanussot, and Paul Gader

Signal and Image Processing in Hyperspectral Remote Sensing

In recent years, it has become clear that hyperspectral imaging has formed a core area within the geoscience and remote sensing community. Armed with advanced optical sensing technology, hyperspectral imaging offers high spectral resolution—a hyperspectral image can contain more than 200 spectral channels (rather than a few channels as in multispectral images), covering visible and near-infrared wavelengths at a resolution of about 10 nm. The

result, on one hand, is significant expansion in data sizes. A captured scene can easily take 100 MB, or more. On the other hand, the vastly increased spectral information content available in hyperspectral images (or large spectral degrees of freedom in signal processing languages) creates a unique opportunity that may have previously been seen as impossible in multispectral remote sensing. We can detect difficult targets, for example, those appearing at a subpixel level. We can perform image classification with greatly improved accuracy. We can also identify underlying materials in a captured scene without prior information of the materials to be encountered, by carrying out blind unmixing.

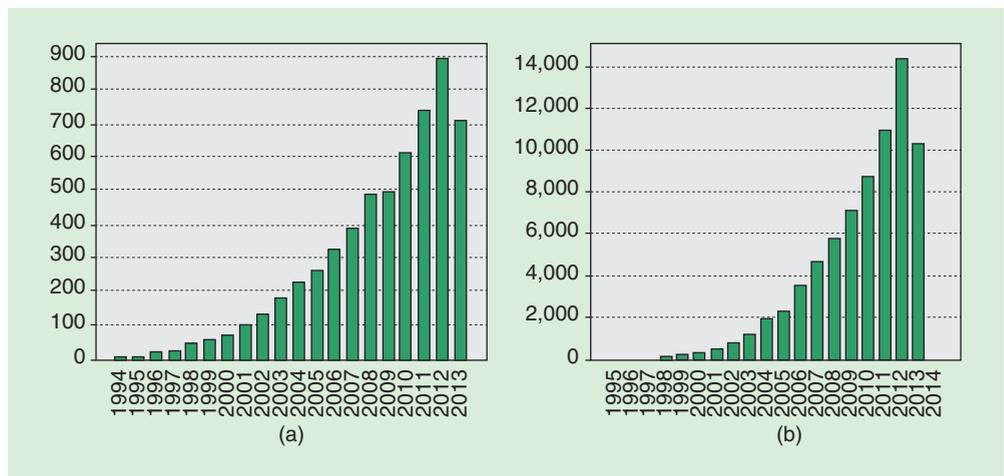
There are many other exciting advances contributed by researchers in hyperspectral remote sensing, and their great effort has resulted in an enormous number of applications, such as

surveillance, reconnaissance, environment monitoring, land-cover mapping, and mineral identification, just to name a few. Hyperspectral imaging is also a key technique for planetary exploration, astrophysics, and nonremote sensing problems such as food inspection and forensics.

There has been much growth in research activities related to hyperspectral imaging lately. Figure 1 shows a report on the number of publications and citations in the “hyperspectral” topic. The results were obtained by searching the Science Citation Index (SCI)-Expanded database of the ISI Web of Science with the topic “hyperspectral” from 1994 to September 2013. A sharp rise with both the publications and citations counts can be observed from 2010 to 2013. While major research activities on hyperspectral remote sensing are in the geoscience and remote sensing community, hyperspectral remote sensing is also an area that contains many interesting and important signal processing problems. In fact, this area has attracted

growing attention and contributions from different communities, such as signal processing, image processing, machine learning, and optimization—and this is what motivates us to organize this special issue.

IEEE Signal Processing Magazine published a special issue on signal processing for hyperspectral image exploitation in 2002, which was particularly relevant at the time. After more than ten years, we believe that now would be an appropriate time to consider another special issue on this topic, chronicling recent advances, challenges, and opportunities. Also, this issue has a unique theme—to provide a balanced collection of tutorial-style articles that introduce prominent and frontier signal processing topics in hyperspectral remote sensing and demonstrate the insight and uniqueness of signal processing techniques established in those topics. We also intend to take this opportunity to bridge the gap between remote sensing and signal processing by showing readers a



[FIG1] The number of published papers having the keyword “hyperspectral” and the corresponding citations. Data is obtained from the SCI-Expanded database, ISI Web of Science. (a) Published items in each year. (b) Citations in each year.

sample of relevant problems in hyperspectral remote sensing.

We would like to thank those who showed interest in this special issue. We received approximately 40 white papers. The topics proposed are very diverse from one another, and many of them are indeed interesting in their own rights: we have seen numerous excellent white papers, and in some cases, we are comparing apples and oranges. However, there are page limitations, and consequently only nine articles can be accommodated. Again, we appreciate the enthusiasm received.

The special issue can roughly be divided into four theme topics: detection, classification, unmixing, and compressive sensing (CS). It begins with the detection topic. Manolakis et al. give an overview on the hyperspectral target detection problem. The authors then show that some state-of-the-art detectors can in fact provide consistently good performance for practically relevant applications by resorting to classical detection theory and physics-based signal models. Performance analysis is presented to support the authors' claims.

Next, Nasrabadi explores the detection topic further by looking into recent advances in hyperspectral target detection techniques. In particular, Nasrabadi's contribution highlights novel detection techniques based on concepts in statistical signal processing and machine-learning theory, such as subspace-based detectors, the support vector machine, kernel-based nonlinear detectors, fusion of detectors, and sparsity-based detectors.

The third article considers the classification topic. Classification in hyperspectral images is far from being a generic image classification problem; it is challenging owing to the high dimensionality of data, few training samples, nonlinearity, and a number of other factors. Camps-Valls et al. overview the topic by presenting a statistical learning theory (SLT) framework for hyperspectral image classification. Under the SLT framework, the article covers techniques such as standard regularization; active, semisupervised, and sparse learning

approaches; spatial-spectral regularization; and adaptation of classifiers and feature representations.

Nonlinear manifold learning is another promising framework for hyperspectral image classification, and it has also received much attention. In this framework, the topology of high-dimensional nonlinear data sets is represented in lower, but still meaningful, dimensions for classification or other purposes. Lunga et al. provide an overview on this representative research direction. The article reviews traditional approaches under a graph embedding framework and describes new techniques for modeling hyperspectral data on manifolds, such as multidimensional artificial field embedding and spherical stochastic neighbor embedding.

The next three articles are related to the unmixing topic. Ma et al. overview blind (or unsupervised) hyperspectral unmixing techniques under the linear mixing model (LMM) setting. It is worthwhile to mention that this blind problem from remote sensing has a strong connection to blind source separation and sensor array processing in signal processing. The authors select four significant blind unmixing approaches—pure pixel search, convex geometry, sparse regression, and nonnegative matrix factorization—and use a signal processing researcher's view to describe each approach and appreciate the methodological beauty within.

The LMM is not always valid in the real world. Recently there has been much interest in unmixing based on nonlinear models. Dobigeon et al. present an overview of recent advances dealing with the nonlinear unmixing problem. Representative nonlinear models, such as intimate mixtures, bilinear models, and postnonlinear mixing models, are presented and their validity discussed. Then, the main classes of unmixing strategies, in supervised and unsupervised frameworks, are described. The article also addresses an emerging subtopic—detecting nonlinear mixtures in hyperspectral images.

In the unmixing topic, most models assume that the endmember signatures are invariant across the whole image. This assumption can be violated in reality,

owing to various reasons related to measurement and environment. In Zare and Ho's article, the authors review a representative set of methods designed to cope with endmember variability. The methods are organized in two classes: 1) endmember sets and 2) endmember as statistical distributions. The former class is nonparametric and deterministic, while the latter class stochastic. The article reviews important methods in both classes and highlights their advantages, limitations, and challenges.

The last two articles describes a relatively new front—CS for hyperspectral images. This is a well-motivated topic since hyperspectral data, in their raw form, are often tremendous in size. Arce et al.'s article is an overview of the fundamental optical phenomena behind compressive spectral imaging sensors. It describes the mathematical concepts and optimization framework for designing optimal coded apertures (i.e., measurements) in hyperspectral image reconstruction, spectral selectivity, and superresolution. All of these ideas and concepts are concretized in a specific type of spectral imagers known as coded aperture snapshot spectral imagers (CASSI). Many practical aspects are described and illustrated with real data and imagery.

The last article, by Willett et al., provides a fundamental overview on how CS can make a difference in the hyperspectral context. It describes how novel sparse models enable the design of new hyperspectral imaging hardware and acquisition methods. Performance limits and tradeoffs arising from practical issues, such as noise, quantization, and dynamic range, are discussed. The authors also consider hyperspectral target detection using CS measurements without having to reconstruct the raw hyperspectral data.

ACKNOWLEDGMENTS

We would like to express our gratitude to the contributing authors and the anonymous reviewers, whose contributions play a key role in making this special issue possible.

SP

[Dimitris Manolakis, Eric Truslow, Michael Pieper, Thomas Cooley, and Michael Brueggeman]

Detection Algorithms in Hyperspectral Imaging Systems



[An overview of practical algorithms]

Hyperspectral imaging applications are many and span civil, environmental, and military needs. Typical examples include the detection of specific terrain features and vegetation, mineral, or soil types for resource management; detecting and characterizing materials, surfaces, or paints; the detection of man-made materials in natural backgrounds for the purpose of search and rescue; the detection of specific plant species for the purposes of counter narcotics; and the detection of military vehicles for the purpose of defense and intelligence. The objective of this article

is to provide a tutorial overview of detection algorithms used in current hyperspectral imaging systems that operate in the reflective part of the spectrum ($0.4\text{--}2.4\ \mu\text{m}$). The same algorithms might be used in the long-wave infrared spectrum; however, the phenomenology is quite different. The covered topics and the presentation style have been chosen to illustrate the strong couplings among the underlying phenomenology, the theoretical framework for algorithm development and analysis, and the requirements of practical applications.

INTRODUCTION

In hyperspectral target detection applications, we seek to determine whether a rare object with a known spectral signature is

Digital Object Identifier 10.1109/MSP.2013.2278915

Date of publication: 5 December 2013

present or not in the imaged scene. The term *rare* is used here to signify a relatively small number compared to the total number of pixels, e.g., a few pixels in a million pixel image.

The processing methods developed for landscape classification over the last few decades are not applicable to target detection for two reasons. First, even if there are targets in the scene, their number is typically too small to support estimation of statistical properties of the target class from training data. Second, depending on the spatial resolution of the sensor, targets of interest may not be clearly resolved and hence may appear in only a few pixels or even as part of a single pixel (subpixel target). The isolated character or absence of targets from the imaged scene makes verification by means of clustering of like samples problematic. However, as we explain in this article, target detection is still possible in these situations by developing algorithms using classical detection theory and physics-based signal models. However, even the best detection algorithms exhibit a large number of false alarms; therefore, in practical systems, detected targets are further analyzed by the false alarm mitigation (FAM) and target identification (TID) algorithms. A large number of hyperspectral detection algorithms have been developed and tried over the last two decades. The reviews in [1] and [2] cover developments up to 2002; more recent developments are reviewed in [3].

The purpose of this tutorial review article is threefold. First, we introduce the physical principles that determine the nature of target and background clutter spectra measured by a hyperspectral sensor. Second, we derive the detectors currently used in practical applications, investigate their properties and performance, and discuss their practical implementation. Third, we discuss the need for FAM and TID algorithms and review some recent work in these areas. Whenever possible, we take the opportunity to mention areas where further understanding of phenomenology or better algorithms would be beneficial.

AT-SENSOR RADIANCE SIGNAL MODEL

We start with a simple discussion of the radiation components in the reflective spectral region. As shown in Figure 1, the total radiation signal reaching the sensor consists of three components:

- 1) the radiation L_1 reflected from the pixel of interest, i.e., the direct and diffused solar radiation incident on the pixel surface and reflected directly into the sensor (reflected radiance)
- 2) the radiation L_2 reflected from the surface surrounding the pixel of interest and scattered by the air volume into the sensor (adjacency radiance)
- 3) the path radiance L_3 , i.e., the photons scattered into the sensor's field of view, without ground contact.

Only the reflected radiance component L_1 contains information about the pixel of interest. The task of atmospheric correction is the calculation and removal of components L_2 and L_3 and the retrieval of the pixel reflectance from component L_1 .

Although the interactions of the radiance terms are not independent, under many atmospheric conditions and wavelengths the total spectral radiance $L(\lambda_k)$ received at the sensor can be well approximated by the linear relation [4]

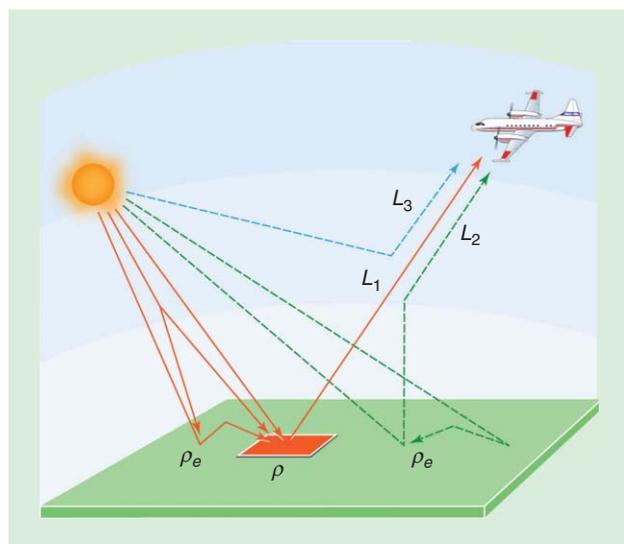
$$L(\lambda_k) = c(\lambda_k)\rho(\lambda_k) + d(\lambda_k), \quad (1)$$

where λ_k is the central wavelength of the k th spectral channel and $\rho(\lambda_k)$ is the pixel surface reflectance. The first term in (1) corresponds to L_1 ; the second term includes L_2 and L_3 .

RADIANCE AND REFLECTANCE DOMAINS

Hyperspectral imaging sensors have a sufficient number of spectral bands to allow extraction of spectra that closely resemble those acquired by laboratory spectrometers. This makes possible, at least in principle, the use of such spectra to identify materials based upon their reflectance properties and provides the basis for most hyperspectral imaging applications. Several libraries of high-resolution reflectance spectra of natural and man-made materials have been compiled by various organizations and are available for public use. We note that quite often the spectral signatures provided for various materials represent “averages” or “typical” examples. To use these library spectra, the raw sensor data must go through a series of processing steps to convert it to a usable format.

The first step in the conversion process is to calibrate the raw sensor data to the physical units of radiance. Radiance quantifies the energy at a given time passing through the input aperture of the sensor within a certain angular extent (solid angle) as a function of direction. The second step is the application of physical radiative transfer models to convert the radiance data to reflectance. This is known as atmospheric compensation. These models account for atmospheric transmission and the atmosphere's thermal emission. The difficulty with radiative transfer models lies in needing a rather detailed understanding of the atmospheric constituents. For example, the number and distribution of the small suspended particles, called aerosols, are typically unknown and often difficult to estimate. Alternatively, there are in-scene methods that assume the presence of materials with known spectral signatures [5] to enable the retrieval of the desired reflectance signature.



[FIG1] The physics-based model for at-sensor measured radiance from a Lambertian ground pixel surface.

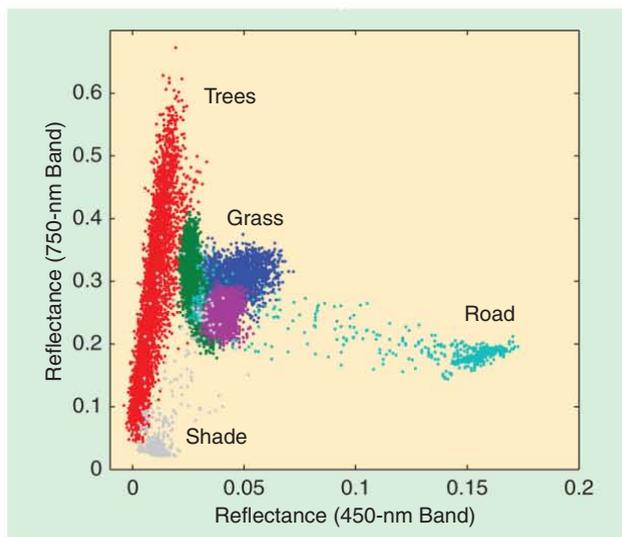


[FIG2] Areas of interest for different ground covers: trees, different types of grass, a road, and a small area dominated by shade.

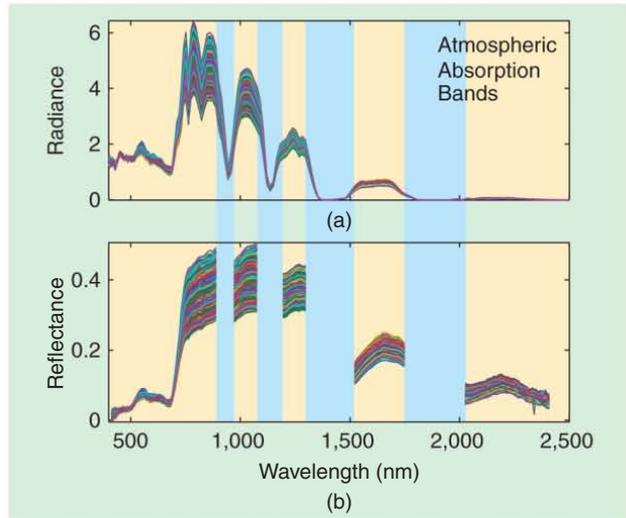
The challenge of atmospheric compensation can be understood by considering the spectra in the rectangular area enclosed by the white frame in Figure 2. Figure 3(a) shows radiance spectra for all pixels of this green area as measured by a calibrated sensor. Figure 3(b) shows the reflectance spectra after in-scene atmospheric compensation. It is obvious that the two graphs for each material appear significantly different and that the rich spectral content introduced by the atmosphere dominates the spectral radiance signatures.

SPECTRAL VARIABILITY

Figure 4 shows a scatter plot for the ground cover areas shown in Figure 2. We note that measured spectra corresponding to pixels with the same surface type exhibit a random variation



[FIG4] The scatter plot for the ground cover areas shown in Figure 2. Note that the grass class includes three different overlapping types of grass.



[FIG3] (a) and (b) Radiance and reflectance spectra retrieved from radiance extracted from the green grass area located on the left of the road in Figure 2 (white rectangular frame).

(spectral variability) that prevents the characterization of homogeneous surface materials by unique spectral signatures. The pixel-to-pixel spectral variability (background clutter) is measured by the signal-to-clutter ratio (SCR), which typically dominates sensor noise. Hence, detection algorithms primarily focus on clutter suppression.

The mathematical representation of the “spectral-swarms” in Figure 4 provides the basis for the development of hyperspectral exploitation algorithms. The statistical approach models the data clouds using probability distributions; the geometrical approach typically uses linear subspaces [2]. The geometrical approach can be justified by the linear mixing model, which states that when a pixel contains several spatially distinct materials, its spectrum is given by

$$x = a_1s_1 + \dots + a_ms_m + e = Sa + e \quad (2)$$

$$a_k \geq 0, \quad a_1 + \dots + a_m = 1, \quad (3)$$

where s_k is a “pure” spectrum, called *endmember*, a_k is the fill-fraction of each endmember, and e is the residual that accounts for the difference between the measured and modeled spectra. The mixing component Sa in (2) lies in an m -dimensional subspace of \mathbb{R}^p . The constraints (3), which ensure physically meaningful values for the fill-fractions, force the mixed spectra to lie on a simplex in \mathbb{R}^p . The enforcement of the nonnegativity constraint $a_k \geq 0$ introduces a major complication in the analysis and application of the linear mixing model.

STATISTICAL MODELING OF BACKGROUND CLUTTER

Inspection of several density scatter plots, like the one in Figure 4, suggests that a plausible model for the PDF of hyperspectral imaging data is the density mixture

$$f(x) = \sum_{k=1}^{N_c} \pi_k f_k(x), \quad \pi_k \geq 0, \quad \sum_{k=1}^{N_c} \pi_k = 1, \quad (4)$$

where π_k is the a priori probability, $f_k(x)$ is the PDF of the k th component (class), and N_C is the number of components. In classification applications, each component is modeled using a normal distribution [6]. A random vector x has a Gaussian or normal distribution, denoted by $x \sim N_p(m, C)$, if its probability density function (PDF) is given by [7]

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |C|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(x-m)^T C^{-1}(x-m)\right], \quad (5)$$

where m is the mean and C is the covariance matrix.

For target detection applications, the components $f_k(x)$ must capture the heavy-tail behavior of natural hyperspectral backgrounds. A useful and flexible model that fulfills this requirement is the multivariate t -elliptically contoured distribution (t -ECD), defined by [7]

$$x = (\zeta/\nu)^{-1/2} z + m, \quad (6)$$

where $z \sim N(0, C)$ and $\zeta \sim \chi_\nu^2(0)$. When $\nu = 1$, the t -ECD is equal to the multivariate Cauchy distribution and has very heavy tails. As $\nu \rightarrow \infty$, the distribution tends toward the multivariate normal distribution and has lighter tails. Modeling of natural hyperspectral backgrounds using t -ECD mixture densities is thoroughly discussed in [8].

MODELING SUBPIXEL TARGETS

Consider a target object that occupies a fraction of the pixel area. The result is a mixed pixel whose spectrum is determined by the linear mixing model defined by (2) and (3). The mixed spectrum is given by

$$x = as + (1-a)v, \quad 0 \leq a \leq 1, \quad (7)$$

where a is the target fill-fraction, s is the target spectrum, and v is the spectrum of the surrounding background. Equation (7) is known as a replacement target model because when the target is present, it “replaces” an equal part of the background. When $a = 1$, we have a resolved or full-pixel target, whereas if $0 < a < 1$, we have a subpixel target. For full-pixel targets, the major obstacle to detection is the variability of target and background spectra. An additional obstacle for subpixel targets is that, according to (7), the background acts as interference, which directly distorts the shape of the observed target spectrum.

Note that the additive target model $x = as + v$, used in radar, is a good approximation to the replacement target model when $a \ll 1$. The simplex for $m = 2$ is the line segment connecting the tips of s and v . For real data, the mixed pixels are clustered around this line because of spectral variability. This is clearly illustrated in Figure 4 in the case of the road and green grass classes; pixels at the edge of the road are mixed with green grass and lie between the road and grass swarms.

A statistical replacement model is obtained if we assume that the target $s \sim N(m_t, C_t)$ is mixed with one of the background classes [9]. Since the tails of the PDF have practically no effect on the probability of detection, we assume that $v \sim N(m_v, C_v)$. This results in the following distribution

$$\begin{aligned} x &\sim N(m(a), C(a)) \\ m(a) &= am_t + (1-a)m_v \\ C(a) &= a^2 C_t + (1-a)^2 C_v, \end{aligned} \quad (8)$$

where m_v and C_v are the parameters of the mixing class. A major practical problem with this model is that we seldom have accurate estimates of the target covariance matrix.

LIKELIHOOD RATIO DETECTORS

It is well known from statistical decision theory that decisions based upon the likelihood ratio test (LRT) are optimum over a wide range of performance criteria [10]. Among them is the Neyman–Pearson (NP) criterion, which maximizes the probability of detection for any desired probability of false alarm. The LRT is defined by

$$\Lambda(x) = \frac{f(x | \text{Target present})}{f(x | \text{Target absent})} \stackrel{H_1}{\underset{H_0}{\gtrless}} \eta, \quad (9)$$

where $f(x | H_i)$ is the conditional PDF under each hypothesis. Using the density mixture model (4) for $f(x | H_0)$ and the replacement model (8) for $f(x | H_1)$ leads to a computationally complicated LRT with many unknown parameters. Efforts to simplify the problem include the use of a single class for the background [9], [11] or clustering the data and using a detector for each class [12]. Despite their intuitive appeal, cluster-based detection techniques do not seem to perform better than the currently used matched filter algorithms [13].

The essential performance metric for target detection algorithm is the receiver operating characteristic (ROC) curve. Reliable estimation of P_D requires between 50 and 100 target pixels with approximately the same fill-fraction. This costly requirement has limited the availability of data sets for the assessment of target detection algorithms. Therefore, we often resort to performance prediction models or we pick a threshold that allows detection of all targets and we count the number of resulting false alarms.

For a fully fair comparison of detection algorithms, we should take the position of an operational user and consider how the use of each detector in a practical application affects its performance and utility. Two very important practical requirements are good performance on a diverse variety of targets and backgrounds and operational ease of use by nonexperts. Furthermore, computational complexity, numerical robustness, robustness to the selection of free parameters, and operation without a man-in-the-loop are also critical for surveillance applications.

MATCHED FILTER DETECTION THEORY

The matched filter algorithms used in current systems are often derived in the hyperspectral literature using the additive model [12], which is a good approximation for small fill-fractions and a bad one for full-pixel targets. In [1], we partially bypass this inconsistency by providing separate derivations for full- and subpixel targets. In this article, we derive the matched filter detectors without any attempt to choose a “good” signal model for the data. Instead, we use the shape of the decision surface to choose algorithms and explain their suitability for hyperspectral target detection.

Consider a $p \times 1$ observation vector x with distribution

$$x \sim N_p(as, \sigma_b^2 C_b), \tag{10}$$

where $a \geq 0$ and $\sigma_b^2 \geq 0$ are scalar quantities, s is a $p \times 1$ vector, and C_b is a $p \times p$ positive definite matrix.

The objective is to find an NP test (detector) to choose between the following hypotheses:

$$H_0: a = 0 \text{ (target absent)} \tag{11a}$$

$$H_1: a > 0 \text{ (target present)}. \tag{11b}$$

The vector s and the matrix C_b are always assumed known. Since the matrix C_b is positive definite, its square-root decomposition $C_b = C_b^{1/2} C_b^{1/2}$ is invertible. Therefore, the whitening transformation

$$\tilde{x} = C_b^{-1/2} x, \quad \tilde{s} = C_b^{-1/2} s \tag{12}$$

results in the random vector

$$\tilde{x} \sim N_p(a\tilde{s}, \sigma_b^2 I) \tag{13}$$

with spherical (“white”) normal distribution. The detection problems specified by (10) and (1) or by (13) and (1) are equivalent; thus, we use the spherical model (13) to simplify subsequent derivation and analysis.

If the distribution in (13) is fully specified, i.e., we know a and σ_b^2 , the likelihood ratio is given by $\Lambda(\tilde{x}) = f(\tilde{x} | H_1) / f(\tilde{x} | H_0)$. If $\Lambda(\tilde{x})$ exceeds a certain threshold η we accept H_1 as true; otherwise, we accept H_0 as true. Since any monotonic function of $\Lambda(\tilde{x})$ results in equivalent decisions, the LRT for the signal model (13) leads to the following test:

$$y_1 = a\tilde{s}^T \tilde{x} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_1 \sigma_b^2. \tag{14}$$

If $a > 0$ under H_1 , we can divide both sides of (14) by a without affecting the direction of the inequality. The resulting statistic

$y_2 = \tilde{s}^T \tilde{x}$ is completely independent of a . In conclusion, if a is unknown but $a > 0$, the matched filter (MF) detector, which is usually defined by the normalized expression

$$y_{MF} = \frac{\tilde{s}^T \tilde{x}}{\sqrt{\tilde{s}^T \tilde{s}}} = \frac{s^T C_b^{-1} x}{\sqrt{s^T C_b^{-1} s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{MF} \tag{15}$$

is NP-optimum for all values $a > 0$. Therefore, the MF detector (15) is a uniformly most powerful (UMP) test. The MF does not have the constant false alarm rate (CFAR) property because the threshold depends on both a and σ_b^2 .

If a and σ_b^2 are unknown, an intuitively appealing approach is to use their maximum likelihood estimates (MLEs) to form a generalized LRT (GLRT). The MLEs of a and σ_b^2 in (13) are

$$\hat{a} = \frac{\tilde{s}^T \tilde{x}}{\tilde{s}^T \tilde{s}}, \quad p\hat{\sigma}_b^2 = \tilde{x}^T \tilde{x} - \frac{(\tilde{s}^T \tilde{x})^2}{\tilde{s}^T \tilde{s}}. \tag{16}$$

The resulting maximum value of the likelihood function is $\exp(-p/2)/(2\pi\hat{\sigma}_b^2)^{p/2}$. Therefore, the GLR is

$$\Lambda_G(\tilde{x}) = \left(\frac{\hat{\sigma}_b^2 \text{ under } H_1}{\hat{\sigma}_b^2 \text{ under } H_0} \right)^{-\frac{p}{2}} = \left[1 - \frac{(\tilde{s}^T \tilde{x})^2}{(\tilde{s}^T \tilde{s})(\tilde{x}^T \tilde{x})} \right]^{-\frac{p}{2}}. \tag{17}$$

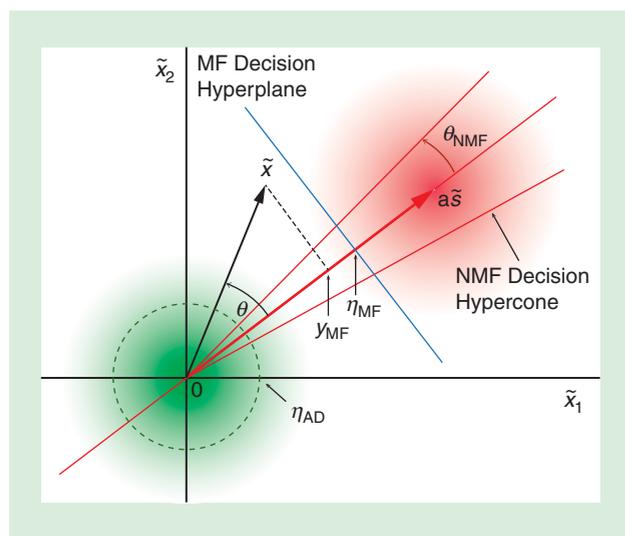
Since $a > 0$ under H_1 and only the term $\tilde{s}^T \tilde{x}$ retains the sign of \hat{a} , the normalized MF (NMF), defined by

$$y_{NMF} = \frac{\tilde{s}^T \tilde{x}}{\sqrt{\tilde{s}^T \tilde{s}} \sqrt{\tilde{x}^T \tilde{x}}} = \frac{s^T C_b^{-1} x}{\sqrt{s^T C_b^{-1} s} \sqrt{x^T C_b^{-1} x}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{NMF} \tag{18}$$

is a logical test for the hypotheses (11). The NMF is also known as a CFAR matched filter [14]. Figure 5 provides an intuitive geometrical interpretation of MF and NMF detectors in the whitened space. The vector $\tilde{s}/\|\tilde{s}\|$, where $\|\cdot\|$ denotes Euclidean length, is a unit vector; hence, the value of y_{MF} is the scalar projection of the observation vector onto the target. In contrast, the value of y_{NMF} is the cosine of the angle between the observation and target vectors. Therefore, the NMF is invariant to scaling of the target and observation measurements and has the CFAR property. In fact, the NMF has the CFAR property for all elliptically contoured distributions. The NMF is uniformly most powerful only for the class of detectors that are invariant to rotations about \tilde{s} and to scaling of target and observation spectra [10]; therefore, under the signal model (13), it cannot perform better than the UMP MF detector. This must be expected because the NMF assumes that σ_b^2 is unknown, i.e., the NMF uses “less” information.

To understand the operation of MF and NMF detectors in the whitened space, we look at their decision boundaries, which are obtained by equating (15) and (18) to a constant threshold value. The decision surface for the MF is a hyperplane perpendicular to \tilde{s} at a location determined by the threshold η_{MF} . Note that all observations vectors with tips “touching” this hyperplane have the same MF response. The decision surface for the NMF is the surface of a cone with vertex at the origin and axis along the vector \tilde{s} . The vertex angle θ_{NMF} is specified by the threshold η_{NMF} .

If the mean value of x in (10) lies in a known subspace, i.e., $E(x) = S_t a_t$, where S_t is a $p \times q_t$ matrix, the GLRT for $a_t = 0$ versus $a_t \neq 0$ is given by [1]



[FIG5] Geometrical interpretation of MF, NMF, and AD decision surfaces in the whitened-observation space. These decision surfaces determine the performance of each detector for any situation.

$$T(x) = x^T C_b^{-1} S_t (S_t^T C_b^{-1} S_t)^{-1} S_t^T C_b^{-1} x. \quad (19)$$

If $q_t = p$ and S_t has full rank, the target subspace spans the entire observation space, and the GLRT statistic (19) is reduced to the square-law detector

$$y_{AD} = x^T C_b^{-1} x = \tilde{x}^T \tilde{x} \stackrel{H_1}{\geq} \eta_{AD}, \quad (20)$$

which is also known as the RX anomaly detector (AD) [15]. We note that the MF and NMF detectors use the target signature s to determine a “preferred” search-direction in the whitened-observation space. Since the AD does not use the target signature, there is no preferred direction; the decision is based solely on the Mahalanobis distance $x^T C_b^{-1} x$ or the Euclidean distance $\|\tilde{x}\|$. Thus, the decision surface of AD is a hypersphere in the whitened-observation space (see Figure 5). Overviews of AD algorithms are provided in [3], [16], and [17].

IDEAL PERFORMANCE ANALYSIS

Detection performance is commonly evaluated in terms of ROC curves. The use of the scaled covariance $\sigma_b^2 C_b$ in the signal model (10) was necessary for the derivation of the NMF detector. To simplify subsequent derivations and notation, without any loss of generality, we assume $\sigma_b^2 = 1$ for the rest of the article.

We first determine the PDFs under the true model (13), which ensures the best theoretical performance for each detector. Since spherical distributions are invariant under orthogonal transformations (rotations) [7], to simplify the derivations, we use the transformation

$$z = Q^T \tilde{x}, \quad Q^T = Q^{-1}. \quad (21)$$

The orthogonal matrix Q is chosen such that

$$Q^T \tilde{s} = \|\tilde{s}\| e_1 \quad (22)$$

with $e_1 = [1 \ 0 \ \dots \ 0]^T$. The first column of Q is the unit vector $q_1 = \tilde{s}/\|\tilde{s}\|$; the remaining columns are arbitrarily chosen to make Q an orthogonal matrix. Since

$$z_1 = q_1^T \tilde{x} = \frac{\tilde{s}^T \tilde{x}}{\|\tilde{s}\|} = y_{MF}, \quad (23)$$

the effect of this rotation is to fit the response of the MF along the direction of the target vector in the whitened-observation space and to fit the remaining coordinates of z along the remaining “target-free” directions.

The distribution of the rotated observation vector is

$$z \sim N_p(a\|\tilde{s}\| e_1, I). \quad (24)$$

Using (20), (23), and (24) and the definition of the chi-squared distribution, we obtain

$$y_{MF} = z_1 \sim N(a\|\tilde{s}\|, 1) \quad (25)$$

$$y_{AD} = \|\tilde{x}\|^2 = \|Qz\|^2 = \|z\|^2 \sim \chi_p^2(a^2\|\tilde{s}\|^2). \quad (26)$$

To derive the distribution of the NMF, we note that

$$y_{NMF} = \frac{z_1}{\|z\|^2} = \frac{t}{\sqrt{p-1+t^2}}, \quad t \triangleq \sqrt{p-1} \frac{z_1}{\|z_2\|}, \quad (27)$$

where $z_2 \triangleq [z_2 \dots z_p]^T$. From (27) and the definition of t -distribution, we have

$$t = \frac{z_1}{\frac{1}{\sqrt{p-1}} \|z_2\|} \sim t_{p-1}(a\|\tilde{s}\|), \quad (28)$$

which is a noncentral t -distribution. The PDFs under H_1 (target present) hypothesis depend on the SCR

$$\text{SCR} \triangleq a^2 \|\tilde{s}\|^2 = a^2 s^T C_b^{-1} s, \quad (29)$$

which is the Mahalanobis distance between the means under the two hypotheses. Note that $a = 0$ under the H_0 hypothesis. Thus, we can use these PDFs to evaluate the performance of MF, NMF, and AD using ROC curves.

PERFORMANCE UNDER MODEL MISMATCHES

Application of the MF and NMF assumes perfect knowledge of the in-scene target signature s and the clutter covariance matrix C_b . We now analyze performance losses when the signature s_0 and covariance C_0 used by the detectors are

$$s_0 \neq s \quad (\text{Signature mismatch}) \quad (30)$$

$$C_0 = C_b + a^2 s s^T \quad (\text{Target in covariance}). \quad (31)$$

We first recall that the linear filter $y = h^T x$ that maximizes the output SCR, defined by

$$\text{SCR}_{\text{out}} = \frac{[E(y|H_1) - E(y|H_0)]^2}{\text{Var}(y|H_0)} = \frac{a^2 (h^T s)^2}{h^T C_b h}, \quad (32)$$

is the MF $h = \kappa C_b^{-1} s$, where κ is a normalization constant; the resulting maximum output SCR is given by (29).

Equivalently, the MF minimizes the output variance $\text{Var}(y) = \text{Var}(h^T x) = h^T C_b h$ while keeping the gain in the direction of the target constant. Thus, the MF solves the optimization problem [18]

$$\min_h h^T C_b h \quad \text{subject to} \quad h^T s = 1 \quad (33)$$

with the inconsequential change that $\kappa = 1/s^T C_b^{-1} s$.

If we use the MF $h_1 = C_b^{-1} s_0 / s_0^T C_b^{-1} s_0$, the output SCR is given by [19]

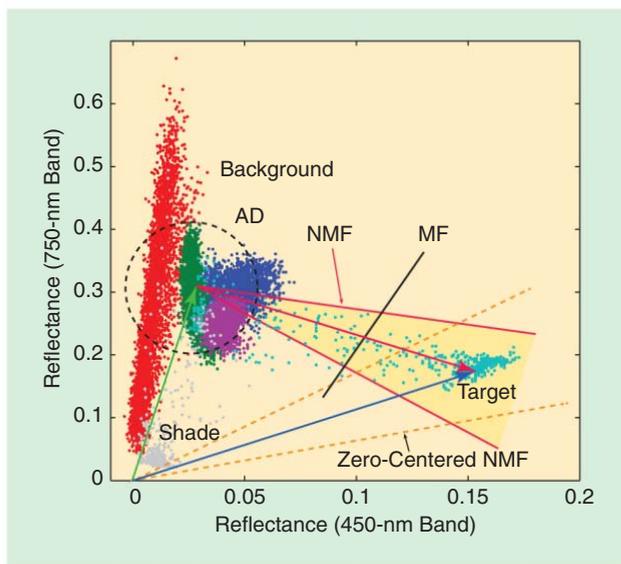
$$\text{SCR}_1 = \cos^2(s, s_0; C_b^{-1}) \cdot \text{SCR}, \quad (34)$$

where $\cos^2(s, s_0; C_b^{-1})$ is the generalized cosine defined by

$$\cos^2(s, s_0; C_b^{-1}) \triangleq \frac{(s^T C_b^{-1} s_0)^2}{(s^T C_b^{-1} s)(s_0^T C_b^{-1} s_0)} \quad (35)$$

and $-1 \leq \cos^2(s, s_0; C_b^{-1}) \leq 1$. The performance loss depends on the mismatch factor $\cos^2(s, s_0; C_b^{-1})$, but it is independent of the strength a of the target. If the target is present in the clutter covariance matrix, [see (31)], the MF is defined by $h_2 = C_0^{-1} s_0 / s_0^T C_0^{-1} s_0$. The output SCR is given by [19]

$$\text{SCR}_2 = \text{SCR}_1 [1 + (2\text{SCR} + \text{SCR}^2) \sin^2(s, s_0; C_b^{-1})]^{-1}, \quad (36)$$



[FIG6] The decision cone (yellow area) of the NMF includes the plurality of subpixel and full-pixel “road” targets (cyan dots). Note the importance of “centering” the cone at the mean of the clutter cloud.

where $\sin^2(s, s_0; C_b^{-1}) = 1 - \cos^2(s, s_0; C_b^{-1})$. The loss due to covariance contamination (target-in-covariance) is determined by the term $(2SCR + SCR^2) \sin^2(s, s_0; C_b^{-1})$, which shows a strong dependence on the target strength through SCR and on mismatch through $\sin^2(s, s_0; C_b^{-1})$. Thus for strong targets, the losses can be significant.

Similar conclusions can be drawn by performance evaluation of both the MF and NMF detectors using ROC curves under the Gaussian model (10). Similar analysis also shows that the NMF, which uses angular information, is more sensitive to mismatch than the MF [20].

Although the effects of target mismatch and covariance contamination have been extensively analyzed, there are not many practical techniques for their mitigation. A plausible way to reduce mismatch losses is to relax the “line-pointing” constraint in (33). This can be done by changing the constraint in (33) to form the following optimization problem:

$$\min_h h^T C_b h \quad \text{subject to} \quad h^T s \geq 1 \quad (37)$$

for all signatures s in the hypersphere $\|s - s_0\|^2 \leq \epsilon^2$. The value of ϵ determines the “selectivity” of the detector. Several solutions to (37) have been proposed in the literature [18]; the most practical ones lead to some kind of diagonal loading $C_b + \xi^2 I$. An interesting solution in [21] provides a direct link between ϵ and the diagonal loading factor ξ . The resulting MFs are robust to signal mismatch and covariance contamination in exchange for a higher false alarm rate.

PRACTICAL HYPERSPECTRAL DETECTORS

The scatter plot in Figure 4 and the decision surfaces in Figure 5 provide insight for the selection of hyperspectral target detection algorithms. We note that an NMF with a vertex at the center of the

grass cloud and axis pointing to the road cloud is the natural choice for the detection of subpixel and full pixel “road” targets. This is illustrated in Figure 6, which also clearly shows why a zero-centered NMF would not be a good choice for subpixel targets. However, we must stress that the ultimate criterion of success is performance in real applications.

The clutter-centered MF and NMF are obtained by subtracting the clutter mean from (15), (18), and (20). This yields

$$y_{MF} = \frac{(s - m_b)^T C_b^{-1} (x - m_b)}{\sqrt{(s - m_b)^T C_b^{-1} (s - m_b)}} \quad (38)$$

$$y_{AD} = (x - m_b)^T C_b^{-1} (x - m_b) \quad (39)$$

$$y_{NMF} = \frac{y_{MF}}{\sqrt{y_{AD}}} \quad (40)$$

In practice, the mean m_b and covariance C_b of the background clutter are replaced by their maximum likelihood estimates

$$\hat{m}_b = \frac{1}{N} \sum_{k=1}^N x_k, \quad \hat{C}_b = \frac{1}{N} \sum_{k=1}^N (x_k - \hat{m}_b)(x_k - \hat{m}_b)^T. \quad (41)$$

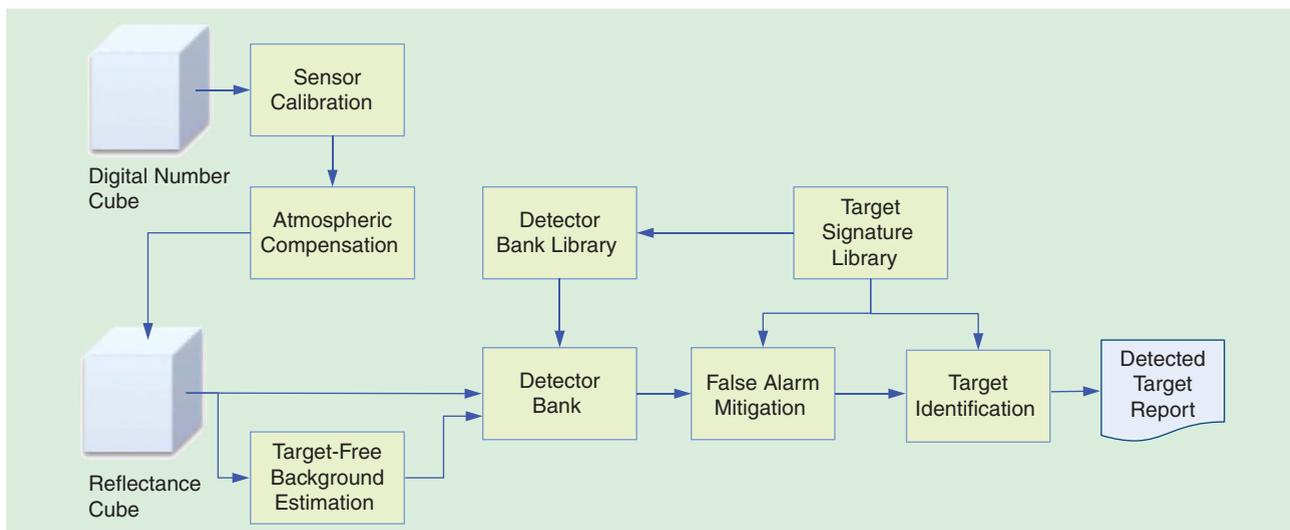
The square of (38) is known as the adaptive MF (AMF) [22]. The square of (40) is known as the adaptive NMF (ANMF) [23] or adaptive cosine/coherence estimator (ACE) [24]. The term *adaptive* is often used to emphasize that the clutter moments are estimated from the data.

We shall use a simple example to illustrate some challenges related to the application and performance evaluation of AMF and ANMF detectors using a data set from the Hyperspectral Digital Imagery Collection Experiment (HYDICE) Forest Radiance I experiment [2]. Figure 7 shows the basic components of a typical target detection system. The cube of raw digital numbers is calibrated to a radiance cube, which is then converted to reflectance to allow the use of target signatures from a library. We focus on a green tarp target having the library signature shown in Figure 8; all other targets have been removed from the data. There are three different sizes of green tarps in the scene; the smallest tarp corresponds to a subpixel target.

We first note from Figure 8 that there is a mismatch between library and in-scene target spectra due mainly to material variability, atmospheric compensation errors, sensor calibration limitations, and mixing for subpixel targets. Furthermore, there is mismatch between the normal distribution assumed in (10) and the distribution of the multiclass background of Forest Radiance I scene. Since the presence of targets in the estimate \hat{C}_b decreases performance, we use the AD algorithm (20) to remove “targetlike” pixels from the training data (target-free background estimation).

In practice, \hat{C}_b is often ill conditioned, and its small eigenvalues and corresponding eigenvectors are difficult to estimate and hard to compute accurately even if $N \gg p$. Numerically stable and more robust detectors can be formed by eliminating the smaller eigenvalues through averaging and loading. We often use dominant mode rejection combined with diagonal loading [25], [26].

Figure 9 shows images of the scene and the responses of the AD, MF, and NMF detectors for each pixel. Inspection of these detection statistics indicates that the MF and NMF score high at



[FIG7] A block diagram of a typical hyperspectral target detection system.

the target pixels; however, we stress that visual assessment is not a dependable way to evaluate detection performance.

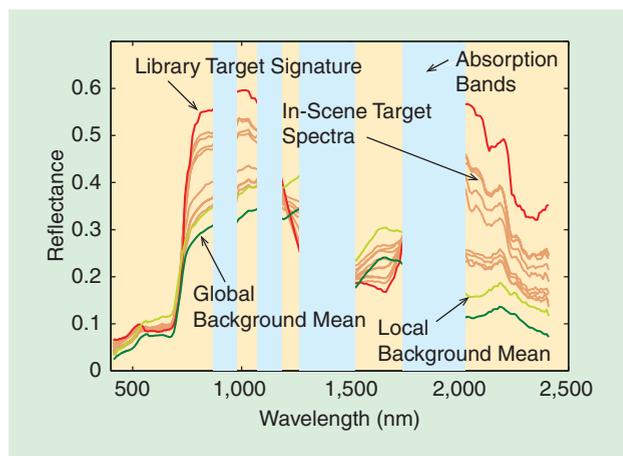
PERFORMANCE FOR MULTICLASS BACKGROUNDS

If we model the background clutter with a density mixture of t -ECDs and the target with the statistical replacement model, we can develop a model to investigate the performance of MF and NMF detectors under more realistic scenarios [27]. Needless to say, in practice, it is impossible to find PDFs that capture all aspects of hyperspectral imaging data. The presence of man-made structures in a scene can produce false alarms that cannot be predicted by any model.

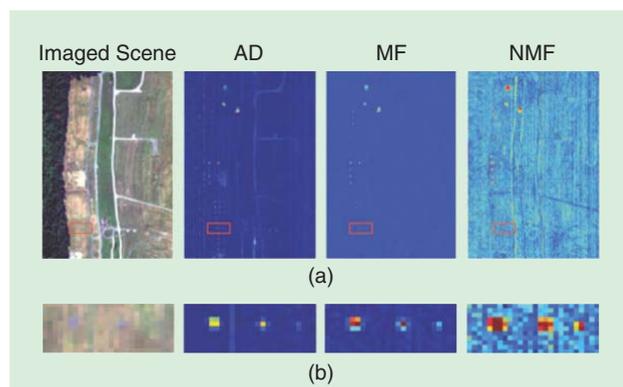
The multiclass model has been used to develop a performance prediction tool useful in obtaining insight and evaluating sensors and detection algorithms under different deployment conditions. Another interesting and useful application of the model is shown in Figure 10, which shows a scatter plot of NMF versus MF responses for the Forest Radiance I data set. The ribbon shows a 95% confidence interval for the NMF and MF responses predicted by a five-class t -ECD model. We first note that the model does a reasonable job at predicting the range of detector responses and the majority of background responses. Clearly, the NMF provides a better separation between target and clutter pixels compared to the MF. Most pixels that exceed the MF threshold belong to other targets and man-made objects present in the scene. We again stress that such results do not prove the validity of the model; a thorough validation of the predictive model, which is work in progress, is a laborious and expensive undertaking.

FALSE ALARM MITIGATION ALGORITHMS

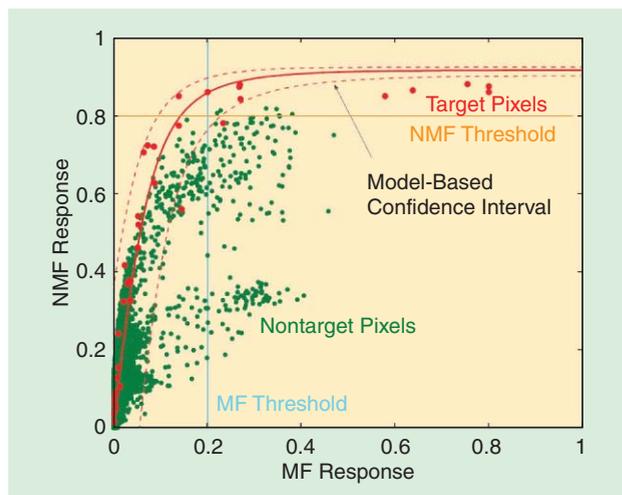
In practice, all detection algorithms exhibit a large number of false alarms [28]. A promising approach to deal with this challenge is postprocessing of the top hits with FAM algorithms [11]. The NMF and MF detectors make decisions by exploiting the statistical distributions of targets and clutter in the clutter-whitened space. Practical experience has shown that both algorithms are relatively



[FIG8] Reflectance spectra for the library target signature, the in-scene target spectra, the local background mean around the targets, and the global background mean for the Forest Radiance I data cube.



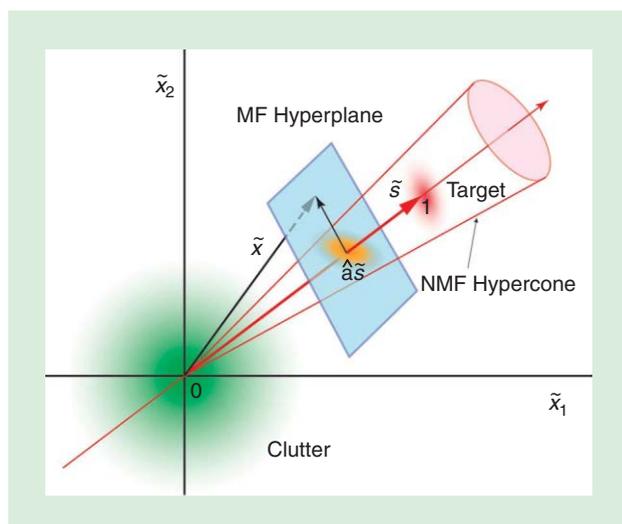
[FIG9] Imaged scene and detection statistics (detector response) for the AD, MF, and NMF algorithms. The images in (b) show the area around the targets of interest, which correspond to (a) the red frame box in the large images.



[FIG10] NMF versus MF values scatter plot and confidence “ribbon” for expected values constructed using the performance prediction model.

robust to model mismatches, including clutter multimodality, covariance estimation, and target signature mismatch. Therefore, FAM algorithms must exploit information not available in the detection statistic. In this sense, the development of FAM algorithms is an art involving statistics, spectroscopy, and spatial image processing. False alarms can also be mitigated using information from coregistered high-resolution video.

A simple but useful FAM algorithm for the MF detector can be obtained by exploiting the geometry in Figure 11. We note that all observation vector with tips touching the MF hyperplane have identical MF responses. However, according to model (8), “true” subpixel targets have a normal distribution centered at the line connecting the target and clutter distributions at a location specified by the target fill factor. The center of the target distribution is the estimated fill factor \hat{a} in (16), which is given by $\hat{a} = y_{MF} / \|\tilde{s}\|$. The Mahalanobis square distance



[FIG11] The geometrical interpretation of MF and NMF detection statistic uncertainties leading to false alarms in the spectral whitened space.

$$\Delta^2(x) = \|\tilde{x} - \hat{a}\tilde{s}\|^2 = (x - \hat{a}s)^T C^{-1}(a) (x - \hat{a}s) \quad (42)$$

is inversely proportional to the likelihood that the test pixel is a subpixel target. The approximation $C(a) \simeq C_b$ works reasonably well in practical applications. Figure 12 shows a typical scatter plot of MD versus MF response for a detection experiment involving a man-made target. We note that with the use of a second threshold, we can significantly reduce the number of false alarms. The use of a CFAR threshold is possible by modeling the distribution of (42). More details about the MF with FAM can be found in [11].

TARGET IDENTIFICATION

The task of a TID algorithm is to determine whether a pixel picked out by the detector contains indeed the target of interest. This is a difficult undertaking because different scene components may result to similar or identical spectra (mimics). The problem is more prominent for subpixel targets because of spectral mixing. In this case, the first step is to estimate the target spectrum using unmixing techniques. Identification algorithms use the estimated spectrum and material spectral libraries to associate this spectrum with the reference spectrum of some material. Typical approaches exploit ideas from the areas of linear unmixing, F-test-based model selection, Bayesian model selection [29], [30], band selection techniques [31], and spectroscopy [32].

CONCLUSIONS

The development of practical hyperspectral target detection algorithms requires the understanding and joint exploitation of phenomenology, sensor technology, and statistical signal processing. Target variability, the mismatch between library and in-scene signatures, and false alarms from spectral “mimics” constitute major challenges to be overcome by future algorithms. Therefore, the development of FAM and TID algorithms is expected to be an active area of research in hyperspectral target detection.

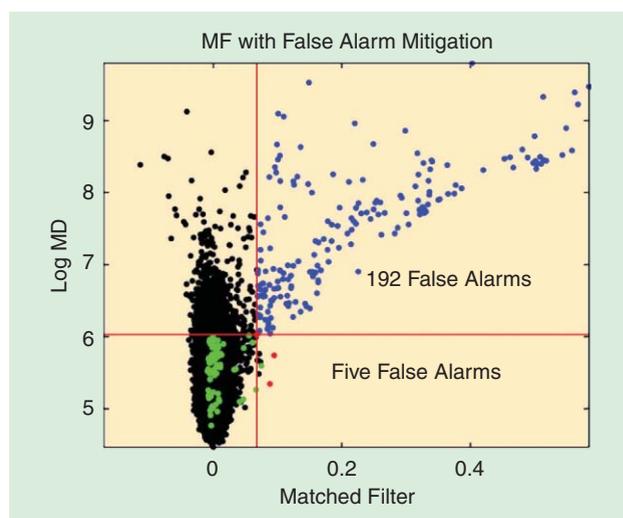
ACKNOWLEDGMENTS

This work is sponsored by the Department of the Air Force under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.

AUTHORS

Dimitris Manolakis (dmanolakis@ll.mit.edu) received his B.S. degree in physics and his Ph.D. degree in electrical engineering from the University of Athens, Greece. He is currently a senior staff member at the Massachusetts Institute of Technology (MIT) Lincoln Laboratory, in Lexington. He is a coauthor of the textbooks *Digital Signal Processing: Principles, Algorithms, and Applications* (Prentice Hall, 2006, 4th ed.), *Statistical and Adaptive Signal Processing* (Artech House, 2005), and *Applied Digital Signal Processing* (Cambridge University Press, 2011).

Eric Truslow (Eric.Truslow@ll.mit.edu) received the B.S. degree in electrical engineering from Union College in Schenectady, New York, in 2010 and the M.S. degree in electrical engineering from Northeastern University in Boston, Massachusetts, in



[FIG12] The MF with FAM uses a “subpixel likelihood metric” and double thresholding to reduce false alarms.

2012. As a research assistant at the MIT Lincoln Laboratory, he has been exploring material detection, data modeling, and detector performance prediction in hyperspectral imagery. His other research interests include digital signal processing, pattern recognition, and image processing.

Michael Pieper (mpieper@ll.mit.edu) received his B.S. and M.S. in electrical engineering from Northeastern University in 2007. He is currently an associate staff member at the MIT Lincoln Laboratory, in Lexington, Massachusetts. His research experience and interests include the application of digital signal processing, adaptive filtering, and pattern recognition to remote hyperspectral imaging sensing.

Thomas Cooley (thomas.cooley@kirtland.af.mil) received the B.S. degree in electrical engineering from Rensselaer Polytechnic Institute, an M.S. degree in electrical engineering and applied physics from the California Institute of Technology (Caltech), and a Ph.D. degree in optical sciences from the University of Arizona. He is currently the mission lead for the Air Force Research Laboratory (AFRL) Space Intelligence, Surveillance, and Reconnaissance. He was the principal investigator of the Artemis payload, a hyperspectral sensor to demonstrate the use of this rich data source for the wide range of applications, and the program manager for the TacSat-3 satellite launched in May 2009.

Michael Brueggeman (Michael.Brueggeman@kirtland.af.mil) received his B.S. degree in physics and mathematics from Florida State University. He is a senior physicist in the Geospatial Intelligence R&D section at the National Air and Space Intelligence Center, where he researches new sensor designs and algorithm development. Prior to this role, he was the program manager of the Imaging Spectroscopy Program at the AFRL and was previously a pilot in the United States Marine Corps.

REFERENCES

- [1] D. Manolakis and G. Shaw, “Detection algorithms for hyperspectral imaging applications,” *IEEE Signal Processing Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.
- [2] D. Manolakis, D. Marden, and G. Shaw, “Target detection algorithms for hyperspectral imaging application,” *Lincoln Lab. J.*, vol. 14, no. 1, pp. 79–116, 2003.

- [3] N. M. Nasrabadi, “Hyperspectral target detection: An overview of recent and future challenges,” *IEEE Signal Processing Mag.*, vol. 31, no. 1, pp. 34–44, 2014.

- [4] G. Felde, G. Anderson, T. Cooley, M. Matthew, S. Adler-Golden, A. Berk, and J. Lee, “Analysis of Hyperion data with the FLAASH atmospheric correction algorithm,” in *IEEE Int. Geoscience and Remote Sensing Symp.*, vol. 1, pp. 90–92, July 2003.

- [5] L. S. Bernstein, S. M. Adler-Golden, R. L. Sundberg, and A. J. Ratkowski, “In-scene-based atmospheric correction of uncalibrated VISible-SWIR (VIS-SWIR) hyper- and multispectral imagery,” in *Remote Sensing of Clouds in the Atmosphere*, vol. 7107. Wales, U.K.: SPIE, 2008, pp. 710706–710706-7.

- [6] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, NJ: Wiley, 2003.

- [7] R. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.

- [8] D. Manolakis, M. Rossacci, D. Zhang, J. Cipar, R. Lockwood, T. Cooley, and J. Jacobson, “Statistical characterization of hyperspectral background clutter in the reflective spectral region,” *Appl. Opt.*, vol. 47, no. 28, pp. F96–F106, 2008.

- [9] A. Schaum and A. Stocker, “Spectrally-selective target detection,” *Proc. ISSSR*, 1997.

- [10] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*. New York, NY: Springer, 2008.

- [11] R. DiPietro, D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, “Hyperspectral matched filter with false-alarm mitigation,” *Opt. Eng.*, vol. 51, no. 1, p. 016202, Jan. 2012.

- [12] M. T. Eismann, *Hyperspectral Remote Sensing*. Bellingham, WA: SPIE, 2012.

- [13] M. Pieper, D. Manolakis, E. Truslow, T. Cooley, and S. Lipson, “Performance evaluation of cluster-based hyperspectral target detection algorithms,” in *Proc. 19th IEEE Int. Conf. Image Processing (ICIP)*. Piscataway, NJ: IEEE, 2012, pp. 2669–2672.

- [14] L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.

- [15] I. S. Reed and X. Yu, “Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.

- [16] D. Stein, S. Beaven, L. Hoff, E. Winter, A. Schaum, and A. Stocker, “Anomaly detection from hyperspectral imagery,” *IEEE Signal Processing Mag.*, vol. 19, pp. 58–69, Jan. 2002.

- [17] S. Matteoli, M. Diani, and G. Corsini, “A tutorial overview of anomaly detection in hyperspectral images,” *IEEE Aerospace Electron. Mag.*, vol. 25, no. 7, pp. 5–28, 2010.

- [18] R. Lorenz and S. Boyd, “Robust minimum variance beamforming,” *IEEE Trans. Signal Processing*, vol. 53, no. 5, pp. 1684–1696, May 2005.

- [19] H. Cox, “Resolving power and sensitivity to mismatch of optimum array processors,” *J. Acoust. Soc. Am.*, vol. 54, no. 3, pp. 771–785, 1973.

- [20] C. D. Richmond, “Statistical performance analysis of the adaptive sidelobe blanker detection algorithm,” in *Proc. 31st Asilomar Conf. Signals and Systems*, 1997, pp. 872–876.

- [21] J. Li, P. Stoica, and Z. Wang, “On robust Capon beamforming and diagonal loading,” *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1702–1715, July 2003.

- [22] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, “A CFAR adaptive matched filter detector,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–218, Jan. 1992.

- [23] E. Conte, M. Lops, and G. Ricci, “Adaptive matched filter detection in spectrally invariant noise,” *IEEE Signal Processing Lett.*, vol. 3, no. 8, pp. 248–250, 1996.

- [24] S. Kraut, L. Scharf, and L. T. McWhorter, “Adaptive subspace detectors,” *IEEE Trans. Signal Processing*, vol. 49, no. 1, pp. 1–16, Jan. 2001.

- [25] H. Cox and R. Pitre, “Robust DMR and multi-rate adaptive beamforming,” *Conf. Rec. of the 31st Asilomar Conf. on Signals, Syst. and Comput.*, vol. 1, pp. 920–924, Nov. 1997.

- [26] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, “Hyperspectral detection algorithms: Use covariances or subspaces?” in *Proc. SPIE Conf. Imag. Spectroscopy XIV*, vol. 7457, pp. 74570Q–74570Q-8, 2009.

- [27] E. Truslow, D. Manolakis, M. Pieper, T. Cooley, M. Brueggeman, and S. Lipson, “Performance prediction of matched filter and ACE hyperspectral target detectors,” *IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing*, no. 99, pp. 1–14, July 2013. DOI: 10.1109/JSTARS.2013.2272697.

- [28] M. Eismann, J. Meola, and A. Stocker, “Automated hyperspectral target detection and change detection from an airborne platform: Progress and challenges,” in *2010 IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, July 2010, pp. 4354–4357.

- [29] M. L. Pieper, D. Manolakis, E. Truslow, T. Cooley, and M. Brueggeman, “False alarm mitigation techniques for hyperspectral target detection,” in *SPIE Defense, Security, and Sensing*, pp. 874304–874304-14, 2013.

- [30] T. Burr, H. Fry, B. McVey, E. Sander, J. Cavanaugh, and A. Neath, “Performance of variable selection methods in regression using variations of the Bayesian information criterion,” *Commun. Stat.-Simul. Comput.*, vol. 37, no. 3, pp. 507–520, Feb. 2008.

- [31] N. Keshava, “Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries,” *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 7, pp. 1552–1565, 2004.

- [32] J. B. Adams and A. R. Gillespie, *Remote Sensing of Landscapes with Spectral Images: A Physical Modeling Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2006.



[Nasser M. Nasrabadi]

Hyperspectral Target Detection



[An overview of current and future challenges]

Over the last decade, hyperspectral imagery (HSI) obtained by remote sensing systems has provided significant information about the spectral characteristics of the materials in the scene. Typically, a hyperspectral spectrometer provides hundreds of narrow contiguous bands over a wide range of the electromagnetic spectrum. Hyperspectral sensors measure the reflective (or emissive) properties of objects in the visible and short-wave infrared (IR) regions (or the mid-wave and long-wave IR regions) of the spectrum. Processing of these data allows algorithms to detect and identify targets of interest in a hyperspectral scene by exploiting the spectral signatures of the materials [1], [2].

Digital Object Identifier 10.1109/MSP.2013.2278992

Date of publication: 5 December 2013

Target detection is basically a binary classifier with the aim of labeling every pixel in the image as a target or background. Since only a small fraction of all the pixels in the image can be labeled anomalies or classified as targets, the overall classification error is not a good measure of performance since pixels of interest are sparse. Therefore, classical target detection algorithms are based on the Neyman–Pearson criterion, which maximizes the probability of detection for any fixed probability of false alarm. A detailed discussion of performance and evaluation of the classical target detection techniques and their implementation issues are discussed in a companion article [3]. However, this article mainly focuses on the use of more recent statistical signal processing and machine-learning techniques for hyperspectral anomaly and target detection. Techniques such as support vector data description (SVDD), sparse representation classifiers, regularization techniques, kernel-based detectors, fusion of detectors, subspace-based

methods, and challenges associated with their implementations are reviewed. In this article we only show experimental results for full-pixel targets. However, most of the algorithms are applicable or can be extended to subpixel or gaseous targets where the strength of the target is unknown.

THIS ARTICLE MAINLY FOCUSES ON THE USE OF MORE RECENT STATISTICAL SIGNAL PROCESSING AND MACHINE-LEARNING TECHNIQUES FOR HYPERSPECTRAL ANOMALY AND TARGET DETECTION.

against a background clutter with unknown spectral covariance. This algorithm, now commonly referred to as the RX anomaly detector, has been successfully applied in many hyperspectral applications and is considered as the benchmark anomaly detection algorithm for multispectral/hyperspectral data. The RX algorithm is

ANOMALY DETECTORS

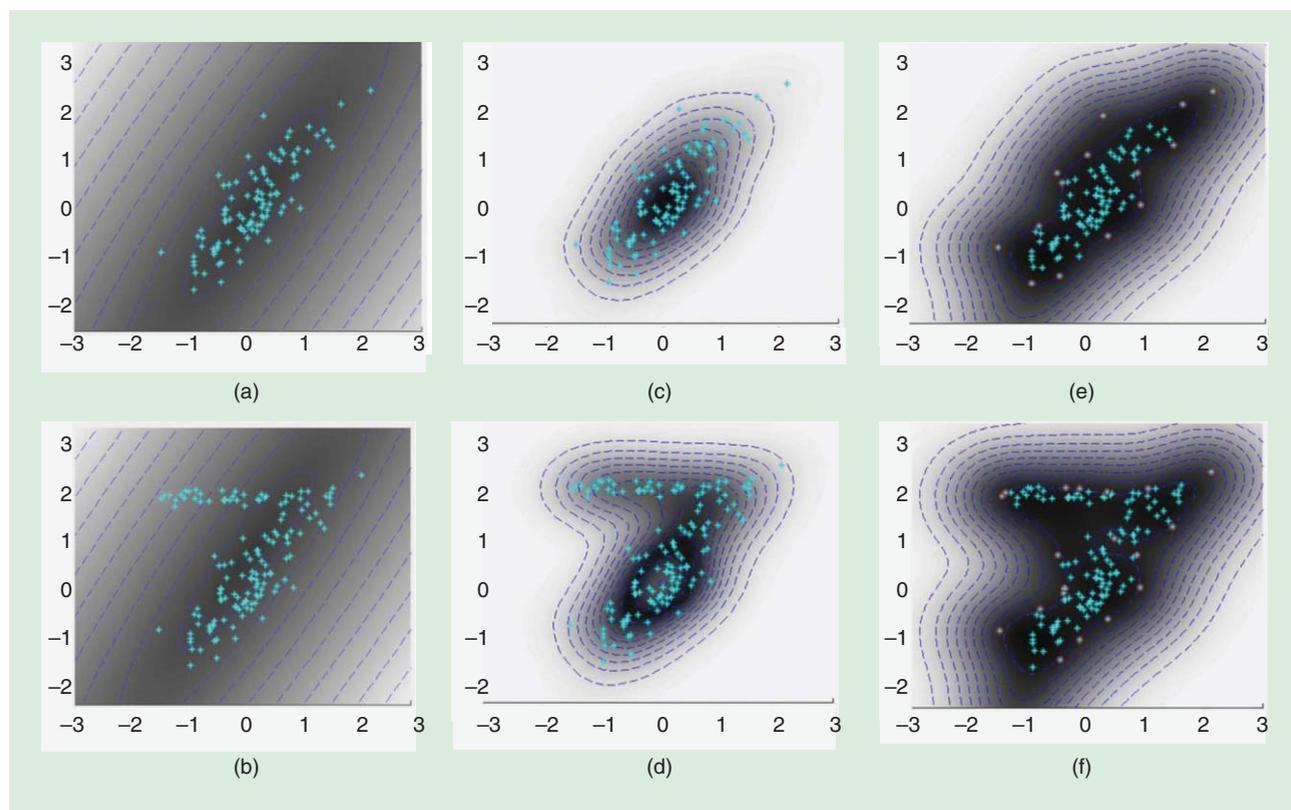
Anomaly detectors are pattern recognition or statistical schemes used to detect objects that stand out from the cluttered background. In spectral anomaly detection algorithms [4], [5], pixels that have a significantly different spectral signature from their neighboring background clutter pixels are identified as spectral anomalies. In such algorithms, no prior knowledge of the target spectral signature is used or assumed. In this section, the classical Reed–Xiaoli (RX) anomaly detector, kernel RX, SVDD, and subspace-based anomaly detector are discussed.

RX ANOMALY DETECTION

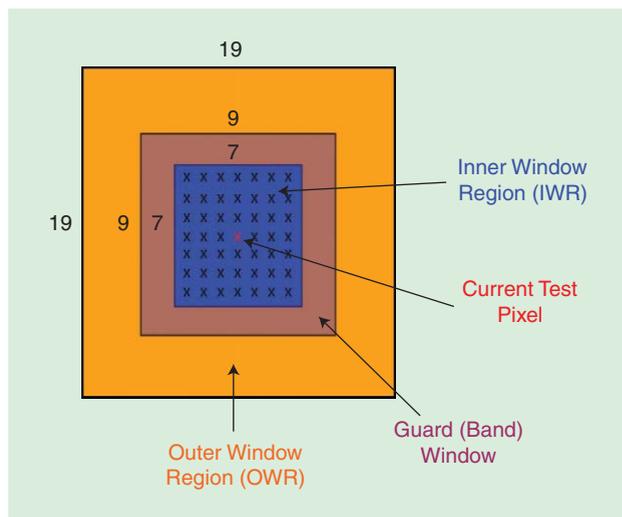
In [6], a spectral anomaly detection algorithm was developed for detecting targets of unknown spectral characteristics

a constant false alarm rate (CFAR) adaptive anomaly detector that is derived from the generalized likelihood ratio test (GLRT) [6]. The CFAR property allows the detector to use a single threshold to maintain a desired false alarm rate regardless of the background variation at different locations in the scene. It is based on the assumption that the background statistics can be modeled as a multivariate Gaussian distribution whose mean and covariance are estimated from the pixels in the image. Assuming a single pixel $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_p]^T \in \mathbb{R}^p$ as the observation test vector consisting of p bands, the output of the RX algorithm is given by

$$RX(\mathbf{x}) = (\mathbf{x} - \hat{\boldsymbol{\mu}}_b)^T \hat{\mathbf{C}}_b^{-1} (\mathbf{x} - \hat{\boldsymbol{\mu}}_b), \quad (1)$$



[FIG1] Decision contours and surface plots of RX, KRX, and SVDD for two illustrative toy distributions: RX for (a) a single Gaussian background distribution and (b) a mixture of Gaussians background distribution; KRX (RBF kernel $\sigma = 1.0$) for (c) a single Gaussian background distribution and (d) a mixture of Gaussians background distribution; SVDD (RBF kernel $\sigma = 1.0$, $C = 0.1$ and support vectors are indicated by the red circles) for (e) a single Gaussian background distribution and (f) a mixture of Gaussians background distribution.



[FIG2] An example of a double concentric sliding window with guard band centered at pixel x .

where $\hat{\mu}_b$ is the estimated background clutter sample mean, and \hat{C}_b is the estimated background clutter covariance obtained from the test image.

The RX algorithm is simply the square of the Mahalanobis distance between the test pixel and the local background mean. This is shown graphically for a toy example in Figure 1(a) with a single Gaussian background distribution and Figure 1(b) with the background made of a mixture of two Gaussian distributions. In this figure, the background data are assumed to be the blue dots and the rest of the space is considered to be the anomaly space. It is clear from Figure 1(b) that the conventional RX algorithm fails to tightly model the background distribution in the case of background being a mixture of two Gaussian distributions.

The background mean and covariance matrix can be estimated globally from the whole hyperspectral image or locally using a double concentric sliding window approach (see Figure 2). To estimate \hat{C}_b globally, a single Gaussian distribution is inadequate to model the whole background clutter distribution. Therefore, more advanced methods are proposed to model the nonstationarity of the background clutter such as using a mixture of multivariate Gaussian distributions, linear or stochastic mixture models or by some clustering techniques that are used to segment the background into several clusters. On the other hand, the local background covariance matrix can be estimated by using a double concentric sliding window centered at each test pixel, consisting of a small inner window region (IWR) centered within a larger outer window region (OWR), as shown in Figure 2. The local background mean vector and covariance matrix are then computed from the spectral pixels falling within the OWR. The size of the inner window is assumed to be the size of the typical target of interest in the image. A guard band surrounding the IWR is sometimes used to prevent the target pixels from corrupting the calculation of the background OWR statistics. It should be

pointed out that the local RX algorithm is computationally very intensive when compared with the global RX algorithm. This is due to the need to estimate and invert a large covariance matrix at each location of the double concentric sliding window centered on each test pixel.

Several variations of the RX detector that attempt to alleviate the limitation of RX have been proposed in the literature [4], [7]. As an example, in [7], a modification to the RX algorithm called subspace RX (SSRX) was outlined that is based on the principal component analysis (PCA) of the background covariance matrix. In the SSRX algorithm, several high-variance background dimensions are deleted before applying the RX algorithm, as these dimensions are assumed to capture non-normal background clutter variance. Another consideration in RX implementation is the potential ill conditioning of the local covariance matrix due to the high correlation, high dimensionality of the hyperspectral data, and a limited background sample size. This ill conditioning is typically addressed by shrinkage methods [8] for regularizing estimates of a large covariance matrix. Regularization procedures such as PCA-based regularization (discarding the contributions from the eigenvectors with low eigenvalues) or adding a scaled identity matrix ($\lambda\mathbf{I}$) to the background covariance matrix (a popular technique in the ridge regression procedure [8]) have been reported in the literature. For example, the output of the ridge-regularized RX algorithm is

$$RX_{\text{reg}}(x) = (x - \hat{\mu}_b)^T (\hat{C}_b + \lambda\mathbf{I})^{-1} (x - \hat{\mu}_b), \quad (2)$$

where λ is a regularization parameter and \mathbf{I} is the identity matrix. Other types of regularization methods can be found in the literature dealing with the rank-deficient and ill-posed problems [9].

SUBSPACE-BASED ANOMALY DETECTION

Anomaly detection techniques formulated as modeling the whole (local) background as a subspace and eliminating it from every pixel have also been investigated in [10]. In subspace anomaly detectors the input spectra is projected onto a subspace, where bases are defined by some projection vectors. The projection separation statistic for an input test pixel, x , is calculated using $s' = (x - \hat{\mu}_b)^T (\mathbf{I} - \mathbf{W}\mathbf{W}^\#) (x - \hat{\mu}_b)$, where $\mathbf{W} = [w_1 w_2 \dots w_m]$ is a matrix with columns that are the m projection bases obtained from the background samples, $\hat{\mu}_b$ is the estimated mean of the background samples, and $\mathbf{W}^\# = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T$ denotes the pseudoinverse of \mathbf{W} . The product $\mathbf{W}\mathbf{W}^\#$ is known as a projection operator and represents the background subspace. An anomaly is detected if the projection separation, s' , is greater than some threshold. There are several different methods to generate the projection basis to obtain the background subspace. The typical approach is to obtain the eigenvalue decomposition of the background covariance matrix and then select the significant eigenvectors (with large eigenvalue) that capture the background subspace without including the sensor noise statistics. Another approach to obtain the projection basis is to extract the

background endmembers from the image, locally or globally, using a hyperspectral unmixing procedure [11].

Both subspace-based and RX anomaly detectors try to suppress the background clutter subspace. For example, in the subspace anomaly detector, the background subspace is completely discarded by projecting the data onto a complement projection operator, $(I - WW^{\#})$, which is generated from the significant eigenvectors (with large eigenvalues) W of the background covariance. However, in the case of RX the background subspace is suppressed by normalizing the significant eigenvectors of the back covariance matrix with their corresponding large eigenvalues and the remaining minor eigenvectors, representing the target subspace, are enhanced by normalizing them by their corresponding low eigenvalues.

KERNEL RX ANOMALY DETECTION

In general, the background clutter is not Gaussian distributed, and it is not even practical to model it as a mixture of Gaussians due to insufficient training data and the lack of knowledge about the number of Gaussian mixtures. To address this problem, a nonlinear version of the RX algorithm called kernel RX (KRX) was developed in [12]. The key concept of KRX is to express the original RX model in a high-dimensional feature space \mathcal{F} , where the decision regions in this new feature space will correspond to more complex decision regions than in the original input space. Given a nonlinear mapping function $\Phi(x)$ that maps an input vector x into a potentially much higher (possibly infinite) dimensional feature space. The output of the RX algorithm in this new feature space is represented as

$$RX(\Phi(x)) = (\Phi(x) - \hat{\mu}_{b\phi})^T \hat{C}_{b\phi}^{-1} (\Phi(x) - \hat{\mu}_{b\phi}), \quad (3)$$

where $\hat{C}_{b\phi}$ and $\hat{\mu}_{b\phi}$ are the estimated covariance and mean of the background clutter samples in the feature space, respectively.

Direct implementation of the RX algorithm in the feature space described by (3) is not possible due to the nonlinear mapping Φ , which produces a data space of high dimensionality. To avoid implementing (3) explicitly, we need to kernelize it, which corresponds to expressing (3) in dot products form and evaluating it using the kernel trick concept [13]. The kernel trick concept allows the computation of the dot products in the feature space using positive definite kernels defined on a pair of vectors

$$k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle. \quad (4)$$

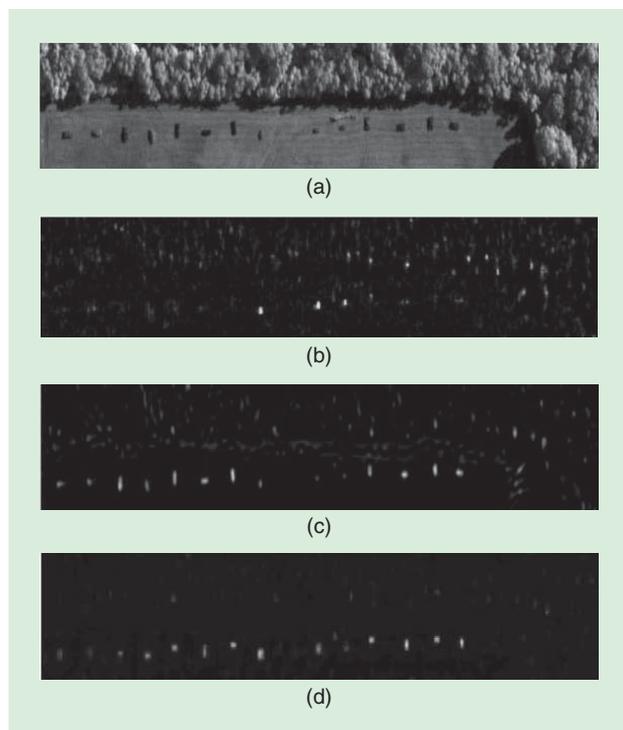
Equation (4) shows that the dot products in \mathcal{F} can be replaced by a kernel function k , a nonlinear function that can be easily calculated without identifying the nonlinear map Φ . A commonly used kernel is the Gaussian radial basis function (RBF) $k(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / (\sigma^2))$, where $\sigma > 0$ is the kernel bandwidth parameter. Detailed information about the properties of different kernels and kernel-based learning theory can be found in [13].

The kernelized version of RX in the feature space (3) is given in [12] as

$$RX_k(x) = (k_x - k_{\hat{\mu}_b})^T \hat{K}_b^{-2} (k_x - k_{\hat{\mu}_b}), \quad (5)$$

where $\hat{K}_b = \Phi(X_b)^T \Phi(X)$ denotes the centered kernel matrix for the mean-removed background pixels $\Phi(X_b)$ in the feature space, $k_x = \Phi(X_b)^T \Phi(x)$ represents the so-called empirical kernel map of the test pixel $\Phi(x)$, and $k_{\hat{\mu}_b} = \Phi(X_b)^T \Phi(\hat{\mu}_b)$ is the corresponding empirical kernel map of the background mean $\Phi(\hat{\mu}_b)$. Equation (5) can now be implemented with no knowledge of the mapping function Φ . The only requirement is a good choice for the kernel function k , which can produce a positive definite Gram matrix.

Figure 1(c) and (d) shows the decision surfaces of KRX for the background clutter statistics made of a single Gaussian or a mixture of two Gaussian distributions. Comparing Figure 1(b) and (d), it is clear that when the background statistics is not a pure Gaussian, the kernel RX can provide the better nonlinear decision surfaces. Figure 3(a) shows a typical hyperspectral image, the Forest Radiance I data collection (FR-I), obtained from a hyperspectral digital imagery collection experiment (HYDICE) sensor, which consists of 210 bands across the whole spectral range from 0.4 to 2.5 μm and includes the visible and short-wave IR bands. Figure 3(b)–(d) shows the RX, KRX, and SVDD (discussed in the next section) anomaly detection results



[FIG3] HYDICE hyperspectral FR-I image, RX, KRX, and SVDD anomaly detected images. (a) Sample band image (48th) from the FR-I image; detection results based on the local background statistics using a double concentric sliding window for (b) local RX algorithm, (c) local kernel RX algorithm (RBF kernel $\sigma = 40$), and (d) local SVDD algorithm (RBG kernel $\sigma = 16$, $C = 0.05$).

using the local background statistics obtained by a double concentric sliding window, respectively. Visually, it is clear that the KRX and SVDD do a better job in capturing and eliminating the background clutter.

SVDD FOR ANOMALY DETECTION

There are two major issues with the RX algorithm that limit its performance. In many applications, it has been shown empirically that the Gaussian model for the background, even locally, provides an inadequate representation of the underlying distribution [5] leading to poor false alarm performance. This is especially true when the local background contains multiple classes of material. In addition, the local RX algorithm is computationally intensive when operating on a typical hyperspectral image, due to the need to estimate and invert large covariance matrices. In [14], an anomaly detector was proposed based on SVDD [15], which is a single-class support vector machine (SVM) classifier that is able to directly estimate the support region for a given data set. SVDD avoids prior assumptions about the distribution of the data and models the background by a minimum enclosing hypersphere. Spectral pixels that fall outside this hypersphere model are considered anomalies. The whole process of finding the minimum enclosing hypersphere is performed in a high-dimensional feature space using the ideas of kernel function [13] for efficient implementation. As in any kernel-based detection algorithm, this high-dimensional hypersphere in the feature space will correspond to very complex decision boundaries in the original input space.

The main concept of the kernel SVDD for noise-free data is to find the smallest hypersphere in the induced feature space $\mathcal{F} = \{\Phi(\mathbf{x}) : \|\Phi(\mathbf{x}) - \mathbf{c}\|^2 < R^2\}$, that includes the entire set of mapped training examples $\Phi(\mathbf{X}) = \{\Phi(\mathbf{x}_i), i = 1, \dots, N\}$, where \mathbf{c} is the center of the hypersphere in the feature space. Therefore, we need to solve the following constrained optimization problem:

$$\min(R) \quad \text{subject to} \quad \Phi(\mathbf{x}_i) \in \mathcal{F}, \quad i = 1, \dots, N. \quad (6)$$

After applying the Lagrange multipliers α_i , for $i = 1, \dots, N$ and Karush–Kuhn–Tucker conditions (see [15] for details), the dual problem can be written as

$$\begin{aligned} \min L(\alpha_i) &= \sum_{i,j} \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) - \sum_i k(\mathbf{x}_i, \mathbf{x}_i) \\ \text{subject to} \quad &0 \leq \alpha_i \leq C \quad \forall i = 1, 2, \dots, N, \quad \sum_i \alpha_i = 1, \end{aligned} \quad (7)$$

where the parameter C controls the tradeoff between the volume of the hypersphere and the errors. The SVDD decision rule for test data \mathbf{y} can be written in terms of the kernel function as

$$\text{SVDD}(\mathbf{y}) = k(\mathbf{y}, \mathbf{y}) - 2 \sum_i \alpha_i k(\mathbf{y}, \mathbf{x}_i) + \sum_{i,j} \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) \geq R^2. \quad (8)$$

This decision rule provides a measure of similarity between the test pixel and its background. One can determine if a test pixel

is an anomaly or not by applying a threshold on the decision statistic. Figure 1(e) and (f) shows the SVDD decision contours as well as the support vectors ($\sigma = 1.0$ and $C = 0.1$) for a toy example with a single and a mixture of two Gaussian background distributions. It is clearly seen that the SVDD algorithm can capture the statistics of the whole (noise-free) data set. Similar decision statistic can also be derived for real training data (noisy samples) where a hypersphere is maximized to include most of the data, as shown in [15]. Although, SVDD and KRX are both implemented in the induced kernel feature space, but they are fundamentally different. SVDD is based on a discriminative model, which does not assume any distribution for the input data. On the other hand, KRX is a generative model that represents the data as a Gaussian distribution in the kernel induced high-dimensional feature space. Comparing Figure 1(d) and (f), it can be seen that both algorithms can equally capture the mixed background distribution with complex nonlinear decision boundaries.

SIGNATURE-BASED TARGET DETECTION

In some target detection applications, we have prior knowledge about the spectral characteristics of the desired targets. In these situations, the target spectral characteristics can be defined by a single target spectrum [16] or a target subspace [17]. Similarly, the background can be modeled statistically by a Gaussian distribution or with a subspace representing the whole or local background statistics. In this section, we review the concept of several classical target detection algorithms such as the linear spectral matched filter (SMF), matched subspace detector (MSD), adaptive subspace detector (ASD), and orthogonal subspace projection (OSP). The issues and challenges associated with their model assumptions and parameters are also discussed. A detailed discussion of the classical target detection techniques, specially SMF, and its practical implementation including the replacement model, is discussed in a companion article in this special issue [3].

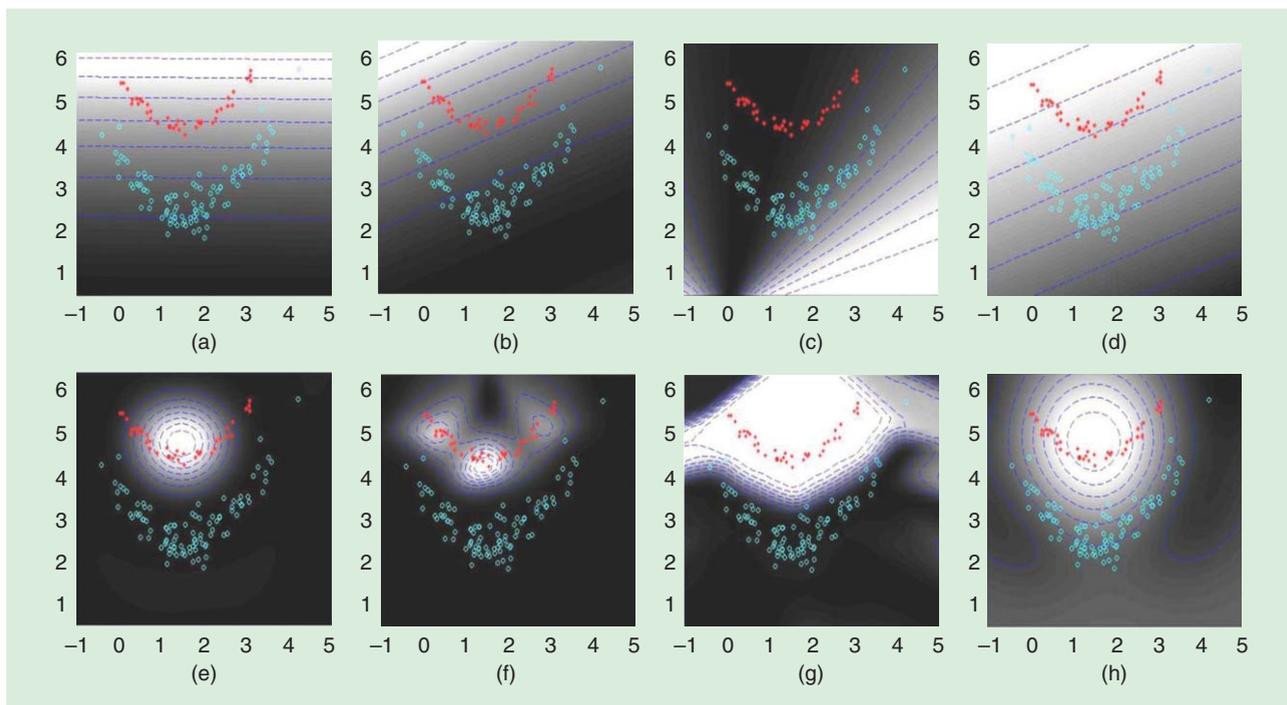
SPECTRAL MATCHED FILTER

The model for SMF is expressed by

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{n}, \quad \text{target absent} \\ H_1 : \mathbf{x} &= a\mathbf{s} + \mathbf{n}, \quad \text{target present}, \end{aligned} \quad (9)$$

where a is the unknown target abundance measure ($a = 0$ when no target is present and $a > 0$ when a target is present), $\mathbf{s} = [s_1 \ s_2 \ \dots \ s_p]^T$ is the spectral signature of the target, and \mathbf{n} is zero-mean Gaussian random additive background clutter noise. The SMF model is based on the assumption that the background clutter noise has a Gaussian distribution $\mathcal{N}(0, \hat{\mathbf{C}}_b)$ and the target distribution is also a Gaussian $\mathcal{N}(a\mathbf{s}, \hat{\mathbf{C}}_b)$ having the same covariance statistics, but with a mean of $a\mathbf{s}$, where a is a scalar abundance value representing the target strength. Then, using GLRT, the output of SMF for a test input \mathbf{x} is given in [16] as

$$D_{\text{SMF}}(\mathbf{x}) = \frac{\mathbf{s}^T \hat{\mathbf{C}}_b^{-1} \mathbf{x}}{\sqrt{\mathbf{s}^T \hat{\mathbf{C}}_b^{-1} \mathbf{s}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{SMF}}, \quad (10)$$



[FIG4] Contour and surface plots of the decision boundaries for the (a)–(d) classical target detectors and (e)–(h) their corresponding RBF kernel versions on a two-dimensional nonlinear toy data set: in this toy example, the red star-shaped and blue circle-shaped symbols represent the target and background data points, respectively. (a) SMF. (b) MSD. (c) ASD. (d) OSP. (e) KSMF. (f) KMSD. (g) KASD. (h) KOSP.

where \hat{C}_b represents the estimated covariance matrix for the centered observation data and η_{SMF} represents a threshold.

MATCHED SUBSPACE DETECTOR

In the matched subspace model, the target pixel vectors are expressed as a linear combination of target spectral signature and background spectral signature, which are represented by subspace target spectra and subspace background spectra, respectively. The hyperspectral target detection problem is expressed as two competing hypotheses H_0 and H_1

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{B}\boldsymbol{\zeta} + \mathbf{n}, \text{ target absent} \\ H_1 : \mathbf{x} &= \mathbf{S}\boldsymbol{\theta} + \mathbf{B}\boldsymbol{\zeta} + \mathbf{n}, \text{ target present,} \end{aligned} \quad (11)$$

where \mathbf{S} and \mathbf{B} represent matrices whose p -dimensional independent columns span the known target and background subspaces, respectively; $\boldsymbol{\theta}$ and $\boldsymbol{\zeta}$ are unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of \mathbf{S} and \mathbf{B} , respectively; \mathbf{n} represents Gaussian random noise ($\mathbf{n} \in \mathbb{R}^p$) distributed as $\mathcal{N}(0, \sigma_b^2 \mathbf{I})$, where σ_b is an unknown scalar value; and $[\mathbf{S} \ \mathbf{B}]$ is a concatenated matrix of \mathbf{S} and \mathbf{B} .

The GLRT for the MSD model (11) was derived in [17], is given by

$$D_{\eta_{MSD}}(\mathbf{x}) = \frac{\mathbf{x}^T (\mathbf{I} - \mathbf{P}_b) \mathbf{x}}{\mathbf{x}^T (\mathbf{I} - \mathbf{P}_{tb}) \mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{MSD}, \quad (12)$$

where $\mathbf{P}_b = \mathbf{B}\mathbf{B}^\#$ is a projection matrix associated with the background subspace $\langle \mathbf{B} \rangle$; $\mathbf{P}_{tb} = [\mathbf{S} \ \mathbf{B}][\mathbf{S} \ \mathbf{B}]^\#$ is a projection matrix associated with the target-and-background subspace $\langle \mathbf{S}\mathbf{B} \rangle$. $D_{\eta_{MSD}}(\mathbf{x})$ is compared to a threshold η_{MSD} to make a final decision about which hypothesis best relates to \mathbf{x} .

ADAPTIVE SUBSPACE DETECTOR

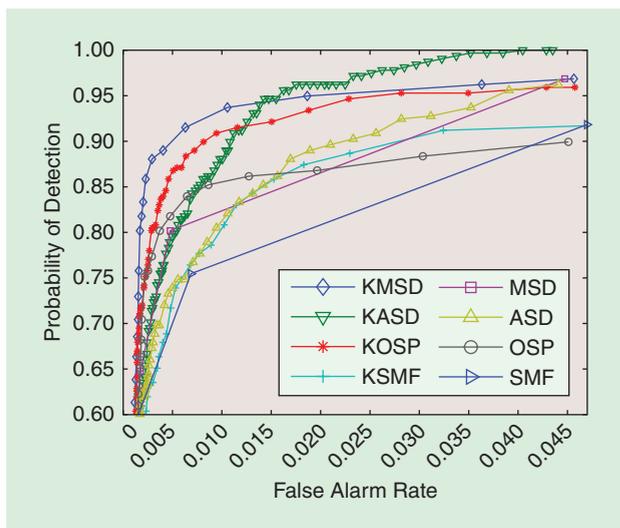
The hypotheses H_0 and H_1 for the ASD detector are

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{n}, \text{ target absent} \\ H_1 : \mathbf{x} &= \mu \mathbf{S}\boldsymbol{\theta} + \mathbf{n}, \text{ target present,} \end{aligned} \quad (13)$$

where \mathbf{S} spans the known target subspace; $\boldsymbol{\theta}$ is an unknown vector, where entries are coefficients that account for the abundances of the corresponding column vectors of \mathbf{S} ; the target signal $\mathbf{S}\boldsymbol{\theta}$ is scaled by μ (target strength); \mathbf{n} represents Gaussian random noise distributed as $\mathcal{N}(0, \sigma^2 \mathbf{C}_b)$ where \mathbf{C}_b is the background noise structure obtained from the training data, and σ represents an unknown scalar value associated with the test data.

In model (13), \mathbf{x} is assumed to be a background noise, under H_0 with $\mu = 0$ and a linear combination of a target subspace signal and a scaled background noise, distributed as $\mathcal{N}(\mu \mathbf{S}\boldsymbol{\theta}, \sigma^2 \mathbf{C}_b)$, under H_1 with $\mu > 0$. The GLRT for the problem described by (13) is given in [18] as

$$D_{\eta_{ASD}}(\mathbf{x}) = \frac{\mathbf{x}^T \hat{\mathbf{C}}_b^{-1} \mathbf{S} (\mathbf{S}^T \hat{\mathbf{C}}_b^{-1} \mathbf{S})^{-1} \mathbf{S}^T \hat{\mathbf{C}}_b^{-1} \mathbf{x}}{\mathbf{x}^T \hat{\mathbf{C}}_b^{-1} \mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{ASD}, \quad (14)$$



[FIG5] ROC curves obtained by the classical target detectors and their corresponding RBF kernel versions for the HYDICE hyperspectral FR-I image [Figure 3(a)] with several military targets. (Figure reproduced and used with permission from [20].)

where \hat{C}_b is the maximum likelihood estimate of the covariance, C_b , from the data and η_{ASD} represents a threshold. Equation (14) has a CFAR property and is also referred to as the CFAR ASD for subspace or multirank CFAR ASD in [18]. When the signal is coherent (S is rank one) (14) is referred to as the adaptive coherence/cosine estimator (ACE), because (14) measures the \cos^2 of the angle between the whitened \tilde{x} and $\langle \tilde{S} \rangle$, where $\tilde{x} = \hat{C}_b^{-1/2} x$ and $\tilde{S} = \hat{C}_b^{-1/2} S$.

ORTHOGONAL SUBSPACE PROJECTION

The OSP algorithm [19] is based on maximizing the signal-to-noise ratio (SNR) of the target data in the subspace orthogonal to the background subspace. Rewriting the linear mixture model in terms of the desired target s and the known background subspace B as

$$x = \alpha_t + B\xi + n, \quad (15)$$

where the columns of B are the undesired background endmembers spectra, ξ is an unknown column vector with elements that are the abundances associated with the background endmembers, α_t is the unknown abundance measure associated with the target spectrum, and n is an additive noise. The output of the OSP classifier is given in [19] as

$$D_{\text{OSP}} = q_{\text{OSP}}^T x = s^T P_b^\perp x, \quad (16)$$

where $q_{\text{OSP}}^T = s^T P_b^\perp$ is the OSP operator consisting of a background spectral signature rejecter $P_b^\perp = (I - BB^{\#})$ followed by a matched filter s .

KERNEL-BASED TARGET DETECTORS

The classical target detection techniques described so far are based on first- and second-order statistics and do not exploit higher-order statistics (nonlinearities). Similar to the kernel-based anomaly detector, all the above classical target detectors

can be extended to their corresponding nonlinear versions through the use of kernel-based machine learning [13]. The implicit exploitation of nonlinear features through kernels provides crucial information about a given data, which the learning methods based on linear models cannot achieve in general. The classical target detectors, discussed in the previous subsection, have all been extended to their nonlinear (kernel) versions. A comparative review of kernel-based target detectors and their performance compared with the corresponding linear versions is given in [20].

Figure 4 shows contour and surface plots of decision boundaries for the classical target detectors and their kernel versions on a two-dimensional nonlinear toy data set. In the contour and surface plots, data points for the desired target were represented by the star-shaped symbols and the background points were represented by light blue circles. As shown in Figure 4, the contours generated by the kernel-based detectors are highly nonlinear and naturally capture the dispersion of the data and more successfully separate the two classes, compared to the linear contours obtained by the classical target detectors. Therefore, the kernel-based detectors clearly provide significantly improved discrimination over the conventional detectors for non-Gaussian data. Figure 5 illustrates the receiver operating characteristic (ROC) curves (probability of detection versus probability of false alarm) obtained by the classical target detectors and their corresponding kernel versions for the HYDICE hyperspectral FR-I image. Typically, the kernel versions of the target detectors outperform the linear versions, as seen in Figure 5. Among the kernel-based detectors, kernel MSD (KMSD) and kernel ASD (KASD) outperform kernel OSP (KOSP) and kernel SMF (KSMF), mainly because targets in KMSD and KASD are better represented by the associated target subspace than by a single target spectral signature used in KOSP and KSMF.

Another kernel-based binary classifier, the SVM [13], has been extensively used for hyperspectral image classification, where some labeled training data from the hyperspectral image itself is used to train the SVM classifier. However, SVM has not been very popular for target detection applications mainly due to the lack of data, especially for the target spectrum. For the background clutter spectra, there is often a period of target-free data collection. One possible approach to obtain sufficient training data for the target spectrum is to synthesize a large number of target spectral signatures using the MODTRAN atmospheric modeling tool as is done in [22] for the design of an invariant matched subspace detector.

One advantage of kernel methods is that by using an appropriate composite kernel [23], the contextual information from multiple pixels on the target can be incorporated directly into the classifier or the kernel detector [24] without the need for any postprocessing [see the ROC curve in Figure 6 for an SVM target detector with a composite kernel (SVM-CK)].

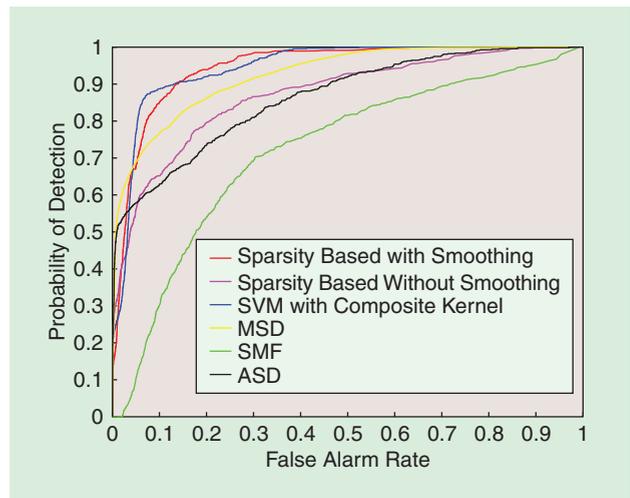
CONTINUUM FUSION DETECTORS

One of the major problems with the target detection models is the variability of the model parameters in practical applications. For

example, in the case of the spectral matched filter, the laboratory target spectral signature used for target detection could be significantly different in the measured test image. One approach is to model the target variation by a probability distribution (Gaussian) with unknown parameters (mean and variance) and then use GLRT to obtain the best-guess estimate of the unknown parameters. Another approach is to use the continuum fusion (CF) detector proposed in [21] and [25]. As an example, to illustrate the concept of CF, a particular flavor of CF detector constructs optimal CFAR-SMF detectors for every known value of target parameter (mean) and then perform a CF-based fusion of the decision boundaries of these detectors into a single fused decision boundary. The toy example in Figure 7(a) depicts the hyperplanar decision boundaries for a CFAR-SMF detector for every allowed position of the target signature on a one-dimensional affine subspace (fixed straight line), shown as red balls. The blue ball represents a Gaussian clutter distribution that has been whitened with the covariance matrix estimated from the background pixels. Figure 7(b) shows the continuum CFAR-fused decision boundary obtained by the union of all the SMF detectors critical regions. The critical region of a CFAR-SMF detector is defined as the area to the right of its decision boundary (see Figure 7). Figure 7(c) illustrates the GLRT decision boundaries and the CFAR continuum fusion decision boundaries for several different threshold values. It should be pointed out that the CFAR-based SMF continuum fusion discussed above is only one possible flavor of CF. The real flexibility of CF arises from the infinite number of ways that individual decision boundaries from different type of detectors can be fused [21].

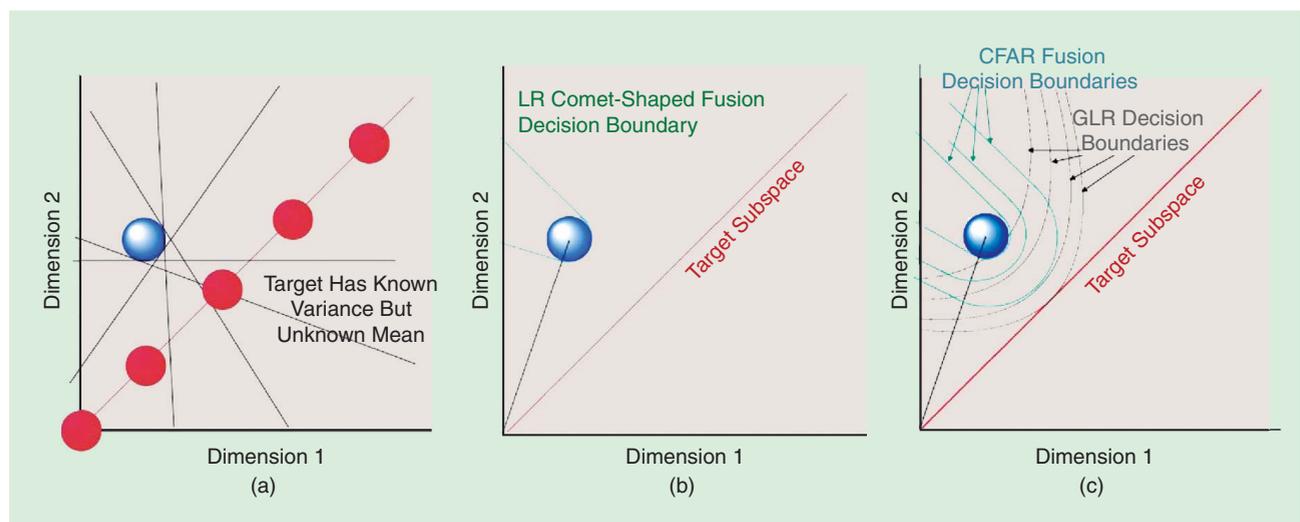
CHALLENGES ASSOCIATED WITH TARGET DETECTORS

In this section, we discuss issues associated with the model assumptions and their implementations for the classical target detection techniques. For example, as was discussed in previous

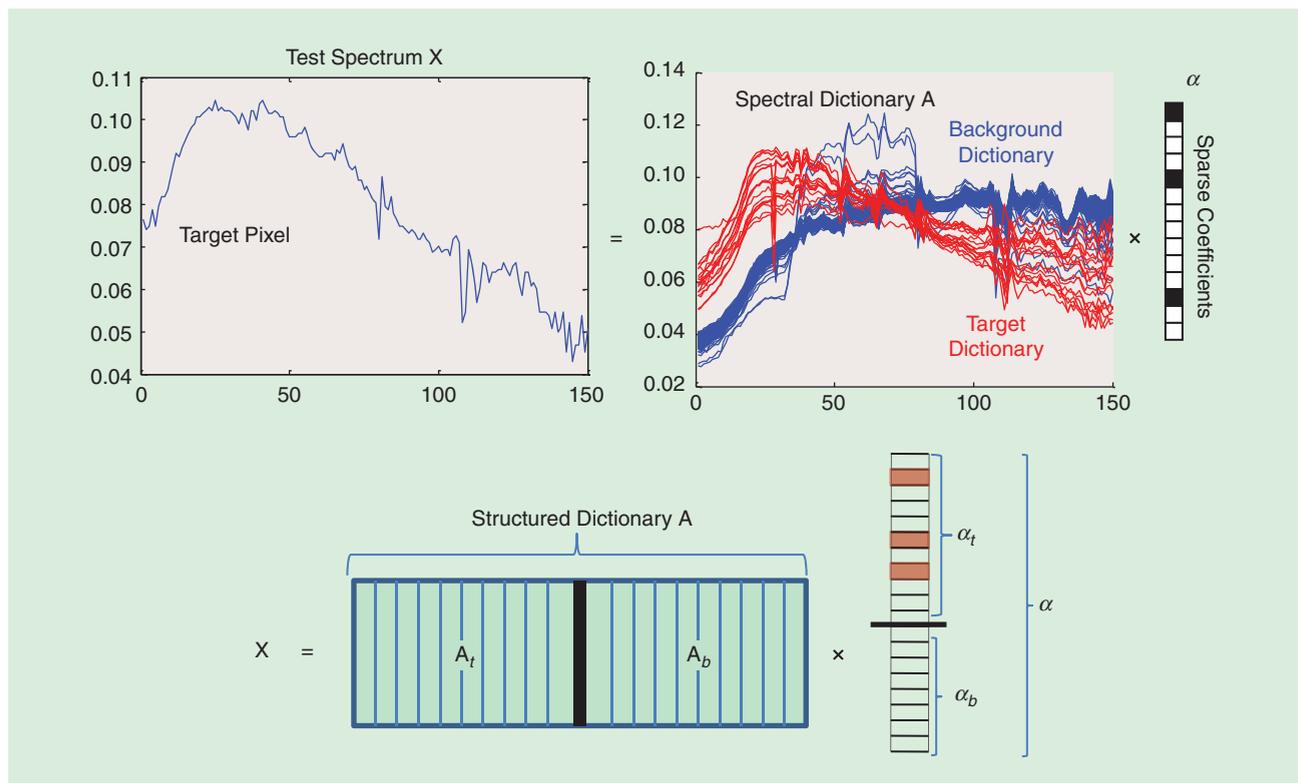


[FIG6] ROC curves for the HYDICE hyperspectral FR-I image [Figure 3(a)] with several military targets. (Figure reproduced and used with permission from [34].)

sections, in the case of SMF or OSP, a fixed laboratory target signature is assumed in the model, which could be significantly different from the actual target test pixels in the scene. This difference involves understanding a myriad of interactions of the target signal with other pixels (adjacency effects) and variations in illumination that depend on many parameters with unknown values. These include the relative contributions of the sky and sun and the amounts of atmospheric gases, vapors, and aerosols that alter spectral signals before being acquired by the sensor. For this reason, researchers have proposed using a preprocessing stage to estimate and compensate for the atmospheric effects on the data to transform the known target spectrum and measurement data into a common domain where a target detection algorithm can be applied.



[FIG7] (a) Hyperplanar decision boundaries for a CFAR SMF solution to the affine target subspace problem, (b) continuum-fused CFAR decision surface for the affine target subspace problem, and (c) GLRT versus CFAR decision surfaces (at different threshold values) for the affine target subspace problem. (Figure reproduced and used with permission from [21].)



[FIG8] The SRC process for target detection. It shows a structured dictionary consisting of a target and a background subdirectories.

On the other hand, some researchers have proposed to use multiple target signatures to try to capture the variability in the target signature. By using physics-based modeling and the MODTRAN atmospheric-modeling tool [26], meaningful target spectral signatures can be generated that can capture the target signature appearance over a wide range of atmospheric conditions. But, these target signatures still need to be represented by a subspace as required in the MSD and ADS models. For example, in [22] and [27], a large number of synthetic target (or background) signatures using MODTRAN under various atmospheric conditions were generated, which were then used to construct an invariant target (or background) subspace by retaining only the significant eigenvectors of the target (or background) covariance matrix with eigenvalues above the sensor noise level.

Furthermore, an important consideration in SMF and ASD implementation is the potential ill conditioning of the covariance matrix due to the high correlation, high dimensionality of the hyperspectral data, and a limited background sample size [28], [29]. Representing the inverse covariance matrix in terms of its eigenvector-eigenvalue decomposition, it becomes clear that the behavior of the inverse covariance matrix depends heavily on the small eigenvalues, which could render it unstable. For example, to reduce SMF sensitivity to statistical and numerical errors, eigenvectors corresponding to eigenvalues below an appropriate condition number or estimated sensor noise level can be discarded or a scaled identity matrix ($\beta\mathbf{I}$) is

added to the background clutter covariance before inverting, which is equivalent to including an l_2 -norm smoothness penalty (regularization) term on the matched filter coefficients during the design of the SMF [30]. A regularized matched filter is given by

$$D_{\text{RegSMF}}(\mathbf{x}) = \frac{\mathbf{s}^T (\hat{\mathbf{C}}_b + \beta\mathbf{I})^{-1} \mathbf{x}}{\sqrt{\mathbf{s}^T (\hat{\mathbf{C}}_b + \beta\mathbf{I})^{-1} \mathbf{s}}}, \quad (17)$$

where β is a regularization parameter related to the sensor noise and \mathbf{I} is the identity matrix. The regularization process acts as a Gaussian prior on the matched filter coefficients. As shown in [30], for small β values, the SMF filter coefficients can have extreme negative and positive values, and at large β values, the filter coefficients tend to become less oscillatory and smoother.

SPARSE REPRESENTATION FOR TARGET DETECTION

Sparse representation classifiers (SRCs) express a signal as a linear combination of very few atoms from an overcomplete dictionary consisting of a set of training data from all the classes. The resulting sparse code can reveal the class information if signals from different classes lie in different subspaces.

SPARSITY MODEL

In the sparsity-based HSI classification model, the spectral signatures of pixels belonging to the same class are assumed to approximately lie in a low-dimensional subspace. Suppose we

have K distinct classes and the k th class has N_k training samples $\{\mathbf{a}_j^k\}_{j=1, \dots, N_k}$. Let \mathbf{x} be a p -dimensional hyperspectral pixel observation. If \mathbf{x} belongs to the k th class, then its spectrum approximately lies in a low-dimensional subspace spanned by the training samples in the k th class.

Based on the above sparsity assumption, in SRC [31] an unknown test sample is modeled to lie in the union of the K subspaces associated with the K classes. By combining the class subdictionaries $\{\mathbf{A}^k\}_{k=1, \dots, K}$, a test sample \mathbf{x} can be written as a sparse linear combination of all the training samples as

$$\begin{aligned} \mathbf{x} &= \mathbf{A}^1 \boldsymbol{\alpha}^1 + \mathbf{A}^2 \boldsymbol{\alpha}^2 + \dots + \mathbf{A}^K \boldsymbol{\alpha}^K \\ &= \underbrace{[\mathbf{A}^1 \ \dots \ \mathbf{A}^K]}_{\mathbf{A}} \underbrace{\begin{bmatrix} \boldsymbol{\alpha}^1 \\ \vdots \\ \boldsymbol{\alpha}^K \end{bmatrix}}_{\boldsymbol{\alpha}} = \mathbf{A}\boldsymbol{\alpha}, \end{aligned} \quad (18)$$

where \mathbf{A} is a $p \times N$ structured dictionary consisting of training samples from all classes with $N = \sum_{k=1}^K N_k$ and $\boldsymbol{\alpha}$ is an N -dimensional sparse vector formed by concatenating the sparse vectors $\{\boldsymbol{\alpha}^k\}_{k=1, \dots, K}$ associated with each class subdictionary. Note that, ideally, if \mathbf{x} belongs to the k th class, then $\boldsymbol{\alpha}^j = 0, \forall j = 1, \dots, K, j \neq k$.

Given a structured dictionary \mathbf{A} , the unknown sparse vector $\boldsymbol{\alpha}$ satisfying $\mathbf{x} = \mathbf{A}\boldsymbol{\alpha}$ can be obtained from the following sparsity-driven optimization:

$$\hat{\boldsymbol{\alpha}} = \arg \min \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \|\mathbf{x} - \mathbf{A}\boldsymbol{\alpha}\|_2 \leq \epsilon_0, \quad (19)$$

where $\|\boldsymbol{\alpha}\|_0$ denotes ℓ_0 -norm, which is defined as the number of nonzero entries in the vector $\boldsymbol{\alpha}$ (in the Bayesian learning framework it represents a sparsity prior on $\boldsymbol{\alpha}$), and ϵ_0 is an approximation error tolerance. The aforementioned optimization problem (19) can be approximately solved by any greedy pursuit algorithm [32]. Furthermore, the optimization problem (19) is NP-hard. But, it can be relaxed by replacing the ℓ_0 -norm with the ℓ_1 -norm $\|\boldsymbol{\alpha}\|_1 = \sum_{i=1}^N |\alpha_i|$ and solved by an standard convex optimization technique [33].

The class label of \mathbf{x} can be obtained directly from the characteristics of the recovered sparse vector $\hat{\boldsymbol{\alpha}}$, which is determined by the class that gives the minimal residual error

$$r^k(\mathbf{x}) = \arg \min_{k=1, \dots, K} \|\mathbf{x} - \mathbf{A}^k \hat{\boldsymbol{\alpha}}^k\|_2, \quad (20)$$

where $\hat{\boldsymbol{\alpha}}^k$ denotes the portion of the recovered sparse coefficients corresponding to the atoms in the k th class.

In the case of target detection, typically the dictionary for the SRC consists of the training samples from the target and background subdictionaries represented by $\mathbf{A} = [\mathbf{A}_t \ \mathbf{A}_b]$, as shown in Figure 8. The sparse representation vector $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_t^T \ \boldsymbol{\alpha}_b^T]^T$ satisfying $\mathbf{x} = \mathbf{A}\boldsymbol{\alpha}$ can be obtained by solving the previous optimization problem (19), where $\boldsymbol{\alpha}_t$ and $\boldsymbol{\alpha}_b$ represent the sparse coefficient vectors corresponding to the target and background dictionaries, respectively. Once the sparse coefficient vector $\hat{\boldsymbol{\alpha}}$ is obtained, the class of the test pixel \mathbf{x} can be determined by comparing the residuals $r_t(\mathbf{x}) = \|\mathbf{x} - \mathbf{A}_t \hat{\boldsymbol{\alpha}}_t\|_2$ and $r_b(\mathbf{x}) = \|\mathbf{x} - \mathbf{A}_b \hat{\boldsymbol{\alpha}}_b\|_2$.

STRUCTURED SPARSE PRIORS

The classification performance of SRC can be improved by incorporating the contextual information from the neighboring pixels into the SRC classifier. If we have multiple pixels on the target, the contextual information can be incorporated into the SRC algorithm by imposing richer sparsity priors (structured sparsity priors) or structural penalty constraints in the SRC optimization. One of the simplest approaches is to use a joint sparsity (collaborative) model, as shown in [35], assuming that the underlying sparse vectors associated with the neighboring pixels share a common sparsity pattern. In the joint sparsity model, a block of T pixels $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_T]$ can be jointly represented for a given dictionary as

$$\hat{\boldsymbol{\Omega}} = \arg \min \|\boldsymbol{\Omega}\|_{\text{row},0} \quad \text{subject to} \quad \|\mathbf{X} - \mathbf{A}\boldsymbol{\Omega}\|_F^2 \leq \epsilon_2, \quad (21)$$

where $\|\cdot\|_F$ represents Frobenius norm, the notation $\|\boldsymbol{\Omega}\|_{\text{row},0}$ denotes the number of nonzero rows of $\boldsymbol{\Omega}$ and ϵ_2 is an error tolerance. Simultaneous greedy algorithms [32] can be used to obtain an approximation to (21). The row-sparsity norm $\|\boldsymbol{\Omega}\|_{\text{row},0}$ can also be replaced by $\|\boldsymbol{\Omega}\|_{1,2}$, which is an $l_{1,2}$ -norm defined as the sum of the l_2 -norms of the rows of $\boldsymbol{\Omega}$, to convert the NP-hard problem (21) into a convex optimization task.

Figure 6 shows the ROC curves for several target detectors on the HYDICE hyperspectral FR-I image. As seen in Figure 6, the SVM-CK and sparsity based with an l_2 -norm smoothing penalty on the reconstructed data outperform the classical signature-based target detection techniques. The experimental implementation for each algorithm and results for couple of other hyperspectral images can be found in [34].

DICTIONARY CONSTRUCTION

Another aspect of the problem that requires careful attention is how to construct appropriate dictionaries \mathbf{A}_b and \mathbf{A}_t . Global dictionaries for target and background can be designed using a given training data. However, in target detection applications, there is usually a lack of training data, especially for the target class. By using physical models and the MODTRAN atmospheric modeling program [26], target spectral signatures can be generated that can capture the target signature appearance over a wide range of atmospheric conditions [22], [27]. These synthetically generated spectral signatures can then be used to construct a redundant target dictionary for the SRC classifier, which could be invariant to environmental variations. The background dictionary is often modeled by randomly selecting some pixels from the test image itself. Furthermore, given sufficient training samples for the background and target classes, a dictionary design technique such as the K-SVD algorithm [32], which alternately minimizes sparsity of the representation and updates the dictionary's atoms to better fit the data, can be used to construct dictionaries used by SRC.

CONCLUSIONS

This article presents a brief overview of recent target detection techniques in HSI processing. The major challenges in target

detection techniques are still the need for developing more robust classification techniques, constructing better models, and designing more effective preprocessing methods that can detect targets in very difficult scenarios such as subpixel targets, camouflaged targets, and hidden targets underneath trees. Estimating the model parameters with limited training data is also a major obstacle in developing robust target detectors. The usefulness of using physics-based synthetically generated spectral signatures to supplement the real data for designing better detection algorithms are currently under investigation by many researchers. More investigation is needed to fully evaluate the performance of the machine-learning-based target detection techniques from laboratory experiments to real field trials. Fusion of multiple target detection algorithms as well as using hyperspectral data to complement other sensors are still being investigated. Kernel-based target detection techniques have shown improvement over the linear techniques, but there is definitely a need to further study nonlinear approaches. We have also not reported the effect of band selection, dimensionality reduction, or compressed sensing techniques on the performance of target detection methods, however, some primary results can be found in [36].

AUTHOR

Nasser M. Nasrabadi (nasser.m.nasrabadi.civ@mail.mil) received the B.Sc. (Eng.) and Ph.D. degrees in electrical engineering from the Imperial College of Science and Technology (University of London), London, United Kingdom, in 1980 and 1984, respectively. He is currently a senior research scientist with the U.S. Army Research Laboratory. He is a Fellow of the IEEE and SPIE. His current research interests are in hyperspectral imaging, automatic target recognition, and statistical machine-learning theory.

REFERENCES

- [1] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.
- [2] M. T. Eismann, A. D. Stocker, and N. M. Nasrabadi, "Automated hyperspectral cueing for civilian search and rescue," *Proc. IEEE*, vol. 97, no. 6, pp. 1031–1055, June 2009.
- [3] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Hyperspectral target detection," *IEEE Signal Processing Mag.*, vol. 31, no. 1, pp. 24–33, Jan. 2014.
- [4] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 5–27, July 2010.
- [5] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Processing Mag.*, vol. 19, pp. 58–69, Jan. 2002.
- [6] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.
- [7] A. P. Schaum, "Joint subspace detection of hyperspectral targets," in *Proc. IEEE Aerospace Conf., Big Sky, MT*, Mar. 2004, vol. 3, pp. 1818–1824.
- [8] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York: Springer, 2001.
- [9] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA: SIAM, 1998.
- [10] H. Goldberg and N. M. Nasrabadi, "A comparative study of linear and nonlinear anomaly detectors for hyperspectral imagery," in *Proc. SPIE, Orlando, FL*, Apr. 2007, vol. 6565, Article ID 656504.
- [11] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing*, vol. 51, no. 11, pp. 354–379, Nov. 2012.
- [12] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 2, pp. 388–397, Feb. 2005.
- [13] B. Schölkopf and A. J. Smola, *Learning with Kernels*. Cambridge, MA: MIT Press, 2002.
- [14] A. Banerjee, P. Burlina, and C. Diehl, "A support vector method for anomaly detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 44, no. 8, pp. 2282–2291, Aug. 2006.
- [15] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Mach. Learn.*, vol. 54, no. 1, pp. 45–66, Jan. 2004.
- [16] F. C. Robey, D. R. Fuhrmann, and E. J. Kelly, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.
- [17] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.
- [18] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Processing*, vol. 49, no. 1, pp. 208–216, Jan. 2001.
- [19] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, no. 4, pp. 779–785, July 1994.
- [20] H. Kwon and N. M. Nasrabadi, "A comparative analysis of kernel subspace target detectors for hyperspectral imagery," *EURASIP J. Advances Signal Process.*, vol. 2007, no. 1, Jan. 2007, Article ID 29250.
- [21] A. P. Schaum and B. J. Daniel, "Continuum fusion methods of spectral detection," *Opt. Eng.*, vol. 5, no. 2, pp. 1–11, Nov. 2012.
- [22] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 6, pp. 2706–2717, Nov. 1999.
- [23] G. Camps-Valls, L. Gomez-Chova, J. Muñoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.
- [24] C. Chen, N. M. Nasrabadi, and T. D. Tran, "Kernel sparse representation for hyperspectral target detection," in *Proc. SPIE, Baltimore, MD*, May 2012, vol. 8390, Article ID 839005.
- [25] A. P. Schaum, "Continuum fusion, a theory of inference, with applications to hyperspectral detection," *Opt. Express*, vol. 18, no. 8, pp. 8171–8181, Nov. 2010.
- [26] A. Berk, G. P. Anderson, L. S. Bernstein, P. K. Acharya, H. Dothe, M. W. Matthew, S. M. Adler-Golden, J. H. Chetwynd, S. C. Richtsmeier, B. Pukall, C. L. Allred, L. S. Jeong, and M. L. Hoke, "MODTRAN4: Radiative transfer modeling for atmospheric correction," in *Proc. SPIE, Orlando, FL*, 1999, vol. 3756, pp. 348–353.
- [27] E. J. Ientilucci and P. Bajorski, "Stochastic modeling of physically derived signature spaces," *J. Appl. Remote Sensing*, vol. 2, no. 1, pp. 023532(1–10), Aug. 2008.
- [28] J. Theiler, G. Cao, L. R. Bachegea, and C. A. Bouman, "Sparse matrix transform for hyperspectral imagery processing," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 3, pp. 424–437, Mar. 2011.
- [29] C. E. Cafer, J. Silvermana, O. Orththal, D. Antonelli, Y. Sharonib, and S. R. Rotman, "Improved covariance matrices for point target detection in hyperspectral data," *Opt. Eng.*, vol. 47, no. 7, pp. 076402(1–13), July 2008.
- [30] N. M. Nasrabadi, "Regularized spectral matched filter for target recognition in hyperspectral imagery," *IEEE Signal Processing Lett.*, vol. 15, pp. 317–320, 2008.
- [31] J. Wright, A. Y. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [32] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal Processing*. New York: Springer, 2010.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [34] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 3, pp. 629–640, Mar. 2011.
- [35] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Simultaneous joint sparsity model for target detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing Lett.*, vol. 8, no. 4, pp. 676–680, July 2011.
- [36] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Effects of linear projections on the performance of target detection and classification in hyperspectral imagery," *J. Appl. Remote Sensing*, vol. 5, no. 1, pp. 053263(1–25), Nov. 2011.

[Gustavo Camps-Valls, Devis Tuia, Lorenzo Bruzzone, and Jón Atli Benediktsson]

Advances in Hyperspectral Image Classification



[Earth monitoring with statistical learning methods]

The technological evolution of optical sensors over the last few decades has provided remote sensing analysts with rich spatial, spectral, and temporal information. In particular, the increase in spectral resolution of hyperspectral images (HSIs) and infrared sounders opens the doors to new application domains and poses new methodological challenges in data analysis. HSIs allow the characterization of objects of interest (e.g., land-cover classes) with unprecedented accuracy, and keeps inventories up to date. Improvements in spectral resolution have called for advances in

signal processing and exploitation algorithms. This article focuses on the challenging problem of hyperspectral image classification, which has recently gained in popularity and attracted the interest of other scientific disciplines such as machine learning, image processing, and computer vision. In the remote sensing community, the term classification is used to denote the process that assigns single pixels to a set of classes, while the term segmentation is used for methods aggregating pixels into objects and then assigned to a class.

One should question, however, what makes HSI so distinctive. Statistically, HSIs are not extremely different from natural grayscale and color photographic images (see [1, Ch. 2]). Grayscale images are spatially smooth: the joint probability density function

Digital Object Identifier 10.1109/MSP.2013.2279179

Date of publication: 5 December 2013

(PDF) of the luminance samples is highly uniform, the covariance matrix is highly nondiagonal, the autocorrelation functions are broad, and have generally a $1/f$ band-limited spectrum. In the case of color images, the correlation between the tristimulus values of the natural colors is typically high. While the three tristimulus channels are equally smooth in generic red/green/blue (RGB) representations, opponent representations imply an uneven distribution of bandwidth between channels. Despite all these commonalities, the analysis of HSI turns out to be more difficult, especially because of the high dimensionality of the pixels, the particular noise and uncertainty sources observed, the high spectral redundancy, and the typically nonlinear relations observed between spectral channels as well as with the corresponding material. Such nonlinearities can be related to a plethora of factors, including the multiscattering in the acquisition process, the heterogeneities at subpixel level, as well as the impact of atmospheric and geometric distortions. These characteristics of the imaging process lead to distinct nonlinear feature relations, i.e., pixels lie in high-dimensional complex manifolds. The high spectral sampling of HSI (the bands usually cover narrow portions of the electromagnetic spectrum, typically 5–10 nm) also leads to strong collinearity issues. Finally, the spatial variability of the spectral signature increases the internal class variability. All of these factors, in conjunction to the few labeled examples typically available, make HSI image classification a very challenging problem. As a result, the accuracy obtained with standard parametric classifiers commonly used for multispectral image classification is typically compromised when applied to HSI [2].

Many of these limitations have been recently addressed under the framework of statistical learning theory (SLT) [3]. SLT is a general framework for learning functions from data, which reduces to finding a linear function defined in a high- (eventually infinite) dimensional Hilbert feature space $f \in \mathcal{H}$ that learns the relation between observed input-output data pairs $(x_1, y_1), \dots, (x_\ell, y_\ell) \in \mathcal{X} \times \mathcal{Y}$, and that generalizes well. Generalization is the capability of a method to extrapolate to unseen situations, i.e., the function f should accurately predict the label $y^* \in \mathcal{Y}$ for a new input example $x^* \in \mathcal{X}$. Generalization has recurrently appeared in statistics literature for decades under the names of *bias-variance dilemma*, *capacity control*, or *complexity regularization tradeoff*. The underlying idea is to constrain too flexible functions to avoid overfitting the training data.

The SLT framework formalizes this intuition [3] and seeks for prediction functions that optimize a functional \mathcal{L}_{reg} that takes into account both an empirical estimation of the training error (loss), \mathcal{L}_{emp} , and an estimate of the complexity of the model (or regularizer), $\Omega(f)$:

$$\begin{aligned} \mathcal{L}_{\text{reg}} &= \mathcal{L}_{\text{emp}} + \lambda \Omega(f) \\ &= \sum_{i=1}^{\ell} V(x_i, y_i, f(x_i)) + \lambda \Omega(f), \end{aligned}$$

where V is a loss function acting on the ℓ labeled samples, and λ is a tradeoff parameter between the cost and the regularization. Different losses and regularizers can be adopted for solving the problem, involving completely different families of models

and solutions. To ensure unique solutions, many SLT algorithms use strictly convex loss functions. The regularizer $\Omega(f)$ limits the capacity of the classifier to minimize \mathcal{L}_{emp} and favors smooth functions.

The hyperspectral image processing community has contributed to the design of specific loss functions and regularizers to take the most out of the acquired images. For example, regularization appears explicitly in many HSI classifiers when trying to impose the spatial homogeneity of images, when including the wealth of user's labeling in active learning (AL), or when exploiting the information contained in the unlabeled pixels to better describe the image manifold in semisupervised learning. Classifiers should also be robust to changes in the image representation: small perturbations of pixels and objects in the image manifold should not produce big differences in the classification. This is why the inclusion of proper image representations and invariances is also an active field.

ADVANCED REGULARIZED IMAGE CLASSIFICATION

Before HSI, most of the classifiers used in remote sensing were parametric, such as Gaussian maximum likelihood or linear discriminant analysis. These methods, based on the estimate of the covariance matrix, were successful when dealing with early multispectral images, whose dimensionality was usually comprised between four and ten bands. HSI changed the rules, as the increased dimensionality of pixels raised to hundreds. Standard parametric methods became either unfeasible or unreliable, since estimating the class-covariance matrices requires many labeled samples, which are usually not available. For that reason, research turned to include regularization, either explicitly through Tikhonov's terms in the involved covariance matrices, or by performing classification in a subspace of reduced dimensionality [4], [5].

Although successful, parametric models make strong assumptions about the normality of the class conditional PDFs or about the linearity of the problem. Due to the complexity of HSI, such assumptions rarely hold, which encouraged research towards nonparametric and nonlinear models. Several approaches have been introduced in the last decade in the field of hyperspectral image classification: kernel methods and support vector machines (SVMs) [2], sparse multinomial logistic regression [6], neural networks [7], and Bayesian approaches like relevance vector machines [8] and Gaussian processes classification [9]. Nevertheless, the SVM has undoubtedly become the most widely used method in HSI classification research [10]. Unlike other nonparametric approaches, such as regularized radial basis function (RBF) neural networks, SVM naturally implements regularization through the concept of maximum margin: given a linear classification function $f(x) = w^\top x + b$, maximizing the linear separability between classes is equivalent to minimize the ℓ_2 -norm of model weights w used as regularizer, $\Omega(f) = \|w\|_2^2$. Nonlinearity is also implemented via reproducing kernels, which allows the ability to work in high-dimensional Hilbert spaces implicitly, while still resorting to linear algebra operations [2].

In spite of those desirable properties, the effectiveness of SVM rapidly turned out to be insufficient to exploit the rich information

contained in HSI. In all methodologies that follow, the functional to be optimized will consider additional information such as the one contained in unlabeled samples, ancillary data, or distinct signal characteristics. Such heterogeneous information is typically included in the HSI classifiers through additional regularizers.

REGULARIZATION WITH UNLABELED SAMPLES

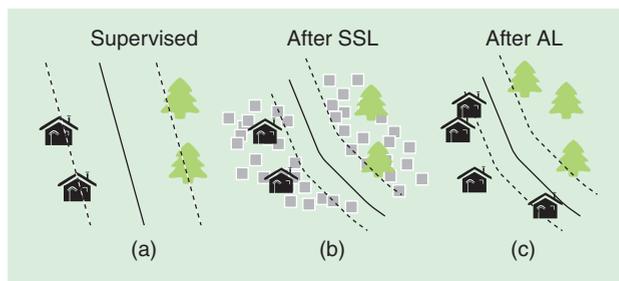
Recently, researchers started to exploit the abundant unlabeled information contained in the image itself, and new forms of regularization and priors were introduced. This is the field of semisupervised learning (SSL) [11]–[13] [see Figure 1(b)], where the minimization functional is modified to take into account the structure of the hyperspectral image manifold. Usually, semisupervised algorithms modify the decision function of the classifier by adding an extra regularization term Ω_u that acts on both labeled and unlabeled examples

$$\mathcal{L}_{\text{reg}} = \mathcal{L}_{\text{emp}} + \lambda \Omega(f) + \lambda_u \Omega_u(f).$$

Several strategies to design the regularizer have been presented. One may use the graph Laplacian as a metric on the predictions to build Ω_u . Since regularization is performed on a proximity graph, the assumption enforced is that decisions on neighboring pixels in the data manifold should be similar [14]. Another possibility considers regularizers that enforce wide and empty SVM margins [11]. Other strategies deform the kernel function by changing the metric induced using the unlabeled samples [13], [15]. Table 1 summarizes the different approaches in semisupervised learning.

We evaluate the performance of semisupervised algorithms in an Airborne Visible Infrared Imaging Spectrometer (AVIRIS) image acquired over the Kennedy Space Center (KSC), in Florida in 1996, with a total of 224 bands of 10-nm bandwidth with center wavelengths from 400 to 2,500 nm. The data was acquired from an altitude of 20 km and has a spatial resolution of 18 m. After removing low signal-to-noise ratio (SNR) bands and water absorption bands, a total of 176 bands remains for the analysis. We merged some subclasses to deal with a more balanced problem of ten classes. The high dimensionality and number of classes and subclasses pose challenging problems for the classifiers, especially when few labeled examples are available.

Figure 2 illustrates classification results for cluster kernels, probabilistic mean map kernel, label propagation, Laplacian SVM (LapSVM), and semisupervised neural networks (SSNNs).



[FIG1] Regularization of models with unlabeled samples: (a) purely supervised solution, (b) semisupervised solution exploiting low-density areas, and (c) active learning solution, where three new samples are labeled by a user.

We used $\ell = 200$ labeled pixels (20 per class) and $u = 1,000$ unlabeled pixels. LapSVM, cluster kernels, and mean map kernels perform similarly, and all improve the results of the label propagation whose training was particularly difficult in this high-dimensional setting. More homogeneous areas and better classification maps are observed in general for the mean map and bag kernels, and particularly for the SSNN, which efficiently deals with complex marsh areas [bottom right part of the classification map in Figure 2(g)] and cope with large scale data sets.

REGULARIZATION VIA USER'S INTERACTION

Another possibility to cope with small sample problems is to provide additional labeled examples. This is possible since HSI represent land surfaces, usually physically reachable or that can be displayed in an image processing software. Therefore, the new samples can be collected either by photointerpretation of the images (only if the classes can be recognized on screen) or by organizing field campaigns. However, since providing additional samples is costly, the samples to be labeled must be selected carefully. To this end, AL [16], [17] [see Figure 1(c)] has gained popularity in recent years: rather than proceeding by random sampling or stratification (i.e., sampling according to a measure of the expected variability within a class), AL uses the outcome of the current model to rank the unlabeled pixels according to their expected importance for future labeling. The aim is to detect the most difficult (and diverse) pixels for the current classifier. The top-ranked pixels are then screened by a human operator, who provides the labels, enlarging the training set. With the enlarged training set, a new improved classifier is built and the process is

[TABLE 1] SUMMARY OF SEMISUPERVISED ALGORITHMS USED IN HSI CLASSIFICATION.

ASSUMPTION	MODEL	IDEA
LOW DENSITY	TSVM [11]	LOOK FOR THE EMPTIEST MARGIN.
MANIFOLD	LABEL PROPAGATION [12]	SPREAD CLASS INFORMATION ON THE GRAPH OF LABELED/UNLABELED (NEARBY SAMPLES ARE CLASSIFIED IN THE SAME CLASS).
	LapSVM [14]	SVM HINGE LOSS PLUS LAPLACIAN EIGENMAPS FOR MANIFOLD REGULARIZATION: PIXELS CLOSE IN THE INPUT SPACE ARE ALSO CLOSE IN THE GRAPH (NEARBY SAMPLES ARE MAPPED CLOSE TOGETHER).
	SSNN [7]	NEURAL NETWORK TRAINED WITH GRADIENT DESCENT REPLACES SVM, GRAPH REGULARIZATION WITH LOSS THAT FORCES SIMILAR PIXELS TO BE MAPPED CLOSELY AND DISSIMILAR ONES TO BE SEPARATED.
CLUSTER	CLUSTER KERNEL [13]	INCREASE THE SIMILARITY MEASURE (KERNEL) IF SAMPLES FALL IN THE SAME CLUSTER, THEN RUN STANDARD SVM.
	MEAN MAP KERNEL [15]	INCREASE SIMILARITY IF SAMPLES ARE MAPPED CLOSE TO CENTROIDS IN HILBERT SPACE, THEN RUN STANDARD SVM.

[TABLE 2] SUMMARY OF AL ALGORITHMS [16].

CRITERION	CLASSIFIER	UNCERTAINTY	DIVERSITY	MODELS TO TRAIN
EQB	ALL	AGREEMENT OF A COMMITTEE	×	p MODELS
AMD	ALL	AGREEMENT OF A COMMITTEE	×	p MODELS
MS	SVM	DISTANCE TO SVM MARGIN	×	ONE SVM
cSV	SVM	DISTANCE TO SVM MARGIN	SPECTRAL DISTANCE TO CURRENT SVs	ONE SVM + DISTANCES TO SVs
MOA	SVM	DISTANCE TO SVM MARGIN	ANGULAR DIFFERENCES	ONE SVM + DISTANCES TO ALREADY SELECTED SAMPLES
MCLU-ECBD	SVM	DISTANCE TO SVM MARGIN	DIFFERENT CLUSTER ASSIGNMENT	ONE SVM + NONLINEAR CLUSTERING OF c SAMPLES
KL-Max	PROB. OUTPUT	DIVERGENCE OF PDF IF ADDING THE CANDIDATE	×	$(c - 1)$ MODELS
BT	PROB. OUTPUT	DIFFERENCE IN POSTERIOR OF MOST CONFIDENT CLASSES	×	ONE MODEL

c : Number of candidates.

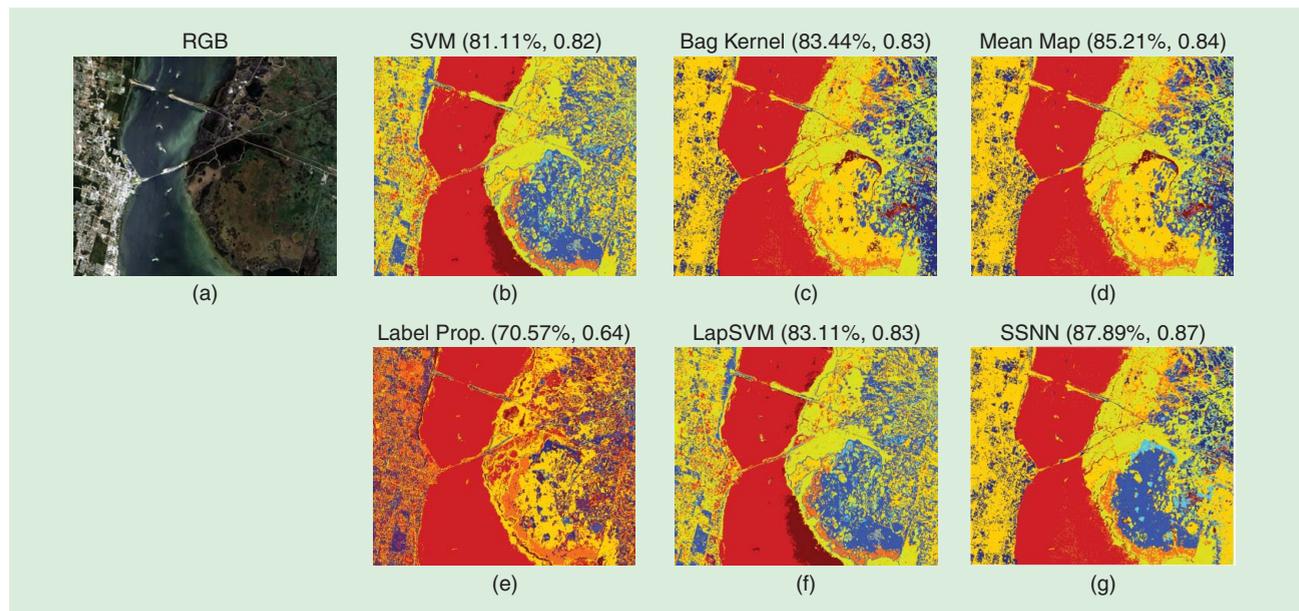
p : Members of a committee of learners.

iterated. Since AL focuses on difficult areas, it boosts the performances with fewer samples than those required by random sampling. Table 2 includes a taxonomical classification of the field of active learning. The section “Adapting the Classifier” will illustrate an example of active learning model regularization for the specific task of adapting classifiers to multiple scenes.

REGULARIZATION THROUGH SPARSITY PROMOTION

Despite their high dimensionality, hyperspectral pixels belonging to the same class typically lie in a low-dimensional subspace. This observation has been recently used in sparse signal representations. Here, the assumption is that pixels can be represented accurately as a linear combination of a few training samples from a structured dictionary.

Note the connection with SVMs that embed the dictionary (the training samples) into a high-dimensional feature space \mathcal{H} . The use of the hinge loss in the SVM functional induces a sparse solution, i.e., few training examples are selected. Recently, sparse kernel methods have been presented, such as the kernel matching pursuit, the ℓ_1 -SVM, the kernel basis pursuit, or the generalized LASSO. In all of these, the dictionary functions are the kernels centered around the selected “support vectors.” Alternatively, in [18], several sparse kernel approaches have been presented with a different philosophy: the target pixel is the test pixel itself, not a similarity evaluation, and the dictionary is composed by the training pixels in the feature space. In this article, a basis projection (BP) approach is used to promote sparsity with $\Omega = \|w\|_1$, as a relaxation of the more computationally demanding problem



[FIG2] (a) RGB composition and classification maps with (b) SVM, (c) cluster (or bag) kernels, (d) probabilistic mean-map kernel, (e) label propagation, (f) LapSVM, and (g) SSNN for the KSC image ($\ell = 200$, $u = 1,000$). Overall accuracy and kappa statistic are given in brackets.

induced by using the ℓ_0 -norm. To solve the BP problem, greedy algorithms, such as the orthogonal matching pursuit (OMP) and the subspace pursuit (SP) methods, can be used. The dictionary can be obtained offline or from the same image. The classification can be additionally improved by incorporating the contextual information from the neighboring pixels into the classifier (see the next section).

In the example illustrated in Figure 3, we compare the performance of the linear (SP, OMP) and kernel (KSP, KOMP) sparse HSI classifiers introduced in [18]. The sparseness factor was tuned for best performance. We also included a linear SVM and the ν -SVM using an RBF kernel, in which the parameter $\nu \in (0, 1)$ controls the degree of sparsity. The RBF σ parameter was tuned by standard tenfold cross-validation. Figure 3 shows the results for these methods and different number of training samples in the standard AVIRIS Indian Pines hyperspectral image (220 spectral channels and spatial resolution 20 m, shown in Figure 4). This is the standard benchmark hyperspectral image, which is used here to allow comparison with results in [18]. We split the data into a training set (20% of the available labeled pixels) and a test set (80%). We trained the classifiers for different rates {1, 5, 10, 15, 20, 25, 30}% of the training set, and show results for the test set that remained constant. Nonlinear methods show a much better performance over linear approaches. In the linear case, SP clearly outperforms the rest, but when the nonlinearity is included all methods perform very similarly.

SPATIAL-SPECTRAL IMAGE CLASSIFICATION

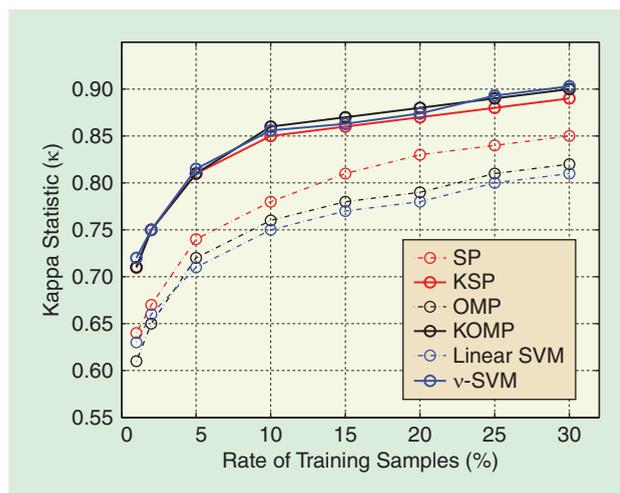
HSI live in a geographical manifold, in the sense that spatially neighboring pixels carry correlated information and that images are usually smooth in the spatial domain [1, Ch. 2]. Accounting for spatial smoothness 1) provides less salt-and-pepper classification maps, 2) reveals the size and shape of the structure the pixel belongs to, and 3) allows the discrimination between structures made of the same materials, but belonging to different land-use types. Spatial regularization has been widely used to improve classification [21]. The joint exploitation of both spectral and spatial information considers that either the loss, the regularizer, or both depend on the spatial neighborhood of a pixel [22]. In the next subsections we review the field of spatial-spectral classification, whose major developments are summarized in Table 3.

SPATIAL FEATURE EXTRACTION

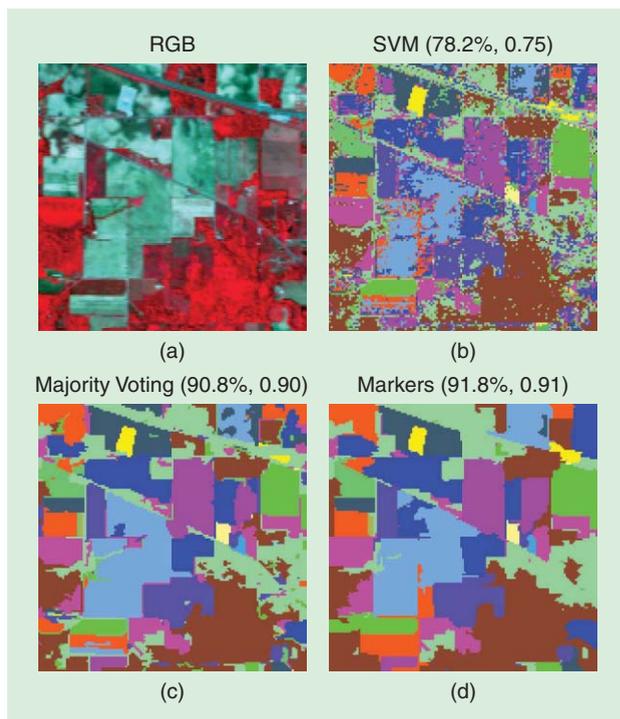
A simple yet effective way to regularize for spatial smoothness is to enrich the input space with features accounting for the neighborhood of the pixels. This is usually done by using moving windows or adaptive filters applied to the spectral bands. These filtered images are then used to learn the classifier. Standard filters based on occurrence or co-occurrence, morphological operators, Gabor filters, or wavelets decompositions generally provide significant improvements over purely spectral classifiers. Among them, morphological filters are the most promising. In [23], filtering was performed at many scales and an extended morphological profile (EMP) was used for classification. Proceeding in a multiscale fashion enables the adaptive

definition of the neighborhood of a pixel according to the structure it belongs to. Filtering in HSI is more challenging than in multispectral images, and one typically resorts to compute the EMP based on only a few principal components (PCs) using morphological reconstruction operators. All of the features are then fed to a classifier, either alone [23] or combined with the original spectral information [24]. Furthermore, feature selection [25] or extraction [23] can be used to find the relevant features. Recently, connected tree-based morphological operators have been investigated for the analysis of HSI [26]. These so-called attribute filters extract thematic attributes of the connected components of an image which are thresholded according to their geometry (area, length, shape factors), or texture (range, entropy). The multiscale version, extended morphological attribute profile (EMAP), has been also introduced.

In Figure 5, the approaches based on mathematical morphology (EMP and EMAP extracted from the first four PCs) are used for classification of ROSIS-03 data from an urban area in Pavia, Italy. We selected this image to illustrate the capabilities of several spatial-spectral classifiers since urban areas monitoring at very high resolution typically requires the extraction of directional, rotational, and scale features from objects. A significant improvement in terms of classification accuracies with respect to the spectral SVM was achieved by applying the EMAP with four different attributes on the image (area and diagonal of the bounding box of connected components, moment of inertia, and standard deviation; see [26]). On the other hand, some redundancy was observed in the original 144-dimensional filter vector of EMAP. Therefore, decision boundary feature extraction (DBFE) was applied on it. After extraction with DBFE, the accuracies improved significantly (+5%), thus confirming the importance of feature extraction routines. Finally, comparison with composite kernels (see the section “Advanced Spatial-Spectral Classifiers” on the next page and in [19] and [20]) yielded improved classification accuracy, but with a strong change of response in the right part of the image, where much more soil is predicted.



[FIG3] The performance measure with the estimated Cohen's kappa statistic, κ , for different sparsity-promoting classifiers.



[FIG4] RGB composition of the standard AVIRIS Indian Pine data set (200 spectral channels and spatial resolution 20 m). Classification maps are shown for spatio-spectral classification with the segmentation and classification with majority voting and segmentation with markers against SVM on the original image only. Overall accuracies and the kappa statistic for each method are reported in parentheses [21].

SPATIAL-SPECTRAL SEGMENTATION

Another approach for the inclusion of spatial information is through image segmentation, typically using watershed, mean shift, and hierarchical segmentation [21]. After segmentation, a supervised scheme assigns the pixels in the segments to the classes. Two approaches are the mostly used: in the first, the regions are treated as input vectors in a supervised classifier. In the second, regions are considered as basins to postprocess the class memberships attributed by a pixel-based classifier within each segment.

The reverse view on the problem is proposed in [27], where a supervised classifier is used to produce confidence values for each pixel. Then, pixels with maximal confidence are used as seeds for a region growing algorithm. In [28] segmentation and classification are linked by user-provided labels: working with a hierarchical segmentation of the data, the labels provided are used to isolate coherent clusters, both spatially and thematically, thus ending with the good segmentation and the labels of the segments. The number of queries is minimized with AL.

In Figure 4, two approaches based on segmentation and classification with majority voting and markers are applied to a 220-bands AVIRIS data set over Indian Pines (Indiana, United States). Significant improvement in terms of overall classification accuracies and kappa statistic were achieved over the pixel-based SVM classifier. Using markers provided the best accuracies: a +1% improvement over the simple majority voting and more than +13% over the traditional pixel-based SVM classifier.

Furthermore, as seen in Figure 4, it is clear that the classification and segmentation approaches provide a significantly more uniform classification map when compared to the purely spectral SVM classification map.

ADVANCED SPATIAL-SPECTRAL CLASSIFIERS

The main problem with spatial-spectral feature extraction approaches is the possibly high dimensionality of the feature vectors used to feed the classifier. This was alleviated in [19] where dedicated kernels for the spectral and spatial information were combined. The framework has been recently extended to deal with convex combinations of kernels through multiple-kernel learning [25] and generalized composite kernels [20]. In both cases, however, the methodology still relies on performing an ad hoc spatial feature extraction before kernel computation. Other alternatives in the literature considered the definition of graph kernels that capture multiscale higher-order relations in a neighborhood without computing them explicitly [29], and the modification of the SVM to seek for the spatial filter that maximizes the margin [30].

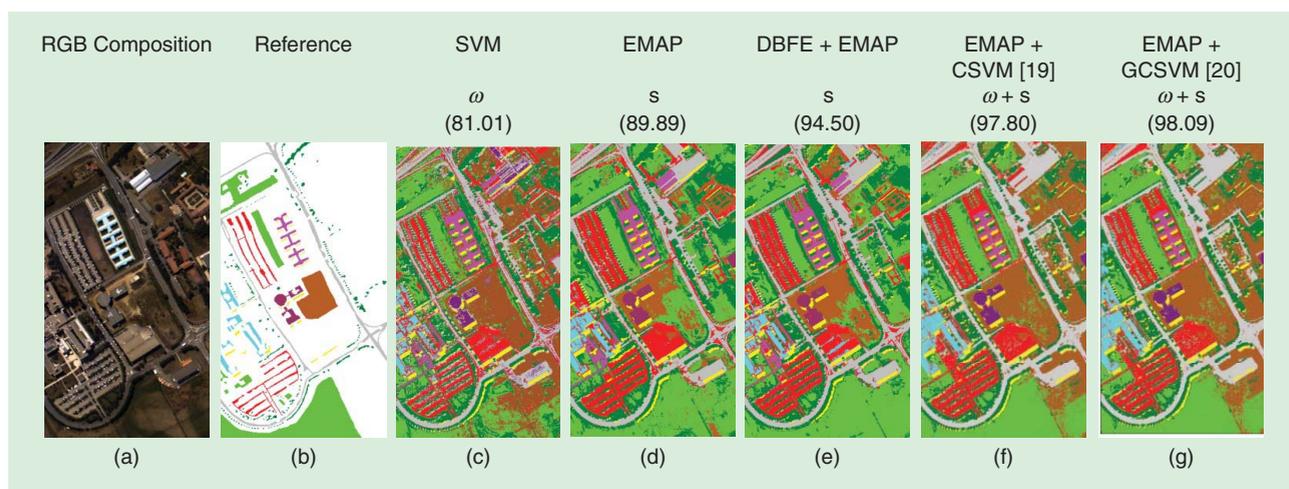
A final alternative is to include contextual information with Markov random fields (MRFs), which naturally include a spatial term on class smoothness in the energy function. However, in the high-dimensional context of HSI, the standard application of the neighbor system definition makes the problem computationally intractable, and therefore recent works have focused on joining MRF spatial priors and discriminative models in HSI classification [31], [32]. An excellent review of MRF spatial-spectral methods can be found in [33].

ADAPTATION AND INVARIANCES

One of the greatest challenges of modern HSI classification is the adaptation of classifiers between acquisitions that differ either by the zone they represent and/or the acquisition conditions such as illumination, angle, and season, among other effects. Adaptation is a central issue in HSI classification. For example, the increase in revisit time of recent satellites has improved multitemporal analysis of scenes. Nevertheless, algorithms must be able to adapt to changing situations. Generally, the direct application of classifiers trained on one image to new images leads to poor results: even if the objects represented in the images are roughly the same, differences in acquisition induce significant local changes in the PDF. These changes must be modeled and introduced in the classifiers. The concept of adaptation can be implemented at the levels of image preprocessing, robust and invariant feature extraction, or in the design of the classification algorithm.

PREPROCESSING

The preprocessing phase can address adaptation through the use of radiometric correction techniques applied to the images. Absolute corrections aim at transforming the radiance measured at the sensor into surface reflectance. Relative calibration techniques adapt the radiometric properties between portions of an image or between images. Generally, absolute correction techniques require additional ground reference data that in many



[FIG5] (a) RGB composition along with (b) the available reference data for the ROSIS-03 Pavia University area data set (103 spectral channels and spatial resolution 1.3 m). Classification maps are shown for (c) SVM on the original image only against spatiospectral classification with the (d) EMAP and (e) EMAP after feature extraction by DBFE (81 data channels), (f) composite kernels with cross-kernels and SVM [19], and (g) generalized composite kernels with multinomial logistic regression [20]. Overall accuracies [%] are reported in parentheses (ω : using spectral bands, s : using spatial filters from PCA).

cases are not available or are difficult to collect. Relative calibration methods are often considered as a pragmatic alternative for adaptation. Among these methods, we recall histogram matching, relative radiometric normalization of time series [34], and multivariate histogram matching [35].

Images can be casted as point clouds in a geometrical space endorsed with an appropriate distance measure. Such a view is quite convenient because it allows us to move from image adaptation to manifold adaptation. Recent methods have explicitly considered the distortions occurring between image manifolds. In [36], multitemporal sequences for each pixel were aligned based on a measure of similarity between sequences barycenters, thus consisting into a global measure of alignment. In [37], spectra of the pixels are spatially detrended using Gaussian processes to avoid shifts related to geometrical differences or to localized class variability. A recent principled approach tries to match graphs representing the data manifolds [38]. There, the graphs of

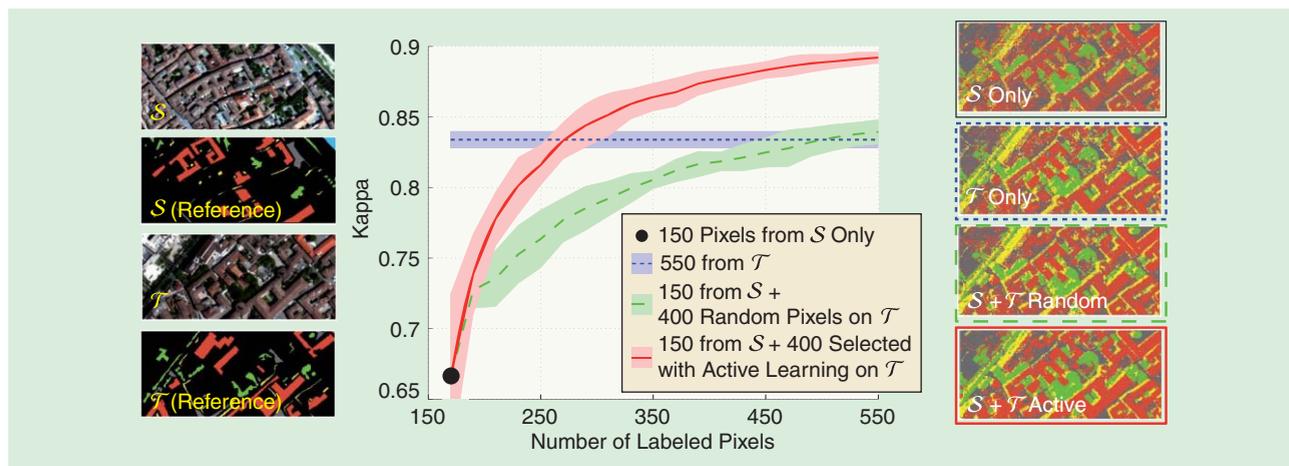
the two domains are matched using a procedure aiming at maximizing their similarity, while at the same time preserving the original structure of the graphs.

ADAPTING THE CLASSIFIER

Learning a transformation between domains may be insufficient to handle all the perturbing factors, so alternative approaches are concerned with the adaptation of the classifier itself. From a machine-learning perspective, the problem of classifier adaptation is studied in the framework of transfer learning, and in particular, of domain adaptation. Domain adaptation reduces to learning from data in a source domain \mathcal{S} (e.g., a portion of an image) to extrapolate to a different target domain \mathcal{T} (another portion of the image or to another image). The problem has been given attention in HSI classification lately [39]. In this setting, source and target domains are assumed to share the same set of information classes (exceptions to this constraint in [40] and

[TABLE 3] SUMMARY OF SPATIAL-SPECTRAL ALGORITHMS.

TYPE OF APPROACH	MODEL	IDEA
SPATIAL FILTERS EXTRACTION	CO-OCCURRENCE	EXTRACT TEXTURE BASED ON STATISTICS OF PAIRS OF PIXELS IN A NEIGHBORHOOD.
	EMP	MULTISCALE MATHEMATICAL MORPHOLOGY (BASED ON SIZE).
	EMAP	MULTISCALE MATHEMATICAL MORPHOLOGY (VARIETY OF ATTRIBUTE TYPES).
SPATIAL-SPECTRAL SEGMENTATION	SEGMENTATION AND CLASSIFICATION BASED ON MAJORITY VOTING	ALL PIXELS ARE ASSIGNED TO THE MOST FREQUENT CLASS INSIDE A SEGMENTED REGION.
	SEGMENTATION AND CLASSIFICATION BASED ON MARKERS	MOST RELIABLY CLASSIFIED PIXELS ARE SELECTED AS "REGION MARKERS" FOR SEGMENTATION.
	SEMI-SUPERVISED HIERARCHICAL CLUSTERING TREE	RETURNS BOTH CLASSIFICATION AND CONFIDENCE MAPS. ACTIVE LEARNING USED TO SELECT INFORMATIVE SAMPLES.
ADVANCED SPATIAL-SPECTRAL CLASSIFICATION	COMPOSITE AND MULTIPLE KERNELS	BALANCES BETWEEN SPATIAL AND SPECTRAL INFORMATION WITH DEDICATED KERNELS.
	GRAPH KERNELS	TAKES INTO ACCOUNT HIGHER ORDER RELATIONS IN EACH PIXEL NEIGHBORHOOD.
	MRF	MARKOV RANDOM FIELD MODELING (PROBABILISTIC).



[FIG6] The use of AL to adapt a maximum likelihood classifier [40]. (a) Given an image and a set of reference pixels in a first source area (S), we want to classify another spatially disconnected target area (T), by adding labels chosen actively in the reference of T . (b) Learning curves for the active (solid red line) and random sampling (dashed green line), evolving between the extremes of a model without adaptation (black dot at 150 samples), and another actively adapted that uses 550 samples randomly selected from T (blue dashed line). (c) Classification maps in the target domain T . The color of the bounding box in each map refers to the legend in the central plot.

[41]) and to follow similar (but not the same) class distributions. Domain adaptation problems in remote sensing have been mainly addressed with semisupervised techniques, which exploit the labeled samples from S and the unlabeled samples from T to derive a classification rule suitable for the target domain. The most recent developments in this sense consider semisupervised and domain adaptation SVMs [39], Gaussian processes [37], [41], and the mean-map kernel methods [15].

Recently, AL has been also used for adaptation assuming that some samples (as few as possible) from the target domain can be labeled by the user and added to the existing training set (defined on S) to adapt the classifier to the target domain S [40], [42]. This makes the adaptation process more robust than in the case of semisupervised learning at the cost of requiring additional labeled samples. Figure 6 illustrates this principle for a 102-bands image of the city center of Pavia acquired by the ROSIS-03 airborne sensor. The whole image (only a portion is shown) has a size of $1,400 \times 512$ pixels and spatial resolution is 1.3 m. Five classes of interest (buildings, roads, water, vegetation, and shadows) are considered, and a total of 206,009 labeled pixels are available. We explore the potential of migrating a classifier built on a source area S with as few labeled pixels as possible from the rest of the image T . First, the model trained with 150 labeled pixels randomly drawn from S was directly applied to T and yielded a κ_S of 0.67. Noticeably, a classifier built on 550 samples randomly selected on T reached $\kappa_S = 0.84$. This suggests that different regions of the same scene follow very different statistics for classification. To improve the first classifier, we enlarged the first training set with labeled samples drawn from T , either taken through random sampling (RS, green dashed line) or with AL (AL, red solid line). Using AL allows for concentrating efforts in areas where the first model is suboptimal, so performance is improved with respect to RS. After 400 queries (thus, a model using 550 training samples in total), random sampling yields similar performance to a model

using 550 randomly drawn pixels ($\kappa_{RS} = 0.84$), while AL improves the results with $\kappa_{AL} = 0.89$. To reach the performance of the model using 550 random pixels, AL requires only 120 active queries (thus, a total of 270 samples in the model).

ENCODING INVARIANCES IN THE CLASSIFIER

Image classifiers must be robust to changes in the data representation within each land cover class. The property of such mathematical functions is called *invariance*. A classifier should be invariant to object rotations, changes in illumination, the presence of shadows, and the spatial scale of the objects to be detected. Extracting robust features (invariants) for classification and domain adaptation has been traditionally pursued by looking at the spatial or the spectral signal characteristics. On the one hand, scale invariants aim to make classifiers invariant to perturbations of object scales. In HSI classification, a single spatial scale is typically suboptimal because different classes exhibit diverse sizes, shapes, and internal variations. Multiscale classification schemes may alleviate these problems. Also, translation invariants have been recently explored by means of scale invariant feature transform descriptors and wavelet-based representations. On the other hand, spectral invariants are considered the fundamental descriptors of object structure and are commonly employed to characterize canopy structure. Spectral invariance to daylight illumination allows, e.g., the improvement of multitemporal image classification.

Incorporating invariances in SVM can be achieved by designing particular kernel functions that encode local invariance under transformations, or to generate artificial examples for training to which the model must be invariant. In the following example, we consider the latter possibility, with the virtual SVM (VSVM) method, which has been successfully exploited to encode scale, rotation, translation, and shadow invariance in HSI classification [43].

In Figure 7, we illustrate the use of the VSVM encoding (spectral) shadow-invariance for image classification. We use the same data acquired by the ROSIS-03 optical sensor of the city center of Pavia (Italy) used before. We perform patch-based classification using only 50 training patches of size $w = 5$. The classes to be detected are, again, buildings, roads, water, vegetation, and shadows. Virtual support vectors (VSVs) were generated according to the observed exponential behavior of the ratio shadow/sunlit as a function of the wavelength [43]. Numerical results, as well as the zoom on a detail of the classification map, show that VSVM leads to more accurate results than the standard SVM: encoding shadow invariance reduces misclassifications on the bridge area and an overall more homogeneous classification over flat areas [see, e.g., the crossroads in Figure 7(b) and (c)].

CONCLUSIONS AND DISCUSSION

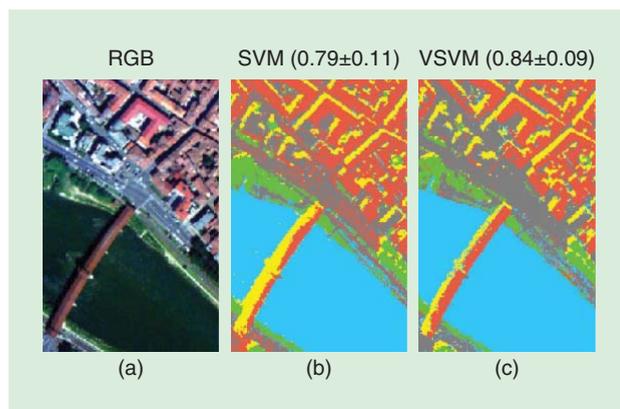
This article reviewed and analyzed the recent developments in hyperspectral image classification. Even though HSI follow similar spatial, spectral, and spatial-spectral image statistics to those conveyed by conventional photographic images, the hyperspectral signals impose additional challenges related to their high dimensionality and heterogeneity. Therefore, even though standard techniques in image processing and computer vision may be transported directly, HSI impose important constraints to develop efficient and effective classifiers.

The use of methods derived from SLT has been a driving factor in recent years. SLT constitutes a proper framework to tackle the problems posed by hyperspectral remote sensing images, which typically involve scenarios with high-dimensional data and few training samples. SLT permits the embedding of numerical regularization in nonlinear classifiers, and also the design of alternative forms of conditioning and incorporation of prior knowledge. Additionally, classification is often improved by including spatially based and manifold-based regularizers. SLT also allows the ability to design sparse methods that are able to work in relevant feature subspaces, where compact and computationally efficient methods can be run. Finally, the SLT framework has allowed the inclusion of prior knowledge in a very fruitful way, e.g., classifiers can now incorporate spatial and spectral invariances that disentangle ambiguities present in land-cover classification.

The field is fast moving, and it is attracting researchers from the computer vision and machine-learning communities. New approaches are introduced regularly, which tackle new scenarios issued from high-resolution imaging (e.g., multitemporal, multi-angular), while learning the relevant features via robust classifiers. It should be also noted that, with upcoming satellites, efficient algorithms for dimensionality reduction before classification and fast/parallel computing solutions will be necessary to accelerate the interpretation and efficient exploitation of HSI.

ACKNOWLEDGMENTS

This work was partially supported by the Spanish Ministry of Economy and Competitiveness (MINECO) under project LIFE-VISION TIN2012-38102-C03-01, and by the Swiss National Research Foundation under grant PZ00P2-136827.



[FIG7] An experiment of patch-based classification with the virtual SVM encoding shadow invariance. (a) True-color composite and classification maps using (b) the standard SVM and (c) the VSVM obtained using 50 training pixels. Results are shown in parentheses in the form of (mean \pm standard deviation of κ in 20 realizations).

AUTHORS

Gustavo Camps-Valls (gustavo.camps@uv.es) received a Ph.D. degree in physics (2002) from the Universitat de València, Spain, where he is currently an associate professor in the Electrical Engineering Department and head of the Image and Signal Processing Group. His research is in machine learning for signal and image processing with a special focus on remote sensing data analysis. He has coauthored many journal papers and edited several books on kernel methods. The Institute for Scientific Information lists him as a highly cited researcher. He is a member of the IEEE Signal Processing Society's Machine Learning for Signal Processing Technical Committee and is an associate editor of *IEEE Transactions on Signal Processing*, *IEEE Signal Processing Letters*, and *IEEE Geoscience and Remote Sensing Letters*. He is a Senior Member of the IEEE.

Devis Tuia (devis.tuia@epfl.ch) received the Ph.D. degree in environmental sciences at the University of Lausanne (Switzerland) in 2009. He was a postdoctoral researcher at both the University of València, Spain, and the University of Colorado at Boulder. He is now a senior research associate at the Laboratory of Geographic Information Systems, at the Ecole Polytechnique Fédérale de Lausanne. His research interests include the development of algorithms for information extraction and classification of very high-resolution remote sensing images using machine-learning algorithms.

Lorenzo Bruzzone (lorenzo.bruzzone@ing.unitn.it) received the Ph.D. degree in telecommunications engineering from the University of Genoa, Italy, in 1998. He is currently a full professor at the University of Trento, Italy, where he is the founder and head of the Remote Sensing Laboratory. His current research interests are in the areas of remote sensing, radar and synthetic aperture radar, signal processing, and pattern recognition. He is the author (or coauthor) of many highly cited scientific publications in referred international journals, conference proceedings, and book chapters. He is the cofounder of the IEEE International Workshop on the Analysis of Multitemporal Remote-Sensing Images series.

Since 2003, he has been the chair of the SPIE Conference on Image and Signal Processing for Remote Sensing. Since 2013, he has been the founding editor-in-chief of *IEEE Geoscience and Remote Sensing Magazine*. He is a member of the Administrative Committee of the IEEE Geoscience and Remote Sensing Society (GRSS) and a Distinguished Speaker of the IEEE GRSS. He is Fellow of the IEEE.

Jón Atli Benediktsson (benedikt@hi.is) received the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, Indiana, in 1990. He is currently the prorector for academic affairs and a professor of electrical and computer engineering at the University of Iceland. His research interests are in remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing. He was a cofounder of Oxymap, Inc. He was the 2011–2012 president of the IEEE Geoscience and Remote Sensing Society and was editor-in-chief of *IEEE Transactions on Geoscience and Remote Sensing* (2003–2008). He is a Fellow of the IEEE and SPIE.

REFERENCES

- [1] G. Camps-Valls, D. Tuia, L. Gómez-Chova, S. Jimenez, and J. Malo, *Remote Sensing Image Processing, Synthesis Lectures on Image, Video, and Multimedia Processing*. San Rafael, CA: Morgan and Claypool, 2011.
- [2] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, 2005.
- [3] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
- [4] T. V. Bandos, L. Bruzzone, and G. Camps-Valls, "Classification of hyperspectral images with regularized linear discriminant analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 47, no. 3, pp. 862–873, 2009.
- [5] A. C. Jensen, M. Loog, and A. H. Schistad Solberg, "Using multiscale spectra in regularizing covariance matrices for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 4, pp. 1851–1859, 2010.
- [6] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 3, pp. 809–823, 2012.
- [7] F. Ratle, G. Camps-Valls, and J. Weston, "Semisupervised neural networks for efficient hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 5, pp. 2271–2282, 2010.
- [8] F. A. Mianji and Y. Zhang, "Robust hyperspectral classification using relevance vector machine," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 6, pp. 2100–2112, 2011.
- [9] Y. Bazi and F. Melgani, "Gaussian process approach to remote sensing image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 1, pp. 186–197, 2010.
- [10] G. Camps-Valls and L. Bruzzone, *Kernel Methods for Remote Sensing Data Analysis*. Hoboken, NJ: Wiley, 2009.
- [11] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sensing*, vol. 44, no. 11, pp. 3363–3373, 2006.
- [12] G. Camps-Valls, T. Bandos, and D. Zhou, "Semisupervised graph-based hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 45, no. 10, pp. 3044–3054, 2007.
- [13] D. Tuia and G. Camps-Valls, "Semi-supervised remote sensing image classification with cluster kernels," *IEEE Geosci. Remote Sensing Lett.*, vol. 6, no. 1, pp. 224–228, 2009.
- [14] L. Gómez-Chova, G. Camps-Valls, J. Muñoz-Marí, and J. Calpe, "Semi-supervised image classification with Laplacian support vector machines," *IEEE Geosci. Remote Sensing Lett.*, vol. 5, no. 4, pp. 336–340, 2008.
- [15] L. Gómez-Chova, G. Camps-Valls, L. Lorenzo Bruzzone, and J. Calpe-Maravilla, "Mean map kernels for semi-supervised cloud classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 1, pp. 207–220, 2010.
- [16] D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Muñoz-Marí, "A survey of active learning algorithms for supervised remote sensing image classification," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 3, pp. 606–617, 2011.
- [17] M. M. Crawford, D. Tuia, and L. H. Hyang, "Active learning: Any value for classification of remotely sensed data?" *Proc. IEEE*, vol. 101, no. 3, pp. 593–608, 2013.
- [18] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 1, pp. 217–231, 2013.
- [19] G. Camps-Valls, L. Gómez-Chova, J. Muñoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sensing Lett.*, vol. 3, no. 1, pp. 93–97, 2006.
- [20] J. Li, P. R. Marpu, A. Plaza, J. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 9, pp. 4816–4829, 2013.
- [21] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, 2013.
- [22] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing Environ.*, vol. 113, Suppl. 1, pp. S110–S122, Sept. 2009.
- [23] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 3, pp. 480–490, 2005.
- [24] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 11, pp. 3804–3814, 2008.
- [25] D. Tuia, G. Camps-Valls, G. Matalasci, and M. Kanevski, "Learning relevant image features with multiple kernel classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 10, pp. 3780–3791, 2010.
- [26] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 10, pp. 3747–3762, 2010.
- [27] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers," *IEEE Trans. Syst. Man, Cybern. B*, vol. 40, no. 5, pp. 1267–1279, 2010.
- [28] J. Muñoz-Marí, D. Tuia, and G. Camps-Valls, "Semisupervised classification of remote sensing images with active queries," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 10, pp. 3751–3763, 2012.
- [29] G. Camps-Valls, N. Shervashidze, and K. M. Borgwardt, "Spatio-spectral remote sensing image classification with graph kernels," *IEEE Geosci. Remote Sensing Lett.*, vol. 7, no. 4, pp. 741–745, 2010.
- [30] R. Flamary, D. Tuia, B. Labbe, G. Camps-Valls, and A. Rakotomamonjy, "Large margin filtering," *IEEE Trans. Signal Processing*, vol. 60, no. 2, pp. 648–659, 2012.
- [31] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sensing Lett.*, vol. 7, no. 4, pp. 736–740, 2010.
- [32] G. Moser and S. B. Serpico, "Combining support vector machines and Markov random fields in an integrated framework for contextual image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 5, pp. 2734–2752, 2013.
- [33] G. Moser, S. B. Serpico, and J. A. Benediktsson, "Land-cover mapping by Markov modeling of spatial-contextual information," *Proc. IEEE*, vol. 101, no. 3, pp. 631–651, 2013.
- [34] M. J. Canty, A. A. Nielsen, and M. Schmidt, "Automatic radiometric normalization of multitemporal satellite imagery," *Remote Sensing Environ.*, vol. 91, no. 3–4, pp. 441–451, June 2004.
- [35] S. Inamdar, F. Bovolo, L. Bruzzone, and S. Chaudhuri, "Multidimensional probability density function matching for pre-processing of multitemporal remote sensing images," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 4, pp. 1243–1252, 2008.
- [36] F. Petitjean, J. Inglada, and P. Gancarski, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern Recognit.*, vol. 44, no. 3, pp. 678–693, 2011.
- [37] G. Jun and J. Ghosh, "Spatially adaptive classification of land cover with remote sensing data," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 7, pp. 2662–2673, 2011.
- [38] D. Tuia, J. Muñoz-Marí, L. Gómez-Chova, and J. Malo, "Graph matching for adaptation in remote sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 1, pp. 329–341, 2013.
- [39] L. Bruzzone and M. Marconcini, "Domain adaptation problems: A DASVM classification technique and a circular validation strategy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 770–787, 2010.
- [40] D. Tuia, E. Pasolli, and W. J. Emery, "Using active learning to adapt remote sensing image classifiers," *Remote Sensing Environ.*, vol. 115, no. 9, pp. 2232–2242, 2011.
- [41] G. Jun and J. Ghosh, "Semisupervised learning of hyperspectral data with unknown land-cover classes," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 1, pp. 273–282, 2013.
- [42] C. Persello and L. Bruzzone, "Active learning for domain adaptation in the supervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 11, pp. 4468–4483, 2012.
- [43] E. Izquierdo-Verdiguier, V. Laparra, L. Gómez-Chova, and G. Camps-Valls, "Encoding invariances in remote sensing image classification with SVM," *IEEE Geosci. Remote Sensing Lett.*, vol. 10, no. 5, pp. 981–985, 2013.

[Dalton Lunga, Saurabh Prasad, Melba M. Crawford, and Okan Ersoy]

Manifold-Learning-Based Feature Extraction for Classification of Hyperspectral Data



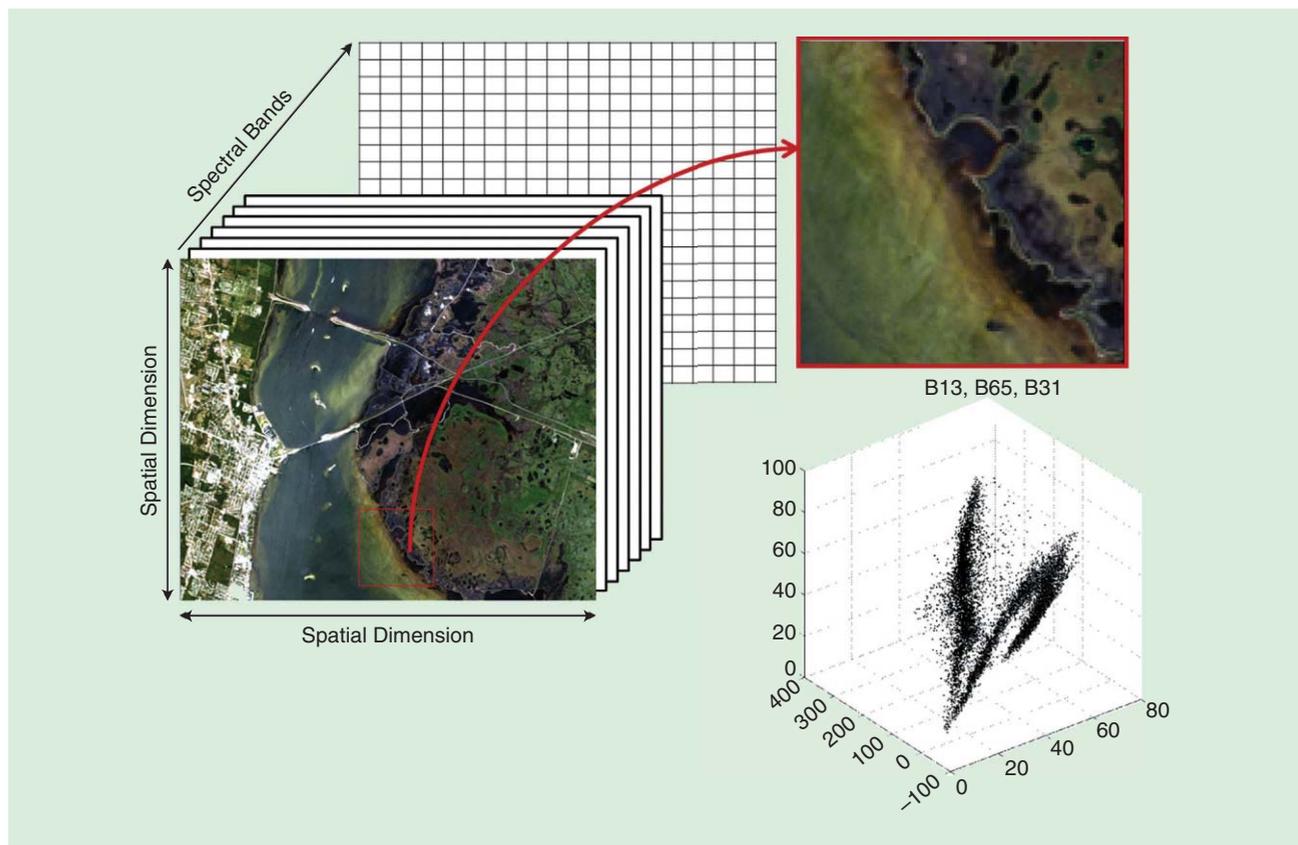
[A review of advances in manifold learning]

Interest in manifold learning for representing the topology of large, high-dimensional nonlinear data sets in lower, but still meaningful, dimensions for visualization and classification has grown rapidly over the past decade, particularly in the analysis of hyperspectral imagery. High spectral resolution and the typically continuous bands of hyperspectral image (HSI) data enable discrimination between

spectrally similar targets of interest, provide capability to estimate within pixel abundances of constituents, and allow for the direct exploitation of absorption features in predictive models. Although hyperspectral data are typically modeled assuming that the data originate from linear stochastic processes, nonlinearities are often exhibited in the data due to the effects of multipath scattering, variations in sun-canopy-sensor geometry, nonhomogeneous composition of pixels, and attenuating properties of media [1]. Because of the dense spectral sampling of HSI data, the associated spectral

Digital Object Identifier 10.1109/MSP.2013.2279894

Date of publication: 5 December 2013



[FIG1] A true color AVIRIS hyperspectral image over the Kennedy Space Center (KSC), Florida. Nonlinearity in the spectral data is exhibited in a plot of bands: 13, 65, and 31.

information in many adjacent bands is highly correlated, resulting in a much lower intrinsic dimensional space spanned by the data (Figure 1). Increased availability of HSIs and greater access to advanced computing have motivated the development of specialized methods for the exploitation of the nonlinear characteristics of these data. In this context, feature selection and feature extraction approaches for dimensionality reduction have received significant attention. While both feature selection and extraction result in some loss of information relative to the original data, both have been demonstrated to be quite successful in the classification arena. Feature selection retains meaningful features for classification, but the algorithms are computationally intensive and often not robust in complex scenes. Alternatively, feature extraction approaches, which project the data to lower-dimensional intrinsic spaces, are typically more robust to variation in spectral signatures across scenes, and most are computationally superior to optimal feature selection, although the interpretation relative to the original spectral signatures is lost. Both feature selection and extraction are flexible relative to the choice of the back-end classifier.

Theoretical contributions and applications of manifold learning have progressed in tandem, with new results providing capability for data analysis and applications highlighting limitations in existing methods. For HSIs, the enormous size

of the data sets and spatial clustering of classes on the image grid provide both challenges and opportunities to extend traditional manifold-learning methods. The machine-learning community has demonstrated the potential of manifold-based approaches for nonlinear dimensionality reduction and modeling of nonlinear structure [2]–[10]. The potential value of manifold learning for HSI analysis has been demonstrated for applications including feature extraction [1], [11], segmentation [12], classification [13]–[15], anomaly detection [16], [17], and spectral unmixing [18]–[21] with some approaches exploiting interband correlation [14], [15] and local spatial homogeneity [21]. Challenges encountered in analyzing data sets have inspired recent advances in manifold-learning methods, particularly related to feature extraction and visualization. This article provides both an overview of traditional approaches and new directions for modeling HSI data on nonlinear manifolds.

A general framework for representing spectral signatures based on graph weights is presented, and traditional unsupervised global and local graph-based methods for dimensionality reduction are summarized. Extensions to exploit labeled data in single image and multitemporal sequences of hyperspectral data are described. Variants of manifold-learning-based projection are particularly suitable as a preprocessing step to traditional Bayesian classification. In this context, locality preserving

discriminant analysis methods are discussed. While traditional eigendecomposition-based methods are computationally advantageous, iterative methods can often provide improved separation of classes in the embedded space for both visualization and classification. Iterative methods are introduced in the context of the affinity matrix, which is utilized to describe multidimensional artificial field embedding (MAFE) and spherical stochastic neighbor embedding (SSNE) [22]. Examples of selected methods applied to a testbed of hyperspectral data are included for illustration of the methods using a one-nearest neighbor classifier.

NONLINEAR MANIFOLD LEARNING IN A GRAPH EMBEDDING FRAMEWORK

Given a data set with training samples $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^n$ in \mathbb{R}^m (m -dimensional feature space) and n is the total number of training samples, nonlinear dimensionality reduction algorithms adapt a graph embedding framework in which $G = \{\mathbf{X}, \mathbf{W}\}$ is the undirected weighted graph and \mathbf{W} is the $n \times n$ data-dependent similarity or *affinity* matrix. The algorithms utilize the notion of affinity weights $W_{ij} \in [0, 1]$ to measure the “distance” between two sample observations. The affinity functions do not use class label information but rather characterize the neighborhood relationships between all pairs of points based on feature differences. A popular approach to measure the affinity between samples \mathbf{x}_i and \mathbf{x}_j makes use of the heat-kernel

$$W_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\gamma_i \gamma_j}\right), \quad (1)$$

where $\gamma_i = \|\mathbf{x}_i - \mathbf{x}_i^{(k_{nn})}\|$ denotes the local scaling of data samples in the neighborhood of \mathbf{x}_i , and $\mathbf{x}_i^{(k_{nn})}$ is the k_{nn} -nearest neighbor of \mathbf{x}_i . Although the heat kernel has been shown to result in effective locality preserving properties, further improvements toward sparse affinity matrices can be achieved by adapting the scaling parameter γ_i to the local data statistics, which often provide a stronger adaptivity to the underlying structure of the embedded image manifolds. The affinity matrix can be modified to include spatial context, which can have a significant impact for manifold learning with HSIs, as discussed in a later section.

When considering multiple data sources (e.g., coregistered gridded imagery data), disparity in the resulting feature spaces can be addressed via separate affinity matrices dedicated to each source. In the realm of kernel methods, a simple approach that has been exploited for geospatial image analysis utilizes composite kernels (e.g., a weighted linear mixture of kernels, each dedicated to a data source: $\mathbf{W}_{ij} = \sum_{k=1}^K \alpha_k \mathbf{W}_k(\mathbf{x}_i^k, \mathbf{x}_j^k)$, s.t. $\alpha_k \geq 0$ and $\sum_{k=1}^K \alpha_k = 1$) to create a unified Gram matrix that characterizes relations across different input sources [23]. In the context of

manifold-learning algorithms, such an approach is particularly relevant for algorithms that operate directly on the affinity matrix, \mathbf{W} . Various complex functional forms for \mathbf{W}_k can be adapted, although the heat-kernel defined in (1) remains a popular choice.

DIMENSIONALITY REDUCTION VIA GRAPH LAPLACIAN OF SPECTRAL FEATURES

Nonlinear manifold-learning methods are broadly characterized as global or locally based approaches and often represented using a graph embedding framework [13]. Global manifold methods retain the fidelity of the overall topology of the data set but have greater computational overhead for large data sets, while local methods preserve local geometry and are computationally efficient because they only require sparse matrix computations. Although global manifolds seek to preserve geometry across all scales of the data and have less tendency to overfit, which is beneficial for generalization in classification, local methods may yield good results for data sets which have significant local variability or submanifolds.

Many popular existing approaches involve models that compute embeddings to preserve pairwise distances, seeking the global structure of data based on local linear fits.

Manifold learning algorithms such as isometric feature mapping (ISOMAP) [2], kernel principal component analysis (KPCA) [3], and locally linear embedding (LLE) [4], for example, have received much attention because of their firm theoretical foundation associated with the kernel and eigenspectrum framework.

In general, given a data matrix \mathbf{X} , the dimensionality reduction problem seeks to find a set of manifold coordinates $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^n$, $\mathbf{y}_i \in \mathbb{R}^p$, where typically, $m \ll p$, through a feature mapping $\Phi: \mathbf{x} \rightarrow \mathbf{y}$, which may be analytical (explicit) or data driven (implicit), and linear or nonlinear. For the hyperspectral data set used in this article, the “optimal” dimensionality is found to be approximately eight for the classical global manifold-learning embeddings, 15–17 for the local embeddings, and eight to ten for the iterative embeddings. Spectral-based dimensionality reduction algorithms adapt a graph embedding platform, i.e., with $G = \{\mathbf{X}, \mathbf{W}\}$, to compute the affinity matrix from which the graph Laplacian \mathbf{L} is derived. Here, $\mathbf{L} = \mathbf{D} - \mathbf{W}$ with a diagonal degree matrix defined by $D_{ii} = \sum_j W_{ij}$, $\forall i$.

In the one-dimensional case, where the resultant manifold coordinate for n samples is a vector $\mathbf{y} = [y_1, y_2, \dots, y_n]$, the dimensionality reduction criterion for eigenspectrum-based methods can be represented as

$$\mathbf{y}^* = \underset{\mathbf{y} \in \mathbb{R}^n}{\operatorname{argmin}} \sum \|y_i - y_j\|^2 W_{ij} \quad (2)$$

$$= \underset{\mathbf{y} \in \mathbb{R}^n}{\operatorname{argmin}} \mathbf{y} \mathbf{L} \mathbf{y}^T, \quad (3)$$

INCREASED AVAILABILITY
OF HSI AND GREATER ACCESS
TO ADVANCED COMPUTING HAVE
MOTIVATED THE DEVELOPMENT
OF SPECIALIZED METHODS
FOR EXPLOITATION OF THE
NONLINEAR CHARACTERISTICS
OF THESE DATA.

[TABLE 1] AFFINITIES AND CONSTRAINTS FOR VARIOUS GRAPH EMBEDDING ALGORITHMS.

ALGORITHM	AFFINITY	CONSTRAINT	APPROXIMATION
LFDA [15]	$\mathbf{W}_{ij}^{(lb)}, \mathbf{W}_{ij}^{(wb)}$	—	NONE
ISOMAP [5]	$\mathbf{W}_{ij} = \ \mathbf{x}_i - \mathbf{x}_j\ ^2$	$\mathbf{B} = \mathbf{I}$	NONE
PCA/KPCA [3]	$\mathbf{W}_{ij} = 1/n, i \neq j$	$\mathbf{B} = \mathbf{I}$	NONE
LLE [4]	$\mathbf{W}_{ij} = (\mathbf{M} + \mathbf{M}^T - \mathbf{M}^T \mathbf{M})_{ij}$	$\mathbf{B} = \mathbf{I}$	NONE
LE [8]	$\mathbf{W}_{ij} = \exp(-\ \mathbf{x}_i - \mathbf{x}_j\ ^2/t)$	$\mathbf{B} = \mathbf{D}$	NONE
SNE [53]	$\mathbf{W}_{ij} = \frac{\exp(-\ \mathbf{x}_i - \mathbf{x}_j\ ^2/2\sigma_j^2)}{\sum_{r \neq i} \exp(-\ \mathbf{x}_r - \mathbf{x}_i\ ^2/2\sigma_r^2)}$	$I_+(f(u))$	$f(u) = \log \sum_{r \neq i} e^{-u^2}$
SSNE [22]	$\mathbf{W}_{ij} = \frac{W(\mathbf{s}_i, \mathbf{s}_j, \mathbf{x}_i, \mathbf{x}_j)}{\sum_{r \neq i} W(\mathbf{s}_i, \mathbf{s}_r, \mathbf{x}_i, \mathbf{x}_r)}$	$I_+(f(v))$	$f(v) = \log(\sum_{j=1}^n v^{-p})$
MAFE-UR [49]	$\mathbf{W}_{ij} = W(\mathbf{s}_i, \mathbf{s}_j, \mathbf{x}_i, \mathbf{x}_j)$	$I_+(f(u))$	$f(u) = \xi_r u^{-2}$

where $u = \|\mathbf{y}_i - \mathbf{y}_j\|$, $v = \|\mathbf{y}_i - \rho \mathbf{y}_j\|$, $f(\cdot)$ approximates the indicator constraint, $\mathbf{H} = \mathbf{I} - \mathbf{e}\mathbf{e}^T/n$, \mathbf{e} is a n dimensional vector with $\mathbf{e} = [1, 1, \dots, 1]^T$.

$I_+(f(u)) = \begin{cases} 0, & u > \epsilon_{ij} \\ \infty, & u \leq \epsilon_{ij} \end{cases}$, ϵ_{ij} is the equilibrium point where the attraction and repulsion forces balances out. $\mathbf{T}_i = [\mathbf{S}_{\text{sup}}]_i^2$, $\mathbf{M} = -\mathbf{H}\mathbf{T}\mathbf{H}$.

where \mathbf{B} is a constraint matrix that depends on the formulation of the dimensionality reduction method. In many algorithms the constraint removes any arbitrary scaling factors in the embedding space. For example, setting \mathbf{B} to a diagonal matrix often yields the required scale normalization. Table 1 summarizes various constraints that are encountered with traditional and modern graph embedding algorithms. The underlying goal is for sample pairs of larger weight to have manifold coordinates that are closer to each other, under a unique data geometry characterized by the graph Laplacian \mathbf{L} . The solution of the optimization problem can be obtained by solving the eigen-decomposition problem $\mathbf{L}\mathbf{y} = \lambda\mathbf{B}\mathbf{y}$, where the one-dimensional manifold coordinates \mathbf{y} are given by the eigenvector with the smallest nonzero eigenvalue. This one-dimensional case can be easily generalized to the multidimensional case through the following expansion

$$\mathbf{Y}^* = \underset{\mathbf{Y}\mathbf{B}\mathbf{Y}^T = \mathbf{I}}{\operatorname{argmin}} \operatorname{tr}(\mathbf{Y}\mathbf{L}\mathbf{Y}^T), \quad (4)$$

where \mathbf{I} is the identity matrix. Analogous to the one-dimensional case, the manifold coordinates \mathbf{Y} of target dimension p can be obtained from the eigenvectors corresponding to the p smallest nonzero eigenvalues. Each of the kernel-based manifold-learning algorithms summarized here can be described in terms of this common framework with different Laplacian matrices and constraints. For a detailed discussion, see [13].

ISOMAP and KPCA are the most widely applied global manifold-learning approaches for nonlinear dimensionality reduction. The ISOMAP method assumes that the local feature space formed by the nearest neighbors is linear, and the global nonlinear transformation can be found by connecting these piecewise linear spaces [2]. Defining \mathbf{X}_i , the set of neighborhood nodes of node \mathbf{x}_i , a distance matrix \mathbf{S}' is computed whereby the Euclidean distance to node $\mathbf{x}_j \in \mathbf{X}_i$ is computed, and the distance beyond \mathbf{X}_i is accumulated along the shortest path to obtain a shortest path network \mathbf{S}_{stp} . Dimensionality reduction is then accomplished through multidimensional scaling (MDS). The computational

burden of computing the geodesic distance matrix scales as $O(n^2 \log n)$, motivating development of approximation methods such as Landmark ISOMAP (L-ISOMAP). These methods avoid the computation for the kernel matrix by selecting a subset of the original points, referred to as “landmark samples,” for which the geodesic distance computation is performed and the remainder of the points are inserted into the “backbone,” thereby reducing the computational cost of the method to $O(\ell n \log n)$, where ℓ is the number of landmark samples [24], [25]. KPCA is a nonlinear extension of linear PCA in a feature space induced by a kernel function [3].

Local kernel-based manifold-learning methods include LLE [4], local tangent space alignment (LTSA) [7], and Laplacian eigenmaps (LE) [8]. All three methods are initiated by constructing a nearest neighborhood for each data point, and the local structures are then used to obtain a global manifold. According to the framework, by solving the eigenvalue problem $\mathbf{L}\mathbf{Y} = \lambda\mathbf{B}\mathbf{Y}$, the embedding \mathbf{Y} is provided by the eigenvectors corresponding to the $2 \sim (p + 1)$ smallest eigenvalues (the eigenvector that corresponds to the smallest zero eigenvalue is a unit vector with equal elements and is discarded). In LLE [4], the local properties of each neighborhood are represented by the linear coefficients that best reconstruct each data point from its neighbors. In LTSA [7], the local geometry is described by the local tangent space of each data point, and the global manifold is determined by aligning the overlapping local tangent spaces. LE [8] obtains the weighted neighborhood graph of each data point by calculating the pairwise distances between neighbors, where the distance is normally obtained using a Gaussian kernel function with parameter σ . The embeddings are obtained by minimizing the total distance between each data point and its neighbors in the low-dimensional space. Parameter settings, including the size of the neighborhood, for both global and local manifold-learning methods and intrinsic dimensionality are selected experimentally and usually robust over a range of values.

Supervised implementations of local manifold learning have also been developed for classification. Unsupervised local manifold-learning approaches search the k spectral neighbors of a given point, whereas supervised local manifold-learning approaches identify only the neighbors that are of the same class as the given point, often making these methods more attractive for classification [26], [27]. Supervised local manifold-learning approaches then map all the training data from the same class onto a single point in the embedded space, resulting in computational complexity of $O(mn_1n_2)$, where n_1 and n_2 represent the number of training and testing samples respectively. Assuming there are c classes, the outputs are c orthogonal vectors $Y^c = [y^1, \dots, y^c] \in \mathbb{R}^{p \times c}$. The kernel out-of-sample extension method is attractive for unsupervised kernel-based embedding of large data sets but is required for testing data when training data are embedded via supervised local manifold-learning methods [27].

NEW DIRECTIONS IN MANIFOLD LEARNING

MANIFOLD LEARNING FOR MULTITEMPORAL IMAGE DATA

Classification of remotely sensed data from multiple scenes acquired at different times or from spatially disjoint areas is an important problem where it is often desirable to exploit labeled data from one time or area to classify data from a different time or area. Although global manifolds are assumed to be similar, spectral shifts in classes over space or time typically manifest themselves as localized variations in the manifold. When the goal is to exploit limited labeled data in a transfer learning mode to classify data in other scenes, changes in the manifold between images can result in misclassification of similar classes. Recent investigations that seek to jointly exploit the global and local characteristics of images [28], [29] and manifold alignment [30], [31] provide the foundation for a correspondence-based framework to classify hyperspectral data acquired in multiple time periods [32] or from spatially disjoint areas.

In [33], a joint manifold over time periods T_1 and T_2 was obtained using the distance matrix

$$W_G = \begin{bmatrix} W_{x^{T_1}, x^{T_1}} & W_{x^{T_1}, x^{T_2}} \\ W_{x^{T_2}, x^{T_1}} & W_{x^{T_2}, x^{T_2}} \end{bmatrix}, \quad (5)$$

where $W_{x^{T_1}, x^{T_1}}$ and $W_{x^{T_2}, x^{T_2}}$ are geodesic distances between points within the two images (intra-image distances) which capture the global geometry of the data manifolds, and $W_{x^{T_1}, x^{T_2}}$ and $W_{x^{T_2}, x^{T_1}}$ represent the connection between the two images (interimage distance). The interimage distances and the resulting alignment are based on u bridging pairs $(x_{c_p}^{T_1}, x_{c_q}^{T_2})$, $i \in [1, u]$, determined from the spatial-spectral optimization

$$\arg \min_{p \leq n_1, q \leq n_2} (\|x_p^{T_1} - x_q^{T_2}\| + a \|s_p^{T_1} - s_q^{T_2}\|), \quad (6)$$

where $\{s_1^{T_1}, s_2^{T_1}, \dots, s_{n_1}^{T_1}\} \in \mathbb{R}^2$, and $\{s_1^{T_2}, s_2^{T_2}, \dots, s_{n_2}^{T_2}\} \in \mathbb{R}^2$ are spatial coordinates of the pair of images with n_1 and n_2 pixels, respectively. Distances between points on the two manifolds are

defined in terms of distances to corresponding pairs within the respective manifold: $W_{x_i^{T_1}, x_j^{T_2}}(i, j) = \min(W_{x_i^{T_1}, x_{i_p}^{T_1}} + W_{x_j^{T_2}, x_{j_q}^{T_2}})$, thereby preserving local relations between arbitrary points and their nearest corresponding pair. The optimal joint manifold feature space $Y = \{y_1, \dots, y_{n_1}, y_{n_1+1}, \dots, y_{n_1+n_2}\} \in \mathbb{R}^p$, $p \ll m$ is computed by minimizing the cost function $E = \|\tau(W_X) - \tau(W_Y)\|$ where W_Y is a distance matrix with elements $W_Y(i, j) = \sqrt{(y_i - y_j)}$, and the τ operator converts distance that characterizes geometry to inner products. The resulting problem is solved using classical multidimensional scaling, yielding the respective eigenvectors. In recent work, Tuia et al. [34] also utilized manifold alignment in conjunction with linear, invertible projections to jointly exploit and synthesize data from multiple sensors.

LOCALITY PRESERVING DISCRIMINATIVE DIMENSIONALITY REDUCTION

PCA, linear discriminant analysis (LDA), and their many variants, such as subspace LDA, stepwise LDA [35], [36], etc. are commonly used for feature extraction prior to classification of hyperspectral data. Under the assumption of homoscedastic Gaussian class-conditional distributions, LDA is optimized for classification tasks but does not perform well when the data are heteroscedastic Gaussian, and it can fail for non-Gaussian data. This makes such projections inappropriate for Bayesian classifiers relying on Gaussian mixture models (GMMs), or for classifiers that assume the decision surfaces to be substantially nonlinear (e.g., nonlinear support vector machines in a kernel-induced space). This issue is particularly relevant for hyperspectral imagery, where several factors can lead to deviation from such assumptions, including variable illumination conditions, and significant mixing between the target pixel and background.

Local Fisher's discriminant algorithm (LFDA) [37] was developed as an extension to LDA to accommodate class distributions that are not unimodal homoscedastic Gaussian, combining the discriminative properties of LDA with properties of unsupervised locality-preserving projections (LPPs) [38]. Unlike LDA or PCA, the LPP seeks to find a linear map that preserves the local neighborhood structure of the data in the projected subspace, i.e., neighborhood points in the original input space remain neighbors in the LPP-embedded space and vice versa. LFDA obtains good between-class separation in the projection while preserving the within-class local structure [37]. It can then be expected that LFDA should be a useful feature reduction algorithm for supervised classification tasks, particularly for problems where local structures convey relevant information (e.g., when the data lie on a complex manifold in the input space) and need to be preserved. In recent work for supervised HSI analysis tasks [14], [15], [39], [40], LFDA and its variants have been found to be very effective feature extraction algorithms, particularly when paired with powerful Bayesian classifiers such as GMMs. The heat kernel's normalized version has been adapted for LFDA to compute the local between-class $W_{ij}^{(b)}$ and within-class $W_{ij}^{(w)}$ weights as defined by

$$W_{ij}^{(lb)} = \begin{cases} W_{ij}(1/n - 1/n_l), & \text{if } z_i = z_j = l, \\ 1/n, & \text{if } z_i \neq z_j, \end{cases} \quad (7)$$

$$W_{ij}^{(lw)} = \begin{cases} W_{ij}/n_l, & \text{if } z_i = z_j = l, \\ 0, & \text{if } z_i \neq z_j. \end{cases} \quad (8)$$

Here n_l is the number of available training samples for the l th class, $\sum_{l=1}^c n_l = n$ and class labels are denoted by $z_i \in \{1, 2, \dots, c\}$, where c is the number of classes.

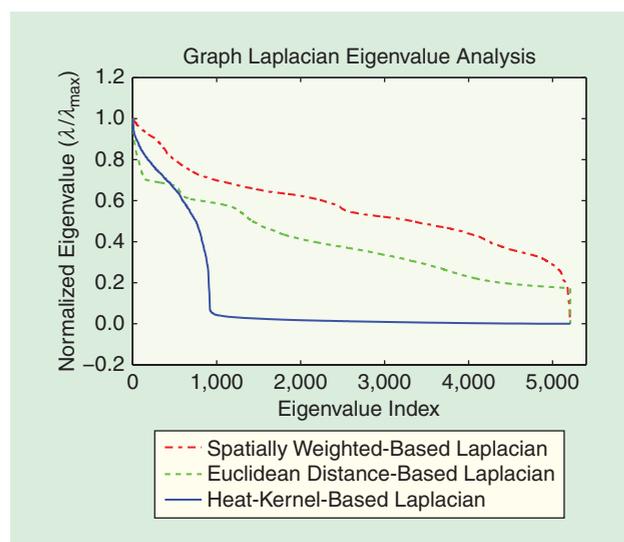
In LFDA, the local between-class $\mathcal{S}^{(lb)}$ and within-class $\mathcal{S}^{(lw)}$ scatter matrices are defined as

$$\mathcal{S}^{(lb)} = \frac{1}{2} \sum_{i,j=1}^n W_{ij}^{(lb)} (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\top, \quad (9)$$

$$\mathcal{S}^{(lw)} = \frac{1}{2} \sum_{i,j=1}^n W_{ij}^{(lw)} (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\top. \quad (10)$$

LFDA seeks to find a projection Φ_{LFDA} that maximizes the “local” Fisher’s ratio as defined using the local scatter matrices defined above. The solution is obtained by solving the generalized eigenvalue problem $\mathcal{S}^{(lb)} \Phi_{\text{LFDA}} = \Lambda \mathcal{S}^{(lw)} \Phi_{\text{LFDA}}$, where Λ is the diagonal eigenvalue matrix.

Based on (9) and (10), LFDA can be thought of as a “localized variant” of LDA since it ensures that local neighborhood structures are preserved by incorporating an appropriate scaling of the scatter matrices. Hence, when the data in the input space lie on nonlinear manifolds, or in general, possess non-Gaussian, even multimodal class-conditional statistics, LFDA is expected to outperform traditional linear projection-based dimensionality reduction approaches. Another benefit of scaling the LFDA-based scatter matrices is that the between-class scatter matrix is no longer rank-limited to $c - 1$. Thus, the “optimal” dimensionality of the projected subspace is no longer restricted to $c - 1$.



[FIG2] Plots of normalized eigenvectors for different graph neighborhoods computed from Euclidean distance, spatially weighted, and heat-kernel-based Laplacian affinity functions for KSC data.

Although LFDA serves as an effective feature reduction strategy for HSIs, it is also prone to statistical ill conditioning when the training sample size is small. In recent work [41], a segmented feature reduction approach was developed wherein the high-dimensional hyperspectral space is partitioned into contiguous subspaces, followed by LFDA-based feature reduction and GMM-based classification. Hyperspectral imagery exploits such an approach naturally, since the correlation structure of the spectral feature space is often strongly block-diagonal (nearby bands are much more correlated than bands that are farther apart). The resulting approach showed substantial robustness to the small-sample-size problem. Other approaches to discriminative feature reduction inspired by manifold learning are also emerging for HSI analysis. For example, in [42], a nearest feature line embedding transformation is proposed for hyperspectral dimensionality reduction, which also seeks to preserve the local manifold structure under the embedding.

MANIFOLD LEARNING FOR SPATIAL-SPECTRAL CLASSIFICATION OF HSI

Traditional nonlinear dimension reduction approaches treat samples as statistically independent, ignoring the local spatial relationships among pixels that occur in patches, as well as the spatially disjoint locations of many spectrally similar classes. Spatial issues have been addressed in many ways by the image processing and remote sensing communities, including Markov random fields, vectors with stacked spectral-spatial features, morphological profiles, and segmentation (see [43] for a comprehensive review).

Recent work related to feature extraction from hyperspectral data has also addressed local spatial relationships via composite and other combined kernels [23], [44]–[46], tensor embedding [47], and iterative methods [22], [48], [49]. Forero and Manian [50] proposed nonlinear diffusion partial differential equations (PDEs) for spatial preprocessing of HSIs, and the results demonstrated a significant improvement in classification performance. Represented in the context of affinities, HSI spectral and spatial neighborhood relations $W_{ij} = W(s_i, s_j, \mathbf{x}_i, \mathbf{x}_j)$ can be computed through a weighted kernel function

$$W(s_i, s_j, \mathbf{x}_i, \mathbf{x}_j) = \exp\left\{-\frac{\|\mathbf{s}_i - \mathbf{s}_j\|^2}{\sigma_s^2}\right\} \cdot \tilde{W}_p(\mathbf{x}_i, \mathbf{x}_j), \quad (11)$$

where s_i denotes the spatial coordinates of pixel i , and \mathbf{x}_i denotes the m -dimensional spectral vector. The expression $\|\mathbf{s}_i - \mathbf{s}_j\|^2$ weights image pixel values as a function of the spatial distance from the center pixel and the variance parameter σ_s and

$$\tilde{W}_p(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^\top \Sigma^{-1}(\mathbf{x}_i - \mathbf{x}_j)\right\} \quad (12)$$

simply weights relations as a function of spectral differences between the center pixel and its neighbor pixel. With additional manipulations as shown in [22], $\tilde{W}_p(\mathbf{x}_i, \mathbf{x}_j)$ can be rewritten as

$$\tilde{W}_p(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{-\frac{n}{2} \text{tr}(\Sigma^{-1} \mathbf{S})\right\}, \quad (13)$$

where S is the sample covariance. Σ^{-1} can be obtained by seeking an orthogonal decomposition of the true covariance matrix Σ . This can be achieved through an efficient, robust sparse matrix transform [48], [51] to decorrelate HSI bands. The resulting affinity function infuses local adaptivity and spatial sensitivity to the neighborhood graph, which leads to preservation of local disjoint neighborhoods that are compact and similar, benefiting hyperspectral data embedding. Figure 2 depicts the eigenspectra corresponding to the spatially weighted Laplacian graph for an HSI. The uniqueness and smoothness of eigenvalues demonstrate the ability of affinity functions to capture both local and global structures in the data. Smooth, rapidly decaying eigenvalues suggest a neighborhood graph with a single, very large connected component, which is the case for using a heat kernel. Alternatively, smooth but slowly decaying eigenvalues suggest a neighborhood graph with various disconnected components, each based on the local spatial details of the image.

GIVEN A SPATIALLY WEIGHTED AFFINITY FUNCTION, HIGH-QUALITY LOWER-DIMENSIONAL HSI VISUALIZATION AND IMPROVED CLASSIFICATION PERFORMANCE MAY BE ACHIEVED BY ADAPTING AN ITERATIVE DYNAMIC EMBEDDING FRAMEWORK.

ITERATIVE GRAPH EMBEDDING FOR DIMENSIONALITY REDUCTION

Nonlinear embedding formulations that ignore spatial relationships often collapse maps toward the center coordinates of the embedding space, thereby increasing the crowding or overlapping of class boundaries. Given a spatially weighted affinity function, high-quality lower-dimensional HSI visualization and improved classification performance may be achieved by adapting an iterative dynamic embedding framework.

In an iterative graph embedding framework, each affinity weight $W_{ij} \in \mathbf{W}$, as computed from (11), is viewed as characterizing spring force properties between a pair of vertices i and j for all $\{(x_i, x_j)\} \in \mathbf{G}$. The affinities can be normalized or unnormalized for each observed pixel pair in \mathbf{X} . The embedding of \mathbf{G} can then be interpreted as an assignment of positions to vertices in a p -dimensional space \mathbb{R}^p . With the notation $\mathbf{Y} = \{y_1, y_2, \dots, y_n\}$ denoting the assigned embedding of \mathbf{G} , where $y_i \in \mathbb{R}^p$ is the position of the map of the i th vertex. An optimal embedding \mathbf{Y} can be obtained through an iterative optimization scheme whose goal is to establish the minimum energy configuration state of \mathbf{G} . The quality of the embedding representation is heavily dependent on the choice of both the objective function and the kernel function used to compute the affinity matrix \mathbf{W} .

Iterative embedding of data is based on an intuitive premise. Assume that $\mathbf{y} = \{y_1^T, y_2^T, \dots, y_n^T\}^T$ is a vector in \mathbb{R}^{Np} that denotes the state of \mathbf{G} . The framework builds on a dynamic model formulation [52], to employ pairwise distance dependent functions and a neighborhood characterized graph to control the grouping of similar vertex maps. The dynamics evolve in continuous time; as such, the velocity as

determined by the additive group effect on each vertex i , and at position y_i is described by

$$\dot{y}_i = \sum_{j \neq i} (y_i - y_j) \{F_r^{ij}(\|y_i - y_j\|) - F_a^{ij}(\|y_i - y_j\|)\}. \quad (14)$$

$F_r^{ij}: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ denotes the repulsion term for dispersing all embedding pixel maps, whereas $F_a^{ij}: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ represents the attraction term for similar pixels.

Functional forms are selected such that an attraction term dominates the pairwise interaction between vertex maps at large distances, while at short distances the repulsion term dominates, and in-between there is a unique distance ϵ_{ij} at which both terms will be in equilibrium—defining a minimum energy configuration state and, hence, an optimal positioning of pairwise vertex maps. To generate

the corresponding force field, the framework assumes the existence of pairwise dependent functions $U_{\text{att}}^{ij}: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ and $U_{\text{rep}}^{ij}: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that

$$\nabla_{y_i} W_{ij} U_{\text{att}}^{ij}(\|y_i - y_j\|) = F_a^{ij}(\|y_i - y_j\|) (y_i - y_j)$$

and

$$\nabla_{y_i} U_{\text{rep}}^{ij}(\|y_i - y_j\|) = F_r^{ij}(\|y_i - y_j\|) (y_i - y_j),$$

where U_{att}^{ij} and U_{rep}^{ij} are viewed as artificial attraction and repulsion potential energy functions that determine the trajectories of vertex maps. The general embedding framework based on the dynamic model has the form

$$\dot{y}_i = \sum_{j \neq i} \nabla_{y_i} \{U_{\text{rep}}^{ij}(\|y_i - y_j\|) - W_{ij} U_{\text{att}}^{ij}(\|y_i - y_j\|)\}. \quad (15)$$

Following the negative gradient, i.e., to achieve an equilibrium state for (15), attraction and repulsion potential functions should be chosen such that the minimum of $U_{\text{att}}^{ij}(\|y_i - y_j\|)$ occurs around $\|y_i - y_j\| = 0$, whereas the minimum of $-U_{\text{rep}}^{ij}(\|y_i - y_j\|)$ occurs around $\|y_i - y_j\| \rightarrow \infty$, and that the minimum of the interactive field $U_{\text{att}}^{ij}(\|y_i - y_j\|) - U_{\text{rep}}^{ij}(\|y_i - y_j\|)$ occurs at $\|y_i - y_j\| = \epsilon_{ij}$, thus defining the equilibrium state of dynamic model. This general framework exhibits strong unifying properties that are applicable for deriving novel iterative MAFE algorithms. Further illustrations in this study demonstrate its use for interpreting some of the existing nonlinear dimensionality reduction models, e.g., reformulation of the SNE [53].

MULTIDIMENSIONAL ARTIFICIAL FIELD EMBEDDING

Following the criteria described in the previous section, an attraction term according to a quadratic form can be chosen, i.e., $U_{\text{att}}^{ij}(\|y_i - y_j\|) = \xi_a \|y_i - y_j\|^2$. The notion of a repulsion force can be interpreted as a barrier constraint that can be

captured by an indicator function, even though its gradient is difficult to compute. There are continuous indicator function approximations that yield useful repulsion terms as summarized in Table 1. An effective repulsion potential function used here has the form $U_{\text{rep}}^i(\|y_i - y_j\|) = (\xi_r) / (\|y_i - y_j\|^2)$. The parameters ξ_a and ξ_r reflect the attraction and repulsion force magnitude. Combining the two terms yields a MAFE unbounded repulsion (MAFE-UR) model [48],

$$U(\mathbf{y}) = \sum_{i=1} \sum_{j \neq i} \left\{ \xi_a \|y_i - y_j\|^2 W_{ij} - \frac{\xi_r}{\|y_i - y_j\|^2} \right\}. \quad (16)$$

Obtaining the optimal embedding maps involves solving a nonconvex optimization problem, $\text{argmin}_{\mathbf{y} \in \mathbb{R}^{np}} U(\mathbf{y})$, whose solution space is known to exhibit many local minima and instabilities for a standard gradient descent algorithm. A much-improved stable and efficient iterative updating scheme can be devised in the form of a local adaptive stochastic descent framework,

$$\mathbf{y}^{(t+1)} = \mathbf{y}^{(t)} - \alpha^{(t)} \nabla U(\mathbf{y}^{(t)}) \quad (17)$$

to yield the optimal maps. Where $\alpha^{(t+1)} = \alpha^{(t)} + \gamma_1 \langle \nabla U(\mathbf{y}^{(t-1)}), \nabla U(\mathbf{y}^{(t)}) \rangle + \gamma_2 \langle \nabla U(\mathbf{y}^{(t-2)}), \nabla U(\mathbf{y}^{(t-1)}) \rangle$ is the common adaptive learning rate. γ_1 and γ_2 are the meta-learning rates. This adaptation scheme exploits gradient-related information from the current as well as the two previous embedding coordinates in the sequence to introduce stability. The computational burden of computing the gradient scales as $O(n^2)$, motivating the need to develop faster approximation methods or finding a closed-form solution to $\text{argmin}_{\mathbf{y} \in \mathbb{R}^{np}} U(\mathbf{y})$.

STOCHASTIC NEIGHBOR EMBEDDING

Hinton and Roweis [53] developed an SNE method for preserving neighbor relations based on probabilities in the lower-dimensional space. The original SNE method assumes that edge weights are antisymmetric Gaussian probabilities W_{ij} (i.e., $W_{ij} \neq W_{ji}$) of pairs of vertices being neighbors in the higher-dimensional space. Considering a symmetric version of W_{ij} , the high-dimensional probability edge weights are defined using the Gaussian functions of the form

$$W_{ij} = \frac{\exp\{-\|x_i - x_j\|^2 / 2\sigma_i\}}{\sum_{r \neq i} \exp\{-\|x_r - x_i\|^2 / 2\sigma_i\}}, \quad (18)$$

where σ_i is computed using a binary search method ensuring that the entropy of the distribution W_i is approximately $\log(k)$, where k is the effective number of neighbors. In the lower-dimensional space, SNE assumes symmetric Gaussian probabilities \tilde{W}_{ij} between embedding coordinates, i.e., embedding graph weights are computed as

$$\tilde{W}_{ij} = \frac{\exp\{-\|y_i - y_j\|^2\}}{\sum_{r \neq i} \exp\{-\|y_r - y_i\|^2\}}. \quad (19)$$

SNE proceeds to compute weights for the maps by minimizing a sum of Kullback–Leibler (KL) objective functions

$$\sum_i KL(W_i || \tilde{W}_i) = \sum_i \sum_{j \neq i} W_{ij} \log\left(\frac{W_{ij}}{\tilde{W}_{ij}}\right). \quad (20)$$

The goal is to minimize the distortion between each of the n high-dimensional neighborhood distributions W_i s and their corresponding lower-dimensional neighborhood distributions \tilde{W}_i s. The difficulty with the original formulation of SNE is encountered in the optimization algorithm, where the antisymmetric assumption poses challenges requiring many experimentally

defined parameters for attaining stability. In a more recent approach, Maaten and Hinton [54] improved on SNE by prescribing a student-t distribution to compute the lower-dimensional probabilities. The improvement led to a tSNE model that preserves meaningful structures in lower-dimensional spaces. A further expansion in (20), while ignoring terms that do not depend

on the unknown probabilities \tilde{W}_{ij} , yields a functional form that makes both SNE and tSNE special cases of (15). In particular, SNE can equivalently be represented by

$$U(\mathbf{y}) = \sum_i \sum_{j \neq i} \|y_i - y_j\|^2 W_{ij} + \log \sum_{r \neq i} \exp\{-\|y_r - y_i\|^2\}.$$

Taking the derivative yields the gradient $\nabla U(\mathbf{y})$ that forms a dynamic equation that can be used to obtain optimal embeddings through an iterative algorithm in (17).

SPHERICAL MANIFOLDS AND STOCHASTIC EMBEDDING

Other than studying manifolds on a flat surface, better visualization and increased classification performance may be achieved by seeking HSI coordinate representations on curved manifolds, which exhibit desirable properties and have been well studied in statistics [55].

To embed data onto a spherical surface, one can consider a unit p -dimensional sphere to be represented as the geometric locations of all unit vectors in \mathbb{R}^{p+1}

$$\mathbb{S}_p = \{y_i \in \mathbb{R}^{p+1} : \|y_i\|_2 = 1\}. \quad (21)$$

For every observed image pixel, the goal is to learn the optimal embedding map y_i and a probability distribution that preserves the neighborhood relations originating from the high-dimensional space. Such a goal can be achieved by applying the SSNE framework, which when given an image X , proceeds to compute the high-dimensional symmetric probability w_{ij} that pixel i would select j as its neighbor as

$$W_{ij} = \frac{W(s_i, s_j, x_i, x_j)}{\sum_{r \neq i} W(s_i, s_r, x_i, x_r)}, \quad (22)$$

OTHER THAN STUDYING MANIFOLDS ON A FLAT SURFACE, BETTER VISUALIZATION AND INCREASED CLASSIFICATION PERFORMANCE MAY BE ACHIEVED BY SEEKING HSI COORDINATE REPRESENTATIONS ON CURVED MANIFOLDS.

where W is a combined spatial-spectral kernel function. The corresponding unit spherical coordinates are obtained from using an exit distribution [56] as a kernel density function that estimates the probability of spherical points being neighbors. The exit distribution has the form

$$f(\mathbf{y}; \mathbf{y}_i, \rho) = \frac{1}{A_p} \frac{1 - \rho^2}{\|\mathbf{y} - \rho \mathbf{y}_i\|^p}, \quad \mathbf{y} \in \mathbb{S}_p, \quad (23)$$

where A_p is the surface area of \mathbb{S}_p , i.e., $A_p = (2\pi^{(p/2)})/(\Gamma(p/2))$, $\Gamma(\cdot)$ is the Gamma function, ρ is the concentration parameter, and \mathbf{y}_i is associated with the mean direction of the distribution. The probability \tilde{W}_{ij} of a spherical map i selecting map j as its neighbor is computed as

$$\tilde{W}_{ij} = \frac{\|\mathbf{y}_j - \rho \mathbf{y}_i\|^{-p}}{\sum_{k \neq i} \{\|\mathbf{y}_k - \rho \mathbf{y}_i\|^{-p}\}}. \quad (24)$$

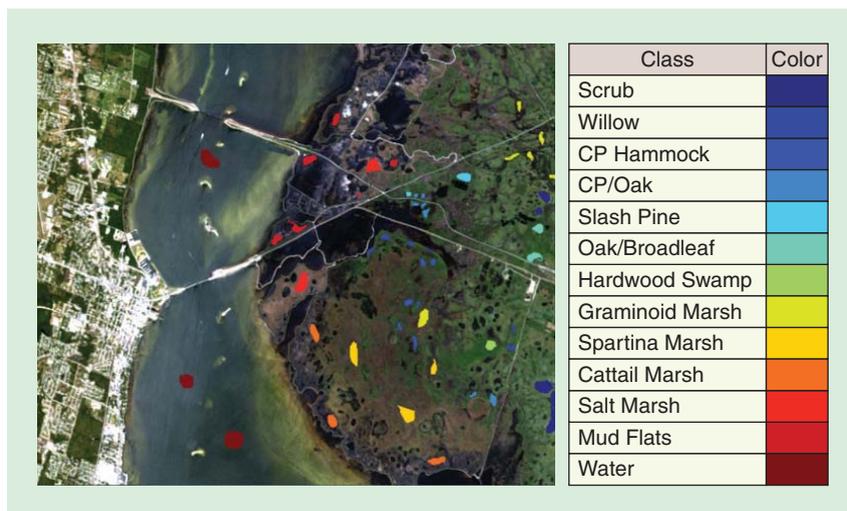
An added benefit of SSNE (and other iterative embedding algorithms) is that they jointly learn the optimal low-dimensional representations and also compute probability distributions over neighborhood relations (or unnormalized relations) with the understanding that spatial proximity should play a role in establishing meaningful manifold structures. SSNE obtains an optimal embedding on a unit (hyper)sphere by iteratively solving an energy minimization problem whose cost function is defined by the sum of KL divergences between the high-dimensional distribution $W_i = (W_{ij})$ and the unknown spherical neighborhood distribution $\tilde{W}_i = (\tilde{W}_{ij})$. The optimization problem is defined as

$$Y^* = \underset{Y = \{y_i^* | y_i \in \mathbb{S}_p\}}{\operatorname{argmin}} \sum_i^n KL(W_i || \tilde{W}_i). \quad (25)$$

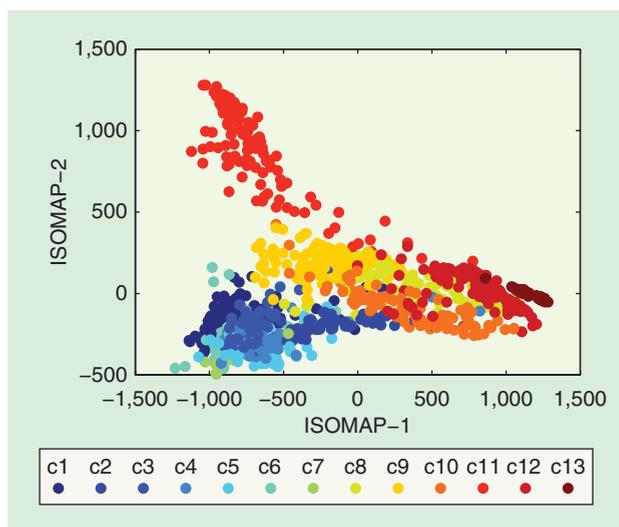
Further manipulations of (25) reveal that SSNE has a functional form of (15) applied to a constant curvature space.

HYPERSPECTRAL IMAGE ANALYSIS EXPERIMENTS

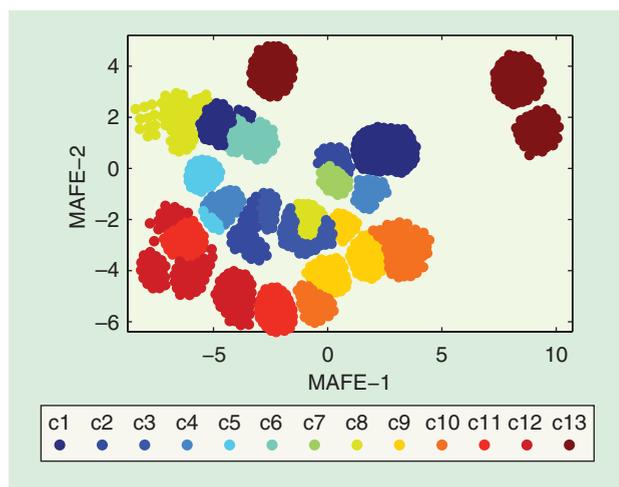
The efficacy of manifold-learning techniques for hyperspectral classification is illustrated using the KSC hyperspectral data—a standard testbed data set that was acquired using the NASA Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor at 18-m spatial resolution. With noisy and water absorption bands removed, 176 features remain for 13 wetland and upland classes of interest. Certain KSC classes that include cabbage palm hammock, and broad leaf/oak hammock upland trees; willow swamp, hardwood swamp, graminoid marsh, and spartina marsh tend to be difficult to separate in lower-dimensional spaces. Their spectral signatures are mixed and often exhibit only subtle differences. Figure 3 illustrates the data set and ground reference information, including the total number of labeled points for each class.



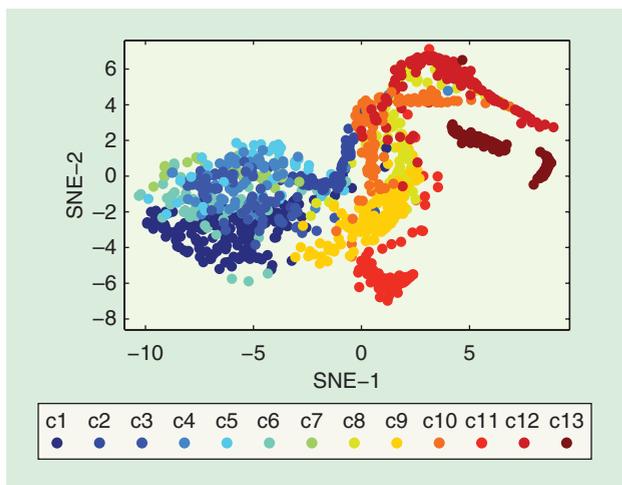
[FIG3] Ground reference information for the KSC hyperspectral data set.



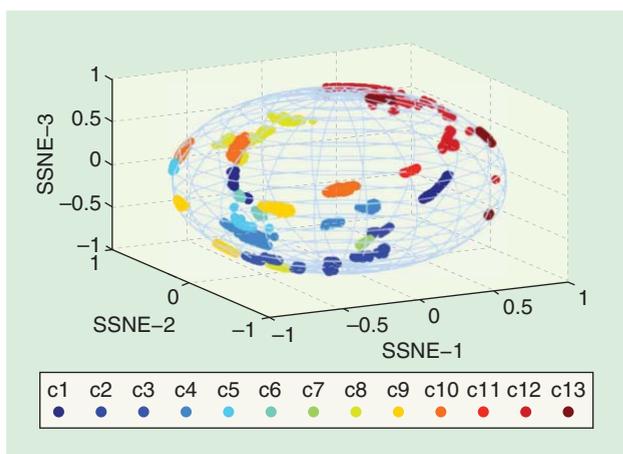
[FIG4] The 2-D scatter plot of the first two dimensions of the ISOMAP embedding of KSC data.



[FIG5] The 2-D scatter plot of the first two dimensions of the MAFE-UR embedding of KSC data.



[FIG6] The 2-D scatter plot of the first two dimensions of the SNE embedding of KSC data.



[FIG7] The 2-D scatter plot of the first two dimensions of the SSNE embedding of KSC data.

VISUALIZATION OF GRAPH EMBEDDING

Figure 4 shows a two-dimensional (2-D) scatter plot after an ISOMAP projection for the KSC hyperspectral data set, respectively. Similar spectral classes such as the lowland marsh grasses (swamp, hardwood swamp, graminoid marsh, and spartina marsh) and the upland woodlands (cabbage palm hammock, and broad leaf/oak hammock upland trees) are

clustered on the manifold. Classes within the groups are difficult to separate because signatures are mixed and often exhibit only subtle differences.

Figures 5–7 depict pixel coordinate representations after the iterative MAFE-UR, SNE, and SSNE 2-D projections, respectively. As illustrated from the embedding visualizations, both SSNE and MAFE map similar pixels onto coordinates with similar values, forming tighter disjoint clusters. The disjoint nature of embeddings is attributed to the spatial information that is captured by the dual spatial-spectral kernel function.

CLASSIFICATION PERFORMANCE VIA MANIFOLD LEARNING

Table 2 depicts the overall classification accuracy, kappa statistic, and the class-specific accuracies using unsupervised (PCA, LLE, ISOMAP) and supervised (LDA, SLLE, LFDA, SSNE) techniques, followed by a 1-NN classifier. Note that all graph-based methods are sensitive to the parameter k_m —the number of neighbors used when constructing the affinity matrix. However, depending upon the data (particularly its local structure) and the embedding algorithm, the classification performance of each algorithm achieves its maximum over a range of k_m values. All the labeled samples (see Figure 3) were used to develop the manifolds via unsupervised methods, and 50% of the labeled samples were used for training and 50% for testing the classifier. Random sampling was repeated 20 times, and the results represent an average accuracy over 20 trials. Manifold-learning techniques outperformed PCA, provided a robust classification performance, and were particularly successful at classifying “hard” classes, such as upland vegetation classes four, five, and six. The intrinsic dimensionality indicated by ISOMAP and the iterative methods was somewhat higher than for PCA and was significantly higher for LFDA, LLE, and SLLE than for PCA. For these data and the 1-NN classifier, higher accuracies were achieved via local methods than global methods, and the value of exploiting correlation structure in the spectral data was demonstrated. The result is consistent with the work of Crawford et al. [13], where a more detailed sensitivity analysis was performed on the parameters for several global and local nonlinear manifold-learning methods. Both the spectral embedding provided

[TABLE 2] THE OVERALL ACCURACY (OA), KAPPA-STATISTIC, AND CLASS-SPECIFIC ACCURACIES FOR THE 13 CLASSES IN THE KSC HYPERSPECTRAL DATA SET.

	OA	KAPPA	1	2	3	4	5	6	7	8	9	10	11	12	13
PCA	88	86.7	93.8	85.8	89	62.2	50.5	44	82	86.5	95.5	92.2	95.7	87.6	99.9
ISOMAP	88.3	86.9	91.6	89.2	84.5	57.3	56.8	42.5	85.2	87.4	95.5	98.4	94.5	89.4	100
LLE	89.5	88.3	92.6	89	84	60.8	54.4	49	81.5	87.5	94.9	98.2	98.6	94.9	100
LDA	94	93.4	95.4	94	84.8	75.4	79.2	78.3	82.8	91.4	97.2	100	98.8	99.3	100
SLLE	93.2	92.4	96.4	94	93	73.1	65.5	62.3	91.7	91.3	98.9	98.6	98.5	95.4	100
LFDA	94.9	93.3	94.7	92.3	89.7	76.9	82.8	82.2	91.8	93.8	98.1	99.8	98.7	99.2	100
SNE	83.5	81.9	91.2	85.3	80.4	51.9	41.2	39.3	82.8	63.6	93.6	93.7	93.8	81.9	100
SSNE	99.4	99	98	100	100	100	95.4	100	100	100	100	100	100	98.5	100
MAFE-UR	99.6	99.7	98.3	100	100	91.2	100	86.2	100	100	100	100	98.97	100	100

by iterative methods and the contribution of localized spatial information were demonstrated by the significantly higher accuracies and high-quality visualizations that were achieved, although the computational overhead of such methods would need to be considered for large remotely sensed data sets.

CONCLUSIONS

Advances in hyperspectral sensing provide new capability for characterizing spectral signatures in a wide range of physical and biological systems, while inspiring new methods for extracting information from these data. HSI data often lie on sparse, nonlinear manifolds whose geometric and topological structures can be exploited via manifold-learning techniques. In this article, we focused on demonstrating the opportunities provided by manifold learning for classification of remotely sensed data. However, limitations and opportunities remain both for research and applications. Although these methods have been demonstrated to mitigate the impact of physical effects that affect electromagnetic energy traversing the atmosphere and reflecting from a target, nonlinearities are not always exhibited in the data, particularly at lower spatial resolutions, so users should always evaluate the inherent nonlinearity in the data. Manifold learning is data driven, and as such, results are strongly dependent on the characteristics of the data, and one method will not consistently provide the best results. Nonlinear manifold-learning methods require parameter tuning, although experimental results are typically stable over a range of values, and have higher computational overhead than linear methods, which is particularly relevant for large-scale remote sensing data sets.

Opportunities for advancing manifold learning also exist for analysis of hyperspectral and multisource remotely sensed data. Manifolds are assumed to be inherently smooth, an assumption that some data sets may violate, and data often contain classes whose spectra are distinctly different, resulting in multiple manifolds or submanifolds that cannot be readily integrated with a single manifold representation. Developing appropriate characterizations that exploit the unique characteristics of these submanifolds for a particular data set is an open research problem for which hierarchical manifold structures appear to have merit. To date, most work in manifold learning has focused on feature extraction from single images, assuming stationarity across the scene. Research is also needed in joint exploitation of global and local embedding methods in dynamic, multitemporal environments and integration with semisupervised and active learning.

AUTHORS

Dalton Lunga (dlunga@csir.co.za) is a senior research scientist at the Meraka Institute, Council for Scientific and Industrial Research, South Africa. He received a B.Eng. degree in electrical and electronics engineering from the University of Johannesburg

DEVELOPING APPROPRIATE CHARACTERIZATIONS THAT EXPLOIT THE UNIQUE CHARACTERISTICS OF THESE SUBMANIFOLDS FOR A PARTICULAR DATA SET IS AN OPEN RESEARCH PROBLEM FOR WHICH HIERARCHICAL MANIFOLD STRUCTURES APPEAR TO HAVE MERIT.

in 2004, an M.S. degree in engineering from Witwatersrand University 2006, and a second M.S. degree in electrical and computer engineering and a Ph.D. degree in electrical and computer engineering, Purdue University, West Lafayette, Indiana, in 2010 and 2012, respectively. He was a Fulbright Scholar at Purdue University. His research interests include statistical machine learning, signal and image processing, opti-

mization, manifold learning, remote sensing, image reconstruction and segmentation, and data fusion.

Saurabh Prasad (saurabh.prasad@ieee.org) is an assistant professor of electrical and computer engineering at the University of Houston. He received a B.Tech. degree from Jamia Millia Islamia, New Delhi, India, in 2003, an M.S. degree from Old Dominion University, Norfolk, Virginia, in 2005, and a Ph.D. degree from Mississippi State University in 2008, all in electrical engineering. He leads a group on geospatial image analysis at the University of Houston. His research interests include statistical signal processing, machine learning, image processing, and data fusion.

Melba M. Crawford (mcrawford@purdue.edu) is the Professor of Excellence in Earth Observation at Purdue University. She received the B.S. and M.S. degrees in civil engineering from the University of Illinois at Urbana-Champaign and the Ph.D. degree in systems engineering from Ohio State University. Her research interests focus on development of advanced methods for image analysis, including manifold learning, active learning, classification, and unmixing, and the application of these methods to remotely sensed data. She is a Fellow of the IEEE and the current president of the IEEE Geoscience and Remote Sensing Society.

Okan Ersoy (ersoy@purdue.edu) is a professor in the School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana. He received the B.S.E.E. degree from Robert College, Istanbul, Turkey, in 1967, and the M.S. Certificate of Engineering, M.S., and Ph.D. degrees from the University of California, Los Angeles, in 1968, 1971, and 1972, respectively. His current research interests include remote sensing, machine learning and pattern recognition, digital signal/image processing and recognition, transform and time-frequency methods, imaging, diffractive optics, and distant learning. He is a Fellow of the IEEE and the Optical Society of America.

REFERENCES

- [1] C. M. Bachmann, T. L. Ainsworth, and R. A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 3, pp. 441–454, 2005.
- [2] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [3] B. Scholkopf, A. J. Smola, and K. R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comput.*, vol. 10, no. 5, pp. 583–588, 1998.
- [4] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by local linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

- [5] V. de Silva and J. B. Tenenbaum, "Global versus local methods in nonlinear dimensionality reduction," in *Proc. Advanced Neural Information Processing Systems*, Hyatt Regency, Vancouver, B.C., Canada, 2002, vol. 15, pp. 713–720.
- [6] L. K. Saul and S. T. Roweis, "Think globally, fit locally: Unsupervised learning of low dimensional manifolds," *J. Mach. Learn. Res.*, vol. 4, pp. 119–155, 2003.
- [7] Z. Zhang and H. Zha, "Principal manifolds and nonlinear dimensionality reduction via local tangent space alignment," *SIAM J. Sci. Comput.*, vol. 26, no. 1, pp. 313–338, 2004.
- [8] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, no. 6, pp. 1373–1396, June 2003.
- [9] D. K. Agrafiotis, "Stochastic proximity embedding," *J. Comput. Chem.*, vol. 24, no. 10, pp. 1215–1221, 2003.
- [10] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin, "Graph embedding and extensions: A general framework for dimensionality reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 40–51, 2007.
- [11] J. He, L. Zhang, Q. Wang, and Z. Li, "Using diffusion geometric coordinates for hyperspectral imagery representation," *IEEE Geosci. Remote Sensing Lett.*, vol. 6, no. 4, pp. 767–771, 2009.
- [12] A. Mohan, G. Sapiro, and E. Bosch, "Spatially coherent nonlinear dimensionality reduction and segmentation of hyperspectral images," *IEEE Geosci. Remote Sensing Lett.*, vol. 4, no. 2, pp. 206–210, 2007.
- [13] M. M. Crawford, L. Ma, and W. Kim, "Exploring nonlinear manifold learning for classification of hyperspectral data," in *Optical Remote Sensing: Advances in Signal Processing and Exploitation Techniques*, S. Prasad, J. Chanussot, and L. B. Eds. London: Springer-Verlag, 2011, pp. 207–234.
- [14] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality preserving dimensionality reduction and classification for hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 4, pp. 1185–1198, 2012.
- [15] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classification," *IEEE Geosci. Remote Sensing Lett.*, vol. 8, no. 5, pp. 895–898, 2011.
- [16] L. Ma, M. M. Crawford, and J. W. Tian, "Anomaly detection for hyperspectral images based on robust locally linear embedding," *J. Infrared Millimeter Terahertz Waves*, vol. 31, no. 6, pp. 753–762, 2010.
- [17] L. Zhang, D. Tao, and X. Huang, "Sparse transfer manifold embedding for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sensing*, 2013. DOI: 10.1109/TGRS.2013.2246837.
- [18] R. Heylen, D. Burazerovic, and P. Scheunders, "Nonlinear spectral unmixing by geodesic simplex volume maximization," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 3, pp. 534–542, 2011.
- [19] A. Halimi, Y. Altmann, N. Dobeigone, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, 2011.
- [20] R. Heylen and P. Scheunders, "Calculation of geodesic distances in nonlinear mixing models: Application to the generalized bilinear model," *IEEE Geosci. Remote Sensing Lett.*, vol. 9, no. 4, pp. 644–648, 2012.
- [21] J. Chi and M. M. Crawford, "Selection of landmark points on nonlinear manifolds for spectral unmixing using local homogeneity," *IEEE Geosci. Remote Sensing Lett.*, vol. 10, no. 4, pp. 711–715, 2013.
- [22] D. Lunga and O. Ersoy, "Spherical stochastic neighbor embedding of hyperspectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 2, pp. 857–871, 2013.
- [23] G. Camps-Valls, L. Gomez-Chova, J. Muñoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sensing Lett.*, vol. 3, no. 1, pp. 93–97, 2006.
- [24] Y. Chen, M. Crawford, and J. Ghosh, "Improved nonlinear manifold learning for land cover classification via intelligent landmark selection," in *IEEE Int. Conf. Geoscience and Remote Sensing Symp.*, 2006, pp. 545–548.
- [25] V. Silva and J. Tenenbaum, "Sparse multidimensional scaling using landmark points," Tech. Rep., Stanford Univ., June 2004.
- [26] X. Geng, D. Zhan, and Z. Xhou, "Supervised nonlinear dimensionality reduction for visualization and classification," *IEEE Trans. Syst., Man, Cybern. B*, vol. 35, no. 6, pp. 1098–1107, 2005.
- [27] M. Li, M. M. Crawford, and J. Tian, "Local manifold learning-based k-nearest-neighbor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 11, pp. 4099–4109, 2010.
- [28] J. Verbeek, "Nonlinear image manifolds by global alignment of local linear models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1236–1250, 2006.
- [29] R. Wang, S. Shan, X. Chen, J. Chen, and W. Gao, "Maximal linear embedding for dimensionality reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 9, pp. 1776–1792, 2011.
- [30] S. Lafon, Y. Keller, and R. R. Coifman, "Data fusion and multicue data matching by diffusion maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1784–1797, 2006.
- [31] C. Wang and S. Mahadevan, "Heterogeneous domain adaptation using manifold alignment," in *Proc. Int. Joint Conf. Artificial Intelligence (IJCAI)*, Barcelona, Spain, July 16–20, 2011, pp. 1541–1546.
- [32] H. L. Yang and M. M. Crawford, "Manifold alignment for multitemporal hyperspectral image classification," in *Proc. 2011 IEEE Int. Geoscience and Remote Sensing Symp.*, Vancouver, B.C., July 25–29, 2011, pp. 4332–4335.
- [33] H. L. Yang and M. M. Crawford, "Learning a joint manifold with global-local preservation for multitemporal hyperspectral image classification," in *Proc. 2013 IEEE Int. Geoscience and Remote Sensing Symp.*, Melbourne, Australia, July 21–26, 2013, pp. 1047–1050.
- [34] D. Tuia, M. Trolliet, and M. Volpi, "Multisensor alignment of image manifolds," in *Proc. 2013 IEEE Int. Geoscience and Remote Sensing Symp.*, Melbourne, Australia, July 21–26, 2013, pp. 1246–1249.
- [35] S. Prasad and L. M. Bruce, "Limitations of principal component analysis for hyperspectral target recognition," *IEEE Geosci. Remote Sensing Lett.*, vol. 5, no. 4, pp. 625–629, Oct. 2008.
- [36] T. V. Bandos, L. Bruzzone, and G. Camps-Valls, "Classification of hyperspectral images with regularized linear discriminant analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 47, no. 3, pp. 862–873, 2009.
- [37] M. Sugiyama, "Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis," *J. Mach. Learn. Res.*, vol. 8, pp. 1027–1061, May 2007.
- [38] X. He and P. Niyogi, "Locality preserving projections," in *Advances in Neural Information Processing Systems*, S. Thrun, L. Saul, and B. Schölkopf, Eds. Cambridge, MA: MIT Press, 2004.
- [39] M. Cui, S. Prasad, W. Li, and L. Bruce, "Locality preserving genetic algorithms for spatial-spectral hyperspectral image classification," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing*, vol. 6, no. 3, pp. 1688–1697, May 2013.
- [40] G. Zhang and X. Jia, "Feature selection using kernel based local fisher discriminant analysis for hyperspectral image classification," in *Proc. 2011 IEEE Int. Geosci. Remote Sensing Symp.*, pp. 1728–1731.
- [41] S. Prasad, M. Cui, W. Li, and J. Fowler, "Segmented mixture of Gaussian classification for hyperspectral image analysis," *IEEE Geosci. Remote Sensing Lett.*, to be published.
- [42] Y.-L. Chang, J.-N. Liu, C.-C. Han, and Y.-N. Chen, "Hyperspectral image classification using nearest feature line embedding approach," *IEEE Trans. Geosci. Remote Sensing*, vol. PP, no. 99, pp. 1–10, 2013. DOI: 10.1109/TGRS.2013.2238635
- [43] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.
- [44] W. Kim, M. M. Crawford, and J. Ghosh, "Spatially adapted manifold learning for classification of hyperspectral imagery with insufficient labeled data," in *Proc. 2008 IEEE Int. Geoscience and Remote Sensing Symp.*, pp. 213–217.
- [45] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial-spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recognit.*, vol. 45, no. 1, pp. 381–392, 2012.
- [46] H. H. Yang and M. M. Crawford, "Exploiting spectral-spatial proximity for classification of hyperspectral data on manifolds," in *Proc. 2012 IEEE Geoscience and Remote Sensing Symp.*, Munich, Germany, pp. 4174–4177.
- [47] L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 1, pp. 242–254, Jan. 2013.
- [48] D. Lunga and O. Ersoy, "Dynamic hyperspectral embedding with a spatial sensitive graph," in *Proc. 2013 IEEE Geoscience and Remote Sensing Symp.*, Melbourne, pp. 2176–2179.
- [49] D. Lunga and O. Ersoy, "Multidimensional artificial field embedding with spatial sensitivity," *IEEE Trans. Geosci. Remote Sensing*, vol. PP, no. 99, pp. 1–1, 2013. DOI: 10.1109/TGRS.2013.2251889
- [50] S. Velasco-Forero and V. Manian, "Improving hyperspectral image classification using spatial preprocessing," *IEEE Trans. Geosci. Remote Sensing*, vol. 6, no. 2, pp. 297–301, 2009.
- [51] J. Theiler, G. Cao, L. R. Bachege, and C. A. Bouman, "Sparse matrix transform for hyperspectral image processing," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 3, pp. 424–437, 2011.
- [52] V. Gazi and K. M. Passino, "Stability analysis of swarms," in *Proc. American Control Conf.*, 2002, pp. 692–697.
- [53] G. Hinton and S. Roweis, "Stochastic neighbor embedding," in *Proc. ICML*, vol. 15, pp. 833–840.
- [54] L. van der Maaten and G. Hinton, "Visualizing data using T-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, 2008.
- [55] K. V. Mardia and P. E. Jupp, *Directional Statistics*. New York: Wiley, 2000.
- [56] S. Kato, "A distribution for a pair of unit vectors generated by Brownian motion," *Bernoulli*, vol. 15, no. 3, pp. 898–921, 2009.

Wing-Kin Ma, José M. Bioucas-Dias, Tsung-Han Chan, Nicolas Gillis, Paul Gader,
Antonio J. Plaza, ArulMurugan Ambikapathi, and Chong-Yung Chi

A Signal Processing Perspective on Hyperspectral Unmixing



Insights from remote sensing

Blind hyperspectral unmixing (HU), also known as unsupervised HU, is one of the most prominent research topics in signal processing (SP) for hyperspectral remote sensing [1], [2]. Blind HU aims at identifying materials present in a captured scene, as well as their compositions, by using high spectral resolution of hyperspectral images. It is a blind source separation (BSS) problem from a SP viewpoint. Research on this topic started in the 1990s in geoscience and remote sensing [3]–[7], enabled by technological advances in hyperspectral sensing at the

time. In recent years, blind HU has attracted much interest from other fields such as SP, machine learning, and optimization, and the subsequent cross-disciplinary research activities have made blind HU a vibrant topic. The resulting impact is not just on remote sensing—blind HU has provided a unique problem scenario that inspired researchers from different fields to devise novel blind SP methods. In fact, one may say that blind HU has established a new branch of BSS approaches not seen in classical BSS studies. In particular, the convex geometry concepts—discovered by early remote sensing researchers through empirical observations [3]–[7] and refined by later research—are elegant and very different from statistical independence-based BSS approaches established in

Digital Object Identifier 10.1109/MSP.2013.2279731

Date of publication: 5 December 2013

the SP field. Moreover, the latest research on blind HU is rapidly adopting advanced techniques, such as those in sparse SP and optimization. The present development of blind HU seems to be converging to a point where the lines between remote sensing-originated ideas and advanced SP and optimization concepts are no longer clear, and insights from both sides would be used to establish better methods.

This article uses an SP researcher's perspective to review blind HU. We will consider several key developments, which include pure pixel search, convex geometry, dictionary-based sparse regression and nonnegative matrix factorization. We will not cover Bayesian techniques [8], although readers should note that they also represent key developments in blind HU. Our emphasis will be on insights, where we will showcase how each approach fundamentally works, and highlight significant results from a viewpoint of SP theory and methods. Some forefront advances will also be discussed. Note that this article does not aim at survey; please see a recent overview paper [2] that provides a comprehensive coverage of numerous blind HU methods and many other aspects.

Our notations are standard in SP. In addition, given a matrix X , x_i , and x^i denote its i th column and i th row, respectively; " \geq " represents elementwise inequality; $\mathbf{1}$ is an all-one vector of appropriate length; A^\dagger is the pseudoinverse of A ; $P_A^\perp = I - A(A^T A)^\dagger A^T$ is the orthogonal complement projector of A ; $\sigma_{\min}(A)$ and $\sigma_{\max}(A)$ denote the minimum and maximum singular values of A , respectively; $\|\cdot\|_p$ denotes the ℓ_p norm; and $\|\cdot\|_F$ denotes the Frobenius norm.

SIGNAL MODEL

Modeling hyperspectral signals is a difficult problem. It depends on numerous factors; some crucial ones include: the types of materials encountered in the acquired scene, the ways the materials are physically mixed and constitute the scene topologically, the way light interacts with the materials, gets reflected and measured by the hyperspectral instrument, and the measurement environment. Over decades, the geoscience and remote sensing community has devoted tremendous efforts to various modeling aspects,

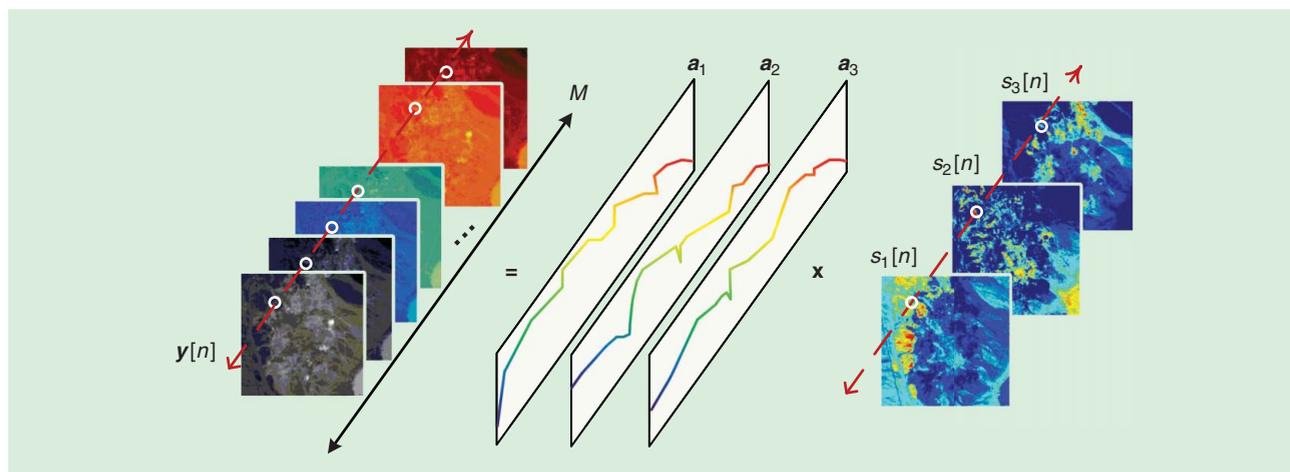
from which we have now significantly improved our understanding of the true problem nature. Nevertheless, modeling can be an overwhelmingly complex process if one wants to treat every aspect very precisely. In particular, while radiative transfer theory (RTT) is well known to be able to provide accurate characterizations of photons' interactions with the materials (see [2] and the references therein), the resulting models are generally too difficult to use for signal analysis and processing. There is a compromise to make between model accuracy and tractability.

We focus on a relatively simplistic but very representative model, specifically, the linear mixing model (LMM). The LMM lies at the center of interest of many important developments in blind HU. Despite the fact that the LMM is not always true, especially under certain scenarios that exhibit strong nonlinearity, it is generally recognized as an acceptable model for many real-world scenarios. The LMM is described as follows. We assume a macroscopic mixing scale in which the incident light interacts with only one material before reflecting off. Let $y_m[n]$ denote the hyperspectral camera's measurement at spectral band m and at pixel n . Letting $\mathbf{y}[n] = [y_1[n], y_2[n], \dots, y_M[n]]^T \in \mathbb{R}^M$ where M is the number of spectral bands, the LMM is given by

$$\mathbf{y}[n] = \sum_{i=1}^N \mathbf{a}_i s_i[n] + \mathbf{v}[n] = \mathbf{A} \mathbf{s}[n] + \mathbf{v}[n] \quad (1)$$

for $n = 1, \dots, L$, where each $\mathbf{a}_i \in \mathbb{R}^M$, $i = 1, \dots, N$, is called an *endmember signature vector*, which contains the spectral components of a specific material (indexed by i) in the scene; N is the number of endmembers, or materials, in the scene; $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N] \in \mathbb{R}^{M \times N}$ is called the *endmember matrix*; $s_i[n]$ describes the contribution of material i at pixel n ; $\mathbf{s}[n] = [s_1[n], \dots, s_N[n]] \in \mathbb{R}^N$ is called the *abundance vector* at pixel n ; L is the number of pixels; and $\mathbf{v}[n] \in \mathbb{R}^M$ is noise. Figure 1 illustrates the mixing process under the LMM.

There are several important aspects concerning the LMM formulation. First, since hyperspectral cameras have wide spectral ranges and fine spectral resolution, M is often large—typically more than 200. Such large spectral degrees of freedom allow us to distinguish an endmember signature from another, as well as



[FIG1] The linear mixing model.

mixtures of endmember signatures, provided that the materials are sufficiently different from one another. Hence, it is reasonable to assume that $\{a_1, \dots, a_N\}$ is linearly independent, and we will assume that this condition holds throughout the article. Second, the mixing process in (1) is a consequence of limited spatial resolution of hyperspectral cameras. Specifically, one pixel may not be spatially fine enough to contain one material only. For example, each pixel is about $4 \text{ m} \times 4 \text{ m}$ to $20 \text{ m} \times 20 \text{ m}$ for airborne visible/infrared imaging spectrometer, depending on the altitude of the flight. Third, while the noise vector $\nu[n]$ is commonly used to represent background and instrument noise, one may also use it to incorporate errors arising from modeling inaccuracies. From such a perspective, (1) can serve as a reasonable approximate model when nonlinear effects are not too strong. Fourth, by nature, the abundance vectors $s[n]$ should satisfy

$$s_i[n] \geq 0, i = 1, \dots, N, \text{ and } \sum_{i=1}^N s_i[n] = 1, \quad (2)$$

for every $n = 1, 2, \dots, L$. The second constraint above, commonly referred to as the *abundance sum constraint* or the *sum-to-one constraint*, means that abundances give the fractional proportions, or percentages, of the different materials in a pixel. For convenience, we will write

$$s[n] \in \mathcal{S} = \{s \in \mathbb{R}^N \mid s \geq 0, s^T \mathbf{1} = 1\}, \quad (3)$$

where \mathcal{S} denotes the feasible set of abundance vectors. Note that \mathcal{S} is a unit simplex.

The LMM introduced above is considered standard. That said, there are some hidden complications. Here we briefly mention them; interested readers can find further clarifications in [2]. First, in the model (1), $y[n]$ s are actually processed measurements. Raw measurements from the hyperspectral camera usually undergo a series of processing steps, such as radiometric calibration, geometric correction, and atmospheric compensation [9], before arriving at the simple LMM. Second, for simplicity we have associated an endmember with a material, presumably pure. However, an endmember could also be a composition of several materials; i.e., a material made of several materials. The definition of an endmember can be subjective, and is dependent on applications. Third, we have assumed that the sum-to-one constraint in (3) holds. In practice, the sum-to-one constraint may be violated under the so-called endmember variability (EV) effects. Besides modeling issues, it is worth noting that recently there has been growing interest in considering specific but more treatable nonlinear mixture models for HU; the same applies to EV. In these scenarios, insights learned from LMM-based HU remain vital and provide building blocks for non-LMM HU problems there. We refer readers to [10] and [11] in this issue of *IEEE Signal Processing Magazine* for a coverage of nonlinear HU and EV, respectively.

PROBLEM STATEMENT

We are concerned with the HU problem, under the model setting in (1)–(3). Specifically, HU aims at recovering $s[n]$ from $y[n]$, thereby retrieving every material's abundance map $\{s_i[n]\}_{n=1}^L$ from the hyperspectral measurements. Assuming full knowledge

of the endmember matrix A , we can carry out unmixing by solving constrained linear least squares (LS) problems:

$$\hat{s}[n] = \arg \min_{s[n] \in \mathcal{S}} \|y[n] - As[n]\|_2^2, \quad (4)$$

for $n = 1, \dots, L$. Fundamentally, the above problem is considered an “easy” problem—it is a convex optimization problem, and a simple way to obtain a solution is to call some general-purpose convex optimization software, such as the widely used CVX [12]. Alternatively, one can design dedicated algorithms for (4) to have more efficient implementations; this is a more popular option in the field [13]–[15]. What makes HU fundamentally challenging is not (4) (or other variants), but the fact that we often do not have full knowledge of A —for if we do, it means that we know exactly all the materials in the scene, which is unlikely in reality.

Blind HU amounts to recovering $\{s[n]\}_{n=1}^L$ from $\{y[n]\}_{n=1}^L$ without knowledge of A . The problem can also be stated as that of identifying A from $\{y[n]\}_{n=1}^L$ without knowledge of $\{s[n]\}_{n=1}^L$. At this point, readers who are familiar with BSS may have realized that the problem formulation of blind HU is the same as that of BSS: The endmember matrix A and abundance vectors $s[n]$ are the mixing matrix and true source vectors in BSS, respectively. While this observation is true, and in fact has been noticed for a while [16], classical BSS methods established in the SP field usually do not fall in any of the mainstream blind HU approaches. The key reason is that under the unit simplex constraint (3), the sources $\{s[n]\}_{n=1}^L$ do not satisfy the statistical independence assumption, which is a very essential assumption in many BSS methods, particularly the well-known independent component analysis (ICA). The violation of source independence makes many existing BSS methods an inappropriate choice for blind HU from the outset.

Before delving into blind HU, we should point out that we will generally assume N , the number of endmembers, to be known. As in BSS and sensor array processing in the SP field, where the same aspect has been extensively studied under the name of *model order selection* (see, e.g., [17]), the problem of identifying the number of endmembers can be seen as a separate problem; see [2, Sec. III] for a description. One may also build on an existing blind HU approach to provide joint blind HU and endmember number identification.

PURE PIXELS PURSUIT

Our review begins with a very simple class of methods that hinges on a special model assumption called *pure pixels*.

DEFINITION OF PURE PIXELS

We say that endmember i (or material i) has a pure pixel if for some index denoted by ℓ_i , we have

$$s[\ell_i] = e_i, \quad (5)$$

where $e_i \in \mathbb{R}^N$ is a unit vector with the nonzero element at the i th entry (that is, $[e_i]_j = 0$ for all $j \neq i$, and $[e_i]_i = 1$). Moreover, we say that the pure pixel assumption holds if every endmember has a pure pixel.

Physically, the existence of pure pixels means that while hyperspectral pixels are generally mixtures of several materials, there are certain pixels that are constituted by one material only. This can be seen from the model (1). Assuming pure pixels and no noise, the observed vector at pixel ℓ_i is

$$\mathbf{y}[\ell_i] = \mathbf{a}_i, \quad (6)$$

for $i = 1, \dots, N$, which are the endmembers. In practice, there are scenarios where the pure pixel assumption holds. For example, imagine a scene that consists of water and soil. If there exist some local pixel regions that contain either water or soil only, then those regions contain pure pixels. Note that since more than one pure pixel may exist for a particular endmember, ℓ_i may not be unique. However, we should also note that the pure pixel assumption does not always hold, e.g., in a scene consisting of highly mixed minerals, or if the spatial resolution of the hyperspectral camera is too low.

Pure pixels provide a unique opportunity for blind HU. In essence, if we know the pure pixel indices ℓ_1, \dots, ℓ_N , then $[\mathbf{y}[\ell_1], \dots, \mathbf{y}[\ell_N]] = [\mathbf{a}_1, \dots, \mathbf{a}_N]$ is the endmember matrix itself—and the problem is solved—in the noiseless case. However, the pure pixel indices are not known a priori, and the problem is to find them.

SUCCESSIVE PROJECTIONS ALGORITHM

We introduce a simple algorithm for finding the pure pixels of all endmembers. The prerequisite required to understand the algorithm is just basic knowledge of linear algebra.

Again, consider the noiseless case and assume that the pure pixel assumption holds. We notice that for any n ,

$$\|\mathbf{y}[n]\|_2 = \left\| \sum_{i=1}^N s_i[n] \mathbf{a}_i \right\|_2 \leq \sum_{i=1}^N \|s_i[n] \mathbf{a}_i\|_2 \quad (7a)$$

$$= \sum_{i=1}^N s_i[n] \|\mathbf{a}_i\|_2 \quad (7b)$$

$$\leq \max_{i=1, \dots, N} \|\mathbf{a}_i\|_2, \quad (7c)$$

where (7a) is due to the LMM and the triangle inequality, and (7b) and (7c) to the unit simplex constraint (3). It can be seen that equality in (7) holds when $s[n] = \mathbf{e}_j$, where $j = \arg \max_{i=1, \dots, N} \|\mathbf{a}_i\|_2$; which holds at $n = \ell_j$, i.e., $\mathbf{y}[n]$ is a pure pixel corresponding to the j th endmember [cf. (6)]. Also, equality in (7) cannot be attained by nonpure pixels, by the equality condition of the triangle inequality and the linear independence of $\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$. Assuming without loss of generality (w.l.o.g.) that $j = 1$, we can identify the first endmember signature by

$$\hat{\mathbf{a}}_1 = \mathbf{y}[\hat{\ell}_1], \quad \hat{\ell}_1 = \arg \max_{n=1, \dots, L} \|\mathbf{y}[n]\|_2. \quad (8)$$

Note that $\hat{\mathbf{a}}_1$ is a perfect estimate of \mathbf{a}_1 under the aforementioned settings.

The next question is to identify pure pixels corresponding to other endmembers. Suppose that we have previously identified

$k-1$ endmember signatures, denoted by $\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_{k-1}$, and that the identification is perfect, i.e., $\hat{\mathbf{a}}_i = \mathbf{a}_i$ for $i = 1, \dots, k-1$. The idea to identify the next endmember is to perform nulling—a standard SP trick that has appeared many times (e.g., [17]), but proves very useful in various fields. Let $\hat{\mathbf{A}}_{1:k-1} = [\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_{k-1}]$, and construct its orthogonal complement projector $\mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp$. Since $\mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \mathbf{a}_i = \mathbf{0}$ holds for any $i < k$, we have that

$$\|\mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \mathbf{y}[n]\|_2 = \left\| \sum_{i=k}^N s_i[n] \mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \mathbf{a}_i \right\|_2 \quad (9a)$$

$$\leq \max_{i=k, \dots, N} \|\mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \mathbf{a}_i\|_2, \quad (9b)$$

where (9b) is obtained in the same way as (7). And like (7), it can be shown that equality in (9) holds only for a pure pixel corresponding to a previously unidentified endmember, which we can assume w.l.o.g. to be that at $n = \ell_k$. The k th endmember signature can therefore be identified via

$$\hat{\mathbf{a}}_k = \mathbf{y}[\hat{\ell}_k], \quad \hat{\ell}_k = \arg \max_{n=1, \dots, L} \|\mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \mathbf{y}[n]\|_2. \quad (10)$$

Hence, by induction, we can identify all the endmembers.

The algorithm presented on the next page is called the *successive projections algorithm* (SPA). Algorithm 1 gives the pseudocode of SPA, which is very simple. From the above algebraic development, we conclude that in the noiseless case and under the pure pixel assumption, SPA perfectly identifies all the endmember signatures $\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$.

We should provide a brief historical note on SPA, since it has been repeatedly rediscovered. To our best knowledge, SPA first appeared in chemometrics in 2001 by Araújo et al. [18]. Later, a very similar algorithm, called the *automatic target generation process* (ATGP), was proposed by Ren and Chang in 2003 in remote sensing [19]. Curiously, the development we just displayed, which shows why SPA works from an algebraic SP viewpoint and pins down its endmember identifiability, was not seen until recently; see [20, Appendix F]. There are other ways to derive SPA, which will be described later. It is worth pointing out that SPA has been used successfully for rather different purposes. In numerical linear algebra, SPA is closely related to the so-called modified Gram–Schmidt algorithm with column pivoting, used for example to solve linear LS problems [21]. In machine learning, SPA has been used for document classification where the pure pixel assumption is referred to as the *separability* assumption and requires that, for each topic, there exists at least one word used only by that topic; see [22] and the references therein.

The above SPA development is based on the noiseless argument. An interesting question is therefore on sensitivity against noise. A provable performance bound characterizing noise sensitivity has been proposed very recently in [23], and is briefly described here. Let us denote $\sigma = \sigma_{\min}(\mathbf{A})$, which is positive since $\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ is linearly independent, and $K = \max_{1 \leq i \leq N} \|\mathbf{a}_i\|_2$. Let us also denote the noise level $\epsilon = \max_{1 \leq n \leq L} \|\mathbf{v}[n]\|_2$. Then, under the pure pixel assumption and assuming that the noise

Algorithm 1 SPA.

input $\{\mathbf{y}[n]\}_{n=1}^L$, N .
 1: $\mathbf{P}^\perp = \mathbf{I}$
 2: for $k = 1, \dots, N$ do
 3: $\hat{\ell}_k = \arg \max_{n=1, \dots, L} \|\mathbf{P}^\perp \mathbf{y}[n]\|_2^2$
 4: $\hat{\mathbf{a}}_k = \mathbf{y}[\hat{\ell}_k]$
 5: $\mathbf{P}^\perp := (\mathbf{I} - (\mathbf{P}^\perp \hat{\mathbf{a}}_k)(\mathbf{P}^\perp \hat{\mathbf{a}}_k)^T / \|\mathbf{P}^\perp \hat{\mathbf{a}}_k\|_2^2) \mathbf{P}^\perp$
 6: end for
 output $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_N]$.

level satisfies $\epsilon \leq O(\sigma^3/(NK^2))$, SPA identifies all the endmember signatures $\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ up to error $O(\epsilon(K^2/\sigma^2))$; more precisely, we have

$$\max_{1 \leq i \leq N} \min_{1 \leq j \leq N} \|\mathbf{a}_i - \hat{\mathbf{a}}_j\|_2 \leq O\left(\epsilon \frac{K^2}{\sigma^2}\right). \quad (11)$$

The above analysis result provides significant practical implications. We see in (11) that the noise robustness of SPA depends on the ratio K/σ . It can be shown that [23]

$$\frac{\max_{1 \leq i \leq N} \|\mathbf{a}_i\|_2}{\min_{1 \leq i \leq N} \|\mathbf{a}_i\|_2} \leq \frac{K}{\sigma} \leq \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}.$$

Thus, the noise robustness of SPA depends on 1) how different the magnitudes of the endmember signatures are and 2) how well the true endmember signatures are spectrally distributed. In particular, the latter implies that challenging scenarios lie in highly similar endmembers.

Let us further point out two notable facts. First, one can generalize SPA by replacing the ℓ_2 norm in (9)–(10) by any continuously differentiable and locally strongly convex function whose minimizer is zero, e.g., any ℓ_p norm with $1 < p < +\infty$. The corresponding algorithm not only works in the noiseless case, it is also shown to possess a similar error bound as in (11) [23]. According to the analysis, the variant using the ℓ_2 norm has the best robustness against noise among all locally strongly convex functions; see also [24] for numerical evidence. Second, it is possible to improve the error bound above to $O(\epsilon(K/\sigma))$ by using the following postprocessing strategy [22]: Let $\{\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_N\}$ be the N endmembers extracted by SPA. Then, for $i = 1, \dots, N$,

- 1) Project the original data $\{\mathbf{y}[n]\}_{n=1}^L$ onto the orthogonal complement of $\{\hat{\mathbf{a}}_k\}_{k=1, k \neq i}^N$.
- 2) Replace $\hat{\mathbf{a}}_i$ with the column of $\{\mathbf{y}[n]\}_{n=1}^L$ whose ℓ_2 norm of the projection is maximum.

This iterative refinement strategy is identical to a previously proposed blind HU algorithm (but without a robustness analysis); it will be further discussed in the section “Simplex Volume Maximization.”

OTHER ALGORITHMS AND DISCUSSION

There are many other pure pixels search algorithms; see [2, Sec. VI.A] for a review. A representative algorithm in this family is vertex component analysis (VCA), proposed in 2003 [25], [26].

VCA is similar to SPA—it also employs successive nulling, but differs in the way it picks pure pixels. Specifically, in VCA, the right-hand sides (RHSs) of (8) and (10) are replaced by

$$\hat{\ell}_k = \max_{n=1, \dots, L} |\mathbf{w}_k^T \mathbf{y}[n]|, \quad (12)$$

for $k = 1, \dots, N$, where \mathbf{w}_k is a randomly generated vector lying on the orthogonal complement subspace of $\hat{\mathbf{A}}_{1:k-1}$. Specifically, it is given by $\mathbf{w}_k = \mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \boldsymbol{\xi} / \|\mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \boldsymbol{\xi}\|_2$, where $\boldsymbol{\xi}$ is an independent and identically distributed (i.i.d.) zero-mean Gaussian vector. Following the same derivations described above for SPA, one can show that VCA also perfectly identifies all the endmember signatures in the noiseless case and under the pure pixel assumption; this result holds with probability one. Also, we must mention the pixel purity index (PPI) by Boardman et al. in 1995 [6], which is one of the earliest blind HU algorithms. PPI does not have successive nulling. It is analogous to running (12) only for $k = 1$, but for many independent random trials. The number of trials needs to be large enough so to increase the chance of successfully hitting all endmembers’ pure pixels. For numerical comparisons of SPA, VCA, and PPI, please see [23] (also [20]).

Some additional comments are in order.

1) To simplify the presentation, we have intentionally skipped a conventional preprocessing procedure, specifically, dimension reduction (DR). In practice, VCA and PPI would apply DR to the observed data $\{\mathbf{y}[n]\}_{n=1}^L$, prior to pure pixels search. While we have seen that DR is not required in SPA (as well as VCA and PPI), applying DR plays a crucial role in suppressing noise, which in turn helps improve pure pixel identification performance. Readers are referred to [2, Section III] for the state-of-the-art DR methods in HU.

2) SPA can be extended in at least two ways. First, it can be modified to accommodate outliers, which are anomalous pixels that exhibit markedly different behaviors from the nominal model and can cause substantial performance degradation. The idea is to consider outliers as endmembers, identify them together with true endmembers, and discard them from the obtained estimates [23]. Second, one can extend the method for joint blind HU and endmember number identification. We note that if we keep running the SPA step in (10) recursively, then, at stage $k = N + 1$, the projection residuals $\|\mathbf{P}_{\hat{\mathbf{A}}_{1:k-1}}^\perp \mathbf{y}[n]\|_2^2$ become zeros. Thus, the projection residuals may serve as an indicator of the number of endmembers. Similar ideas have been considered in [24] and [27].

CONVEX GEOMETRY

We have previously shown how blind HU may be easily handled under the pure pixel assumption. The pure pixel concept actually came from the study of convex geometry (CG) of hyperspectral signals, where remote sensing researchers examined the special geometric structure of hyperspectral signals and looked for automatic methods for endmember determination, i.e., blind HU. In fact, a vast majority of blind HU developments, if

WHO DISCOVERED CONVEX GEOMETRY FOR BLIND UNMIXING?

In geoscience and remote sensing, the work by Craig in the early 1990s [3], [4] is widely recognized to be most seminal in introducing the notion of CG for hyperspectral signal analysis and unmixing. Craig’s original work not only described simplex volume minimization, which turns out to become a key CG concept for blind HU, it also inspired other pioneers, such as Boardman who made notable early contributions to CG-based blind HU [5] and introduced pure pixel search [6], and Winter, who proposed the simplex volume maximization concept [7] that results in the popularized N-FINDR algorithm class. What is remarkable in these early studies is that they discovered such beautiful blind SP concepts through sharp empirical observations and strong intuitions, rather than through rigorous SP or mathematics.

CG is also an idea that has been discovered several times in different areas. The introduction of CG can be traced back to

as early as 1964 by Imbrie [28]. Imbrie’s work belongs to another branch of geoscience studies wherein CG is used for the analysis of compositional data in earth science, such as mineral assemblages, grain-size distribution data, and geochemical and petrological data; see [29] for an overview. In fact, Imbrie’s Q-mode analysis and the subsequent QMODEL by Klován and Miesch [30] are conceptually identical to vertex or pure pixel search, although the methodology is different. Likewise, Full et al. already considered the same simplex volume minimization principle as Craig’s in the 1980s [31]. CG has also been independently discovered in other fields such as chemometrics [32] and SP [33], [34]. In all the discoveries or rediscoveries mentioned above, the driving force that led researchers on different backgrounds to devise the same idea seems to be with the geometric elegance of CG and its powerful implications on solving blind unmixing problems.

not all, are directly or intuitively related to concepts introduced in early CG studies, such as simplex volume minimization by Craig [4], simplex volume maximization by Winter [7], and the previously reviewed pure pixel search by Boardman et al. [6]. We give a historical review in the “Who Discovered Convex Geometry for Blind Unmixing?”

PRELIMINARIES

We introduce several mathematical notations and facts in convex analysis, whose physical relevance to blind HU will become clear soon. The affine hull of a set of vectors $\{a_1, \dots, a_N\} \subset \mathbb{R}^M$ is defined as

$$\text{aff}\{a_1, \dots, a_N\} = \left\{ y = \sum_{i=1}^N \theta_i a_i \mid \theta \in \mathbb{R}^N, \sum_{i=1}^N \theta_i = 1 \right\}. \quad (13)$$

An affine hull can always be represented by

$$\text{aff}\{a_1, \dots, a_N\} = \{y = Cx + d \mid x \in \mathbb{R}^P\} \quad (14)$$

for some $C \in \mathbb{R}^{M \times P}$, $d \in \mathbb{R}^M$, where $\text{rank}(C) = P$ and $P \leq N - 1$ is the affine dimension of the affine hull. The affine dimension is $P = N - 1$ if $\{a_1, \dots, a_N\}$ is affinely independent.

The convex hull of a set of vectors $\{a_1, \dots, a_N\} \subset \mathbb{R}^M$ is defined as

$$\text{conv}\{a_1, \dots, a_N\} = \left\{ y = \sum_{i=1}^N \theta_i a_i \mid \theta \geq 0, \sum_{i=1}^N \theta_i = 1 \right\}. \quad (15)$$

The set $\text{conv}\{a_1, \dots, a_N\}$ is called an $(N - 1)$ -simplex, or simply a simplex, if $\{a_1, \dots, a_N\}$ is affinely independent. The vertices of a simplex are a_1, \dots, a_N . Given a full-dimensional simplex, i.e., an $(N - 1)$ -simplex lying in \mathbb{R}^{N-1} (or $M = N - 1$), its volume can be determined by

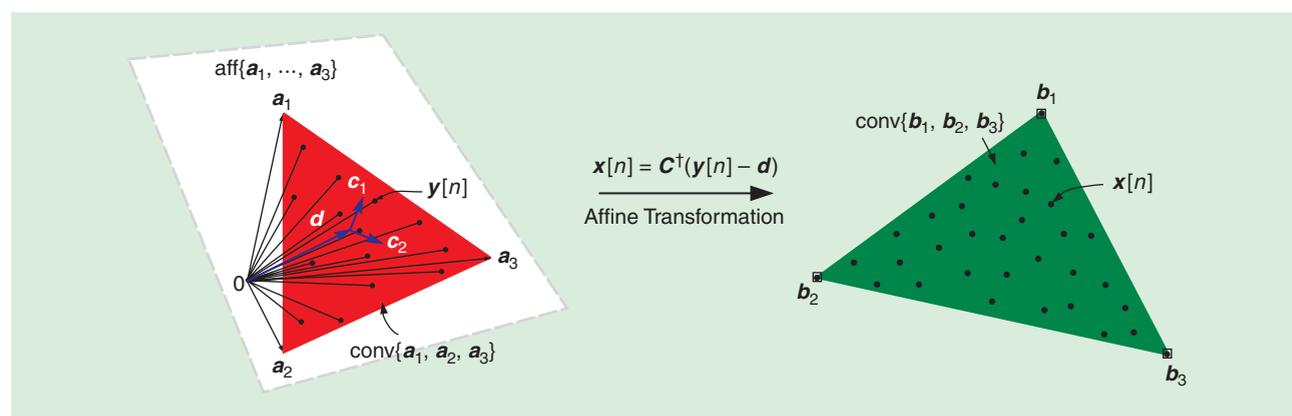
$$\text{vol}(a_1, \dots, a_N) = c \left| \det \begin{pmatrix} a_1 & \dots & a_N \\ 1 & \dots & 1 \end{pmatrix} \right| \quad (16a)$$

$$= c \left| \det \{[a_1 - a_N, \dots, a_{N-1} - a_N]\} \right|, \quad (16b)$$

where $c = 1/(N - 1)!$ For the mathematical details of the above concepts, readers are referred to the literature [35].

CONVEX GEOMETRY IN HYPERSPECTRAL SIGNALS

There is a strong connection between convex analysis and hyperspectral signals. To see it, consider the signal model (1)–(3) in the noiseless case. By comparing the model and the definition of convex hull in (15), we observe that



[FIG2] The convex geometry of hyperspectral signals.

$$y[n] \in \text{conv}\{a_1, \dots, a_N\}, \text{ for all } n = 1, \dots, L,$$

i.e., each measured hyperspectral pixel $y[n]$ is a convex combination of the endmember signatures a_1, \dots, a_N . Also, the set $\text{conv}\{a_1, \dots, a_N\}$ is a simplex, since $\{a_1, \dots, a_N\}$ is linearly independent (and thus affinely independent). The left-hand side of Figure 2 gives a vector space illustration for the case of $N = 3$. As can be seen, $\text{conv}\{a_1, \dots, a_N\}$ is a triangle; note that $\text{conv}\{a_1, \dots, a_N\}$ is a tetrahedron for $N = 4$, and so forth. Also, every $y[n]$ is enclosed by the triangle, and the corners of the triangle, or more formally, the vertices of $\text{conv}\{a_1, \dots, a_N\}$, are the true endmember signatures a_1, \dots, a_N . This observation is simple, but gives a very powerful implication—if we can find all the vertices of $\text{conv}\{a_1, \dots, a_N\}$ from the observation $\{y[n]\}_{n=1}^L$, then blind HU is solved.

Intuitively speaking, CG-based blind HU amounts to finding a set of vectors, say, $\{\hat{a}_1, \dots, \hat{a}_N\}$, such that the corresponding simplex $\text{conv}\{\hat{a}_1, \dots, \hat{a}_N\}$ gives a best fitting to the true endmembers' simplex $\text{conv}\{a_1, \dots, a_N\}$. The previously reviewed pure pixel search algorithms are among one class of such CG solutions; the idea is that pure pixels, if they exist, are also vertices of $\text{conv}\{a_1, \dots, a_N\}$. Hence, pure pixel search is also vertex search in CG, under the pure pixel assumption. Now, we are interested in a different approach where simplex volume is used as the metric to find the best-fitting simplex. Moreover, the pure pixel assumption will not be assumed during the development. We should nevertheless mention a subtle point that the pure pixel assumption will come back when we discuss endmember identifiability.

Before proceeding to the main developments, it is essential for us to introduce a concept related to the affine nature of $y[n]$. Since $y[n] \in \text{conv}\{a_1, \dots, a_N\}$, it also holds true that $y[n] \in \text{aff}\{a_1, \dots, a_N\}$; cf. (13). By the equivalent affine hull representation in (14), we can write

$$y[n] = Cx[n] + d, \tag{17}$$

for some $C \in \mathbb{R}^{M \times (N-1)}$, $\text{rank}(C) = N - 1$, $d \in \mathbb{R}^M$, $x[n] \in \mathbb{R}^{N-1}$, $n = 1, \dots, L$. Suppose that (C, d) is known, and consider the inverse of (17) with respect to (w.r.t.) $x[n]$

$$x[n] = C^\dagger(y[n] - d). \tag{18}$$

From the signal model (1)–(3), it is easy to show that

$$x[n] = \sum_{i=1}^N b_i s_i[n] = Bs[n], \tag{19}$$

where $b_i = C^\dagger(a_i - d) \in \mathbb{R}^{N-1}$, $i = 1, \dots, N$, and $B = [b_1, \dots, b_N] \in \mathbb{R}^{(N-1) \times N}$. We see that (19) takes exactly the same form as the original model (1), but its vector dimension is $N - 1$, which is less than M . Also, $\text{conv}\{b_1, \dots, b_N\}$ is a full-dimensional simplex [36]. Therefore, (19) is a dimension-reduced equivalent model for hyperspectral signals, where the CG structure is preserved. We will employ the equivalent model (19) in our subsequent CG developments. The transformation for the equivalent model is illustrated in Figure 2.

We should discuss how the affine set variable (C, d) is obtained in practice. Since there is no prior knowledge on

$\{a_1, \dots, a_N\}$, we must estimate (C, d) from the observation $\{y[n]\}_{n=1}^L$. This can be done by solving an affine set fitting (ASF) problem

$$\min_{\substack{C, d, \{x[n]\}_{n=1}^L \\ \text{rank}(C) = N-1}} \sum_{n=1}^L \|y[n] - Cx[n] - d\|_2^2, \tag{20}$$

where the rationale is to find an affine set that gives the best fitting w.r.t. the measured pixels $y[n]$, given knowledge of N ; see [36] for details. The ASF solution is as follows. Let $\mu_y = 1/L \sum_{n=1}^L y[n]$ and $\Phi_y = 1/L \sum_{n=1}^L (y[n] - \mu_y)(y[n] - \mu_y)^T$ be the sample mean and sample covariance of $y[n]$, respectively. Also, let q_i be the i th principal eigenvector of Φ_y . The solution to (20) is given by $C = [q_1, \dots, q_{N-1}]$, $d = \mu_y$. There is an interesting coincidence here—the ASF solution is exactly the same as that of principal component analysis (PCA), which is a commonly used DR preprocessing procedure. While ASF and PCA turn out to be equivalent, one should note that they were derived from different principles: ASF is deterministic and concerned with CG-preserving transformation, while PCA is statistical and does not exploit CG.

SIMPLEX VOLUME MAXIMIZATION

This subsection focuses on the simplex volume maximization approach. This approach considers the following problem:

$$\begin{aligned} & \max_B \text{vol}(B) \\ & \text{s.t. } b_i \in \text{conv}\{x[1], \dots, x[L]\}, i = 1, \dots, N. \end{aligned} \tag{21}$$

We will call (21) VolMax for convenience. A picture is illustrated in Figure 3(a) to help us explain the aim of (21). We intend to find a best-fitting simplex, $\text{conv}\{b_1, \dots, b_N\}$, by maximizing its volume while keeping it inside $\text{conv}\{x[1], \dots, x[L]\}$. One can imagine that if the pure pixel assumption holds, then $\text{conv}\{x[1], \dots, x[L]\}$ is also the true endmembers' simplex and the maximum volume simplex should perfectly match the latter—this is Winter's intuition when he first introduced VolMax [7].

We are interested in simple optimization schemes for processing VolMax. Two such schemes are described as follows.

SUCCESSIVE VOLUME MAXIMIZATION

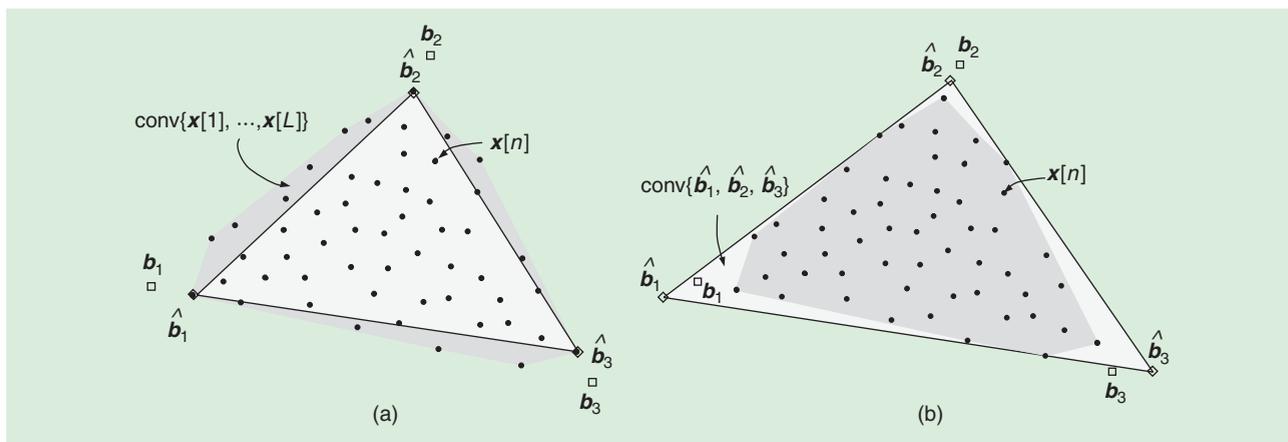
To facilitate our description, let

$$F = \begin{bmatrix} b_1 & \dots & b_N \\ 1 & \dots & 1 \end{bmatrix}, f_i = \begin{bmatrix} b_i \\ 1 \end{bmatrix}, \bar{x}[n] = \begin{bmatrix} x[n] \\ 1 \end{bmatrix}.$$

It can be shown that [20], [37]

$$|\det(F)|^2 = \prod_{k=1}^N \|P_{F_{1:(k-1)}}^\perp f_k\|_2^2, \tag{22}$$

where $F_{1:i} \in \mathbb{R}^{N \times i}$ denotes a submatrix of F , obtained by picking the first i columns of F . We see from the simplex volume formula in (16a) that maximizing $\text{vol}(B)$ is the same as maximizing (22). In successive volume maximization (SVMAX) [20] (also [37]), the principle is to exploit the successive structure of (22) to recursively generate an approximate solution to (21). Specifically, we carry out the following heuristic: for $k = 1, \dots, N$, determine an estimate



[FIG3] (a) Simplex volume maximization. (b) Simplex volume minimization.

$$\hat{b}_k = \arg \max_{b_k} \|P_{\hat{F}_{1:(k-1)}}^\perp f_k\|_2^2 \quad \text{s.t. } b_k \in \text{conv}\{x[1], \dots, x[L]\}, \quad (23)$$

where $\hat{F}_{1:(k-1)}$ is defined in the same way as $F_{1:(k-1)}$, with b_i replaced by \hat{b}_i for all i . Essentially, we estimate one endmember \hat{b}_k based on the previous endmember estimates $\hat{b}_1, \dots, \hat{b}_{k-1}$ and partial maximization of (22). Let us complete the SVMAX algorithm by giving the solution to (23)

$$\hat{b}_k = x[\hat{\ell}_k], \quad \hat{\ell}_k = \arg \max_{n=1, \dots, L} \|P_{\hat{F}_{1:(k-1)}}^\perp \bar{x}[n]\|_2^2, \quad (24)$$

see [20]. Intriguingly, we have seen this algorithm before—SPA in the previous section. To explain, first note that $\bar{x}[n]$ can be expressed as $\bar{x}[n] = F s[n]$, an LMM form. If we apply SPA to $\{\bar{x}[n]\}_{n=1}^L$ to retrieve F , then the resulting SPA is exactly the same as SVMAX. Hence, we conclude that SVMAX is also a pure pixel search algorithm, and SPA has a “dual” identity in VolMax.

SUCCESSIVE N-FINDR

We consider an alternative scheme based on alternating optimization (AO). The idea is to optimize (21) w.r.t. one b_i at a time, while fixing other variables $\{b_j\}_{j \neq i}$. To be specific, given a starting point $\hat{B} = [\hat{b}_1, \dots, \hat{b}_N]$, we update each \hat{b}_k via

$$\hat{b}_k := \arg \max_{b_k} \text{vol}([\hat{B}_{-k}, b_k]) \quad \text{s.t. } b_k \in \text{conv}\{x[1], \dots, x[L]\} \quad (25)$$

for $k = 1, \dots, N$, where \hat{B}_{-k} denotes a submatrix of \hat{B} in which the k th column is removed. Also, we repeat the AO cycle in (25) until some stopping rule (e.g., almost no volume increase) is satisfied. The updates in (25) have a closed form

$$\hat{b}_k = x[\hat{\ell}_k], \quad \hat{\ell}_k = \arg \max_{n=1, \dots, L} \|P_{\hat{F}_{-k}}^\perp \bar{x}[n]\|_2^2, \quad (26)$$

where (26) is obtained by using (22) to turn (25) to (23) (with a proper index reordering), and then applying (24). We call the resulting algorithm successive N-FINDR (SC-N-FINDR) since it is very similar to the SC-N-FINDR proposed in [38]. The pseudocode of SC-N-FINDR is given in Algorithm 2. Note that for

initialization, we can use another algorithm, e.g., SVMAX, or do so randomly. There are several interesting connections here. First, SC-N-FINDR performs pure pixel search. Following [20, Prop. 1], it can be shown that in the noiseless case and under the pure pixel assumption, SC-N-FINDR may perfectly identify all the endmembers’ pure pixels within one AO cycle. Second, since $\bar{x}[n] = F s[n]$, we see from (26) that SC-N-FINDR is performing nulling—this time for all other endmember estimates \hat{F}_{-k} ; cf. the nulling in SPA in (9a). Thus, SC-N-FINDR is also a nulling-based algorithm. Third, we notice that each AO cycle in SC-N-FINDR is essentially the same as the SPA postprocessing strategy we briefly discussed in the section “Successive Projections Algorithm,” which is provably robust against noise.

Algorithm 2 SC-N-FINDR

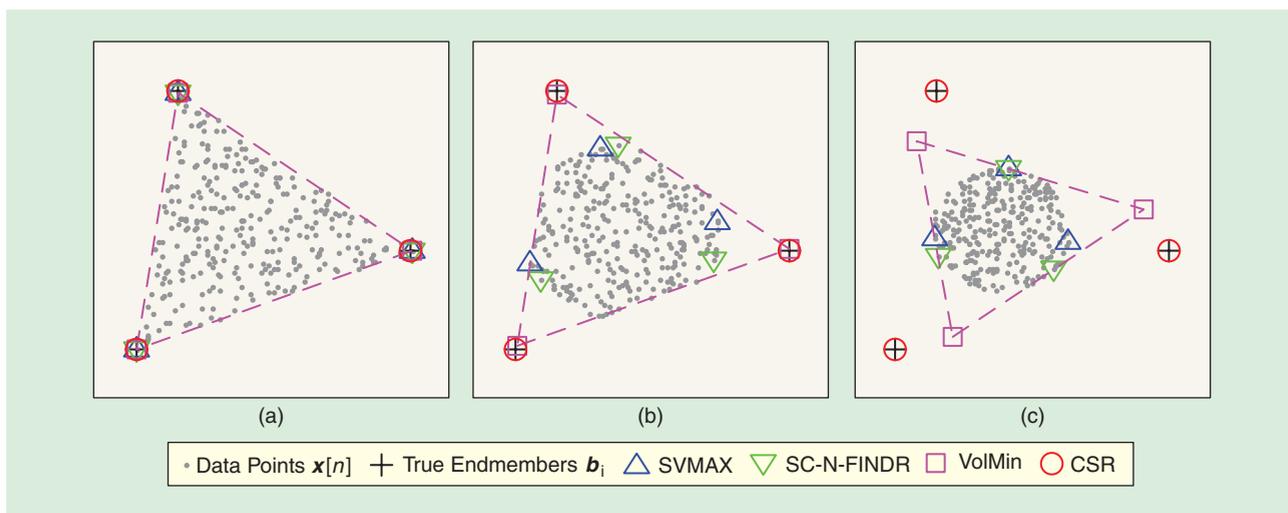
```

input  $\{x[n]\}_{n=1}^L, N, \hat{B}$  (a starting point)
1: repeat
2:   for  $k = 1, \dots, N$  do
3:      $\hat{F} := [\hat{B}^T \mathbf{1}^T]^T$ 
4:      $\hat{\ell}_k := \arg \max_{n=1, \dots, L} \|P_{\hat{F}_{-k}}^\perp \bar{x}[n]\|_2^2$ 
5:      $\hat{b}_k := x[\hat{\ell}_k]$ 
6:   end for
7: until a stopping rule is satisfied
output  $\hat{B} = [\hat{b}_1, \dots, \hat{b}_N]$ 

```

VolMax-based solutions, such as the SVMAX and SC-N-FINDR algorithms above, are usually simple and efficient to implement. Some further discussions are in order.

1) Historically, Winter mainly used VolMax to devise the N-FINDR concept [7] for pure pixel search. There, the intuition is to update one endmember estimate at a time to iteratively increase the volume. N-FINDR is now a popularized algorithm class in blind HU, where we can find many N-FINDR implementation variants in the literature; see [2], [20], and [38]. The SC-N-FINDR we just illustrated is just among one of the many N-FINDR variants, although we have revealed that SC-N-FINDR has several good characteristics.



[FIG4] Numerical comparison of VolMax, VolMin, and sparse regression solutions.

2) VolMax is a provably sound criterion from an endmember identifiability viewpoint. Specifically, the optimal solution of (21) is uniquely the true endmembers' signatures in the noiseless case and under the pure pixel assumption [20]. Also, in this setup, the optimal solution can be easily retrieved by either SC-N-FINDR or SVMAX. However, we should note a fundamental caveat—that SC-N-FINDR and SVMAX are not globally optimal solvers of (21), say, in the presence of noise and/or without pure pixels. In fact, (21) is NP-hard in general [37].

SIMPLEX VOLUME MINIMIZATION

We turn our attention to the simplex volume minimization approach, or simply VolMin, which was first pursued by Craig [4] and Boardman [5] in the blind HU context. VolMin is different from VolMax. It performs simplex fitting by finding a simplex that encloses all the measured pixels, while yielding the minimum volume. This is illustrated in Figure 3(b). Mathematically, VolMin can be formulated as

$$\begin{aligned} & \min_B \text{vol}(B) \\ & \text{s.t. } x[n] \in \text{conv}\{b_1, \dots, b_N\}, n = 1, \dots, L. \end{aligned} \quad (27)$$

VolMin is generally recognized as a more powerful approach than VolMax. Let us illustrate this numerically, before describing VolMin optimization schemes. We simulated a noiseless, three endmember case, where the endmembers were taken from a spectral library [39] and the abundances synthetically generated. Figure 4(a) shows a scenario where the pure pixel assumption holds. We see that both VolMax (via SVMAX or SC-N-FINDR) and VolMin perfectly identify the true endmembers. Figure 4(b) shows another scenario where pure pixels are missing. VolMax is seen to fail, while VolMin can still give accurate endmember estimates. Readers are referred to [2], [20], [36], and [40]–[43] for more numerical comparisons and real-data experiments. Simply speaking, VolMin is numerically found to be robust against lack of pure pixels.

Let us now discuss how VolMin is optimized. VolMin does not have simple closed-form schemes as in VolMax, and requires

numerical optimization. In fact, the VolMin problem in (27) is more difficult to handle; a major obstacle is with the simplex constraints in (27), which are nonconvex. This issue can be overcome by transforming the simplex to a polyhedron (see, e.g., [35, pp. 32–33]). To help the reader understand the idea, an illustration is given in Figure 5. We see that a simplex can be equivalently represented by an intersection of halfspaces, i.e., a polyhedron. More precisely, the following equivalence holds for an affinely independent $\{b_1, \dots, b_N\}$ [36]

$$x[n] \in \text{conv}\{b_1, \dots, b_N\} \iff Hx[n] - g \geq 0, (Hx[n] - g)^T \mathbf{1} \leq 1, \quad (28)$$

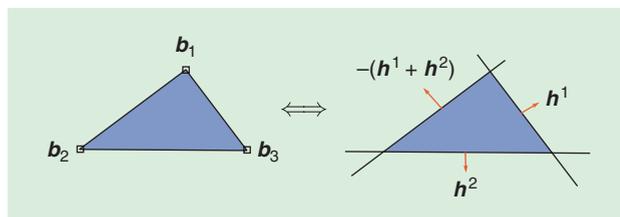
where the RHS is a polyhedron, and

$$H = [b_1 - b_N, \dots, b_{N-1} - b_N]^{-1}, \quad g = Hb_N. \quad (29)$$

By the change of variables in (29), and noting (28) and (16b), we can recast (27) as

$$\begin{aligned} & \max_{H,g} |\det(H)| \\ & \text{s.t. } Hx[n] - g \geq 0, (Hx[n] - g)^T \mathbf{1} \leq 1, n \in \mathcal{L}, \end{aligned} \quad (30)$$

where $\mathcal{L} = \{1, \dots, L\}$. The equivalent VolMin problem in (30) is arguably easier to handle than the original in (27). Specifically, the constraints in (30), which form a data-enclosing polyhedron, are linear (and convex). However, there is still one obstacle—the objective function $|\det(H)|$ is nonconvex. Current



[FIG5] The transformation of a simplex to a polyhedron.

state-of-the-art methods for VolMin tackle this issue by successive convex approximation. Specifically, one can apply iterative linear approximation to the objective function [40], [44]. Another alternative is to perform row-by-row AO w.r.t. (H, g) [36]. These two schemes both operate by solving a sequence of convex optimization problems; see [36], [40], and [44] for the details and comparison.

We complete this subsection by the following comments.

1) As mentioned above, numerical evidence suggests that VolMin may be able to identify the true endmembers accurately in the absence of pure pixels. By analysis, it is known that in the noiseless case, the optimal solution of VolMin is uniquely the true endmembers' signatures if the pure pixel assumption holds [36]. A proof for the no pure pixel case is currently unavailable and is an open fundamental question.

2) While VolMin is deterministic and geometric based, it has a dual identity in stochastic maximum-likelihood (ML) estimation. Specifically, consider the noiseless case, and assume that every abundance vector $s[n]$ is i.i.d. uniformly distributed on the support of unit simplex \mathcal{S} . Then, it can be shown that the corresponding ML estimator is the same as the equivalent VolMin problem in (30) [42]. Note that the authors in [42] also consider a generalization where the abundance prior distribution is nonuniform.

FURTHER DISCUSSION

The CG framework presented above is based on exploitation of the simplex $\text{conv}\{a_1, \dots, a_N\}$. There is an alternative CG formulation where the simplex $\text{conv}\{0, a_1, \dots, a_N\}$ is utilized [40]–[42]; the concepts are identical, though the resulting algorithms exhibit minor differences. Readers should also note other CG interpretations, e.g., [45]. For tutorial purposes, we have focused on the noiseless case only. In the presence of spectrally i.i.d. noise, the ASF preprocessing stage (or equivalently PCA) can be shown to be noise resistant. However, for non-i.i.d. noise, HySime [46] may provide better DR performance. Moreover, both VolMax and VolMin can be modified to improve robustness against noise; e.g., soft constraints [41], chance constraints [43], and robust max-min volume [20]. CG is known to be sensitive to outliers. A robust ASF can be used to identify and discard outliers, before they get into the data [47]. Soft constraints also help “desensitize” VolMin w.r.t. the outliers [41].

DICTIONARY-BASED SEMIBLIND HU

This section describes a relatively new development, where HU is performed by using spectral libraries and techniques arising in compressive sensing (CS). This approach also has a link to sensor array processing in SP, as we will discuss.

SPARSE REGRESSION

When performing blind HU, we generally assume no information on the spectral shapes of the true endmember signatures. The latter is not totally true. In geoscience and remote sensing, a tremendous amount of effort has been spent on measuring and recording spectral samples of many different materials,

which has resulted in spectral libraries for various research purposes. For example, the U.S. Geological Survey (USGS) Library, which has taken over 20 years to assemble, contains more than 1,300 spectral samples covering materials such as minerals, rocks, liquids, artificial materials, vegetations, and even microorganisms [39]. Such valuable knowledge base can be turned to blind HU purposes, or more precisely, semiblind HU.

A slight abuse of notations is required to explain the semiblind formulation. We redefine $A = [a_1, \dots, a_K] \in \mathbb{R}^{M \times K}$ as a dictionary of K hyperspectral samples, where each a_i corresponds to one material (each a_i is also assumed to have been appropriately processed, e.g., atmospherically compensated). We assume that the dictionary A is known, obtained from an available spectral library, and that the true endmembers in each measured pixel $y[n]$ are covered by the dictionary. The measured pixels in the noiseless case (again, for tutorial purposes) can then be represented by

$$y[n] = \sum_{i \in S_n} a_i s_i[n], \quad (31)$$

where $S_n \subseteq \{1, \dots, K\}$ is an index subset that indicates the materials present in the measured pixel $y[n]$, and $s_i[n] > 0, i \in S_n$ are the corresponding abundances. In this representation, note that the sum-to-one constraint $\sum_{i \in S_n} s_i[n] = 1$ may not hold; the measurement conditions of library samples and the actual scene are often different and this can introduce scaling inconsistencies between the library samples and true endmembers. By also letting $s_i[n] = 0$ for all $i \notin S_n$, (31) can be written as

$$y[n] = A s[n], \quad (32)$$

where $s[n] = [s_1[n], \dots, s_K[n]]^T \in \mathbb{R}^K$ is now a sparse abundance vector. The problem now is to recover $s[n]$ from $y[n]$. This is not trivial because we often have $K > M$ and the corresponding system in (32) is underdetermined. However, we know beforehand that $s[n]$ have only a few nonzero components, since the number of materials present in one pixel is often very small, typically within five. Hence, a natural formulation for the semiblind HU problem is to find the sparsest $s[n]$ for the representation in (32). This inference problem turns out to be identical to that investigated in CS, where the objective is to recover a sparse representation of a signal on a given frame from compressive measurements [48]. This connection allows us to capitalize on the wealth of theoretical and algorithmic results available in the CS area.

The sparse regression (SR) problem we describe above can be formulated as

$$\min_{s[n]} \|s[n]\|_0 \quad \text{s.t. } y[n] = A s[n], \quad (33)$$

for each $n = 1, \dots, L$, where $\|s[n]\|_0$ denotes the number of nonzero elements in $s[n]$. The above SR problem possesses provably good endmember identifiability. Specifically, (33) is known to have a unique solution if the true sparse abundance vector $s[n]$ satisfies

$$\|s[n]\|_0 < \frac{1}{2} \cdot \text{spark}(A), \quad (34)$$

where $\text{spark}(A)$ is the smallest number of linearly dependent columns of A [49]. Since every $s[n]$ is highly sparse by nature, (34) should hold in practice. The consequent implication is meaningful—the SR problem (33) can perfectly identify all the true endmembers in general.

While the SR approach sounds promising, there are challenges. Since (33) is NP-hard in general, it is natural to seek approximate solutions. Let us consider the popularized ℓ_1 relaxation solution to (33):

$$\min_{s[n]} \|s[n]\|_1 \quad \text{s.t.} \quad y[n] = As[n], \quad (35)$$

which is convex and has efficient solvers. The CS literature has a series of analysis results telling when (35) gives the same solution as (33), or simply sufficient conditions for exact recovery. Those sufficient conditions usually depend on the conditioning of A . For example, one sufficient exact recovery condition for (35) is $\|s[n]\|_0 < (1/2)(1 + \mu^{-1}(A))$, where

$$\mu(A) = \max_{\substack{1 \leq i, j \leq K \\ i \neq j}} \frac{|a_i^T a_j|}{\|a_i\|_2 \|a_j\|_2} \quad (36)$$

is called the mutual coherence of A [49]. Unfortunately, spectral libraries in practice are strongly correlated, yielding $\mu(A)$ almost being one [50]. A similar issue also occurs in other sufficient conditions, particularly in the restricted isometry property [48]. Thus, one may not obtain a desirable SR solution from a straight ℓ_1 relaxation application.

However, all is not lost. Recall that every $s[n]$ is, by nature, nonnegative. Let us consider a nonnegative ℓ_1 relaxation problem, which is (35) plus the nonnegative constraint $s[n] \geq 0$. As it turns out, exploiting nonnegativity helps a lot. There is a large amount of experimental evidence that indicates that nonnegative ℓ_1 relaxation can yield useful unmixing results [2], [50], [51]. Also, nonnegative ℓ_1 relaxation is theoretically proven to be able to give rather sparse solutions for certain classes of A [52]. Although the above noted theoretical result does not give a direct answer to exact recovery under highly correlated libraries, it gives good insight on the capability of nonnegative ℓ_1 relaxation.

We can also combat the spectral library mutual coherence issue by using the multiple-measurement vector (MMV) formulation [53], which exploits the fact that in a given data set all the spectral vectors are generated by the same subset of library signatures, corresponding to the endmember signatures. Let $S = [s[1], \dots, s[L]] \in \mathbb{R}^{K \times L}$ and $Y = [y[1], \dots, y[L]] \in \mathbb{R}^{M \times L}$, so that we can write $Y = AS$. Also, define $\|S\|_{\text{row-0}}$ to be the number of nonzero rows in S ; i.e., $\|S\|_{\text{row-0}} = |\text{rowsupp}(S)|$, $\text{rowsupp}(S) = \{1 \leq i \leq K \mid s^i \neq 0\}$. We consider a collaborative SR (CSR) problem [54]

$$\min_S \|S\|_{\text{row-0}} \quad \text{s.t.} \quad Y = AS, \quad (37)$$

where the rationale is to use the whole set of measured pixels, rather than one, to strengthen SR performance. It is interesting to note that $\|S\|_{\text{row-0}}$ also represents the number of endmembers. Like the previous SR problem, we can apply a convex relaxation to CSR by replacing $\|S\|_{\text{row-0}}$ in (37) by $\|S\|_{2,1}$, where $\|S\|_{p,q} = (\sum_{i=1}^k \|s^i\|_p^q)^{1/q}$. In theory, there is no extra benefit in using the CSR or MMV formulation in the worst-case sense (think about a special and rather unrealistic case where $s[1] = \dots = s[L]$) [53]. However, an average analysis in [55] gives an implication that increasing the number of measurements (or pixels here) can significantly reduce the probability of recovery failure. In practice, this has been found to be so. Also, the nonnegativity constraint $S \geq 0$ can be incorporated in (37) to improve performance.

A practical SR or CSR solution should also cater for the presence of noise. For CSR, the following alternative convex relaxation formulation may be used to provide HU [54]

$$\min_{S \geq 0} \|Y - AS\|_F^2 + \lambda \|S\|_{2,1} \quad (38)$$

for some constant $\lambda > 0$. The rationale is to seek an LS data fitting, rather than exact, with a sparsity-promoting regularizer $\lambda \|S\|_{2,1}$. It is important to note that while (38) is convex, it is a large-scale optimization problem. An efficient solver for (38) is provided in [54], where a divide-and-conquer optimization strategy, specifically, the alternating direction method of multipliers (ADMM), was implemented.

At this point readers may be wondering: How do we compare SR- and CG-based solutions? Simply speaking, CG relies on exploitation of simplex structures, while SR does not. To illustrate, consider the previous numerical example in Figure 4. In Figure 4(c), we generated a heavily mixed (and noiseless) scenario where data do not possess simplex structures expected in CG. It is seen that even VolMin fails in this scenario. However, CSR, which was run under the USGS Library with 498 spectral signatures, is seen to be able to identify the true endmembers perfectly. Note that the true endmember signatures were taken from the same library, which makes the setting slightly ideal. It would not be too surprising that if the library fails to cover all true endmember signatures (e.g., a new material), then SR solutions would fail. For further numerical results and real-data experiments, see [2], [50], [54], [56], and [57].

SENSOR ARRAY PROCESSING MEETS SEMIBLIND HU

MMV is a powerful concept that has been applied to estimation problems in statistical SP and sensor array processing [58]. Curiously, a classical concept originated from sensor array processing, specifically, subspace methods, also finds its way to MMV research [59]—this provides yet another opportunity for semiblind HU [56].

The idea is simple for readers who are familiar with subspace methods or sensor array processing; or, see classical literatures such as [17]. Consider the block model $Y = AS$ (again, assuming no noise). Let $S = \text{rowsupp}(S)$ be the set of indices of active materials in the measured data Y , and A_S be a submatrix

of A whose columns are $\{a_i\}_{i \in S}$. Note that A_S is the true endmember matrix. Let us assume that $\{s^i\}_{i \in S}$, the set of true abundance maps, is linearly independent; in practice this refers to situations where the abundance maps are sufficiently different. Then, one can easily deduce that $\mathcal{R}(Y) = \mathcal{R}(A_S)$, where \mathcal{R} denotes the range space of its argument. The above expression implies that

$$P_V^\perp a_k = 0 \iff k \in S \quad (39)$$

for all $1 \leq k \leq K$, as far as $\{a_k\} \cup \{a_i\}_{i \in S \setminus \{k\}}$ is linearly independent for any $1 \leq k \leq K$. Since the latter holds for $|S| + 1 < \text{spark}(A)$, we have the following endmember identifiability condition for (39):

$$\|S\|_{\text{row-0}} < \text{spark}(A) - 1. \quad (40)$$

Remarkably, with the mild assumption of linear independence of $\{s^i\}_{i \in S}$, we can achieve such provably good endmember identifiability by the simple subspace projection in (39).

In practice, the identification in (39) can be implemented by the classical multiple signal classification (MUSIC) method [17]; see [56] for implementation details.

FURTHER DISCUSSION

There are a few more points to note.

- 1) As a side advantage, the SR approach does not require knowledge of the number of endmembers N . Note that this does not apply to the subspace approach, which often requires knowledge of N to construct subspace projections.
- 2) Hyperspectral signals are very often piecewise smooth w.r.t. their three dimensional domain (one spectral dimension plus two spatial dimensions). Therefore, one can exploit such spatial/spectral contextual information for improving SR performance by applying piecewise smooth regularization, such as total variations (TVs) [57], on top of an SR formulation, e.g., (38).
- 3) An interesting (but also elusive) question is whether a given dictionary can truly cover the true endmembers. From an end user's viewpoint, it depends on the scene and whether one can preselect a reliable library for that scene specifically. Moreover, there are concurrent studies that consider learning the dictionary from the data, thereby circumventing these issues [51], [60], [61]. Dictionary learning is an active research topic. It is also related to NMF, to be described in the next section. In addition, there has been interest in using the measured data Y itself as the dictionary for MMV [62]. This self-dictionary MMV (SD-MMV) approach is related to pure pixel search. For example, SPA and VCA can both be derived from SD-MMV [63].

NONNEGATIVE MATRIX FACTORIZATION

This section turns the attention back to blind HU, where we review a class of algorithms known as nonnegative matrix factorization (NMF).

NMF was originally proposed as a linear DR tool for analyzing environmental data [64] and for data mining applications [65]. It is posed as a low-rank matrix approximation problem where, given a data matrix $Y \in \mathbb{R}^{M \times L}$, the task is to find a pair of non-negative matrices $A \in \mathbb{R}^{M \times N}$, $S \in \mathbb{R}^{N \times L}$, with $N < \min\{M, L\}$, that solves

$$\min_{A \geq 0, S \geq 0} \|Y - AS\|_F^2. \quad (41)$$

In blind HU, the connection is that the NMF factors obtained, A and S , can serve as estimates of the endmembers and abundances, respectively (note that endmember spectral signatures are nonnegative by nature). However, there are two problems here. First, (41) is NP-hard in general [66]. For this reason, optimization schemes we see in the current NMF-based blind HU developments are rather pragmatic. We should, however, mention that lately, there are new theory-guided NMF developments in optimization [67], [68]. Second, NMF may not guarantee solution uniqueness. This is a serious issue to the blind HU application, since it means that an NMF solution may not necessarily be the true endmembers and abundances, even in the noiseless case.

In blind HU, NMF is modified to fit the problem better. Roughly speaking, we may unify many NMF-based blind HU developments under one formulation

$$\min_{A \geq 0, S \in \mathcal{S}^L} \|Y - AS\|_F^2 + \lambda \cdot g(A) + \mu \cdot h(S), \quad (42)$$

where $\mathcal{S}^L = \{S \mid s[n] \geq 0, 1^T s[n] = 1, 1 \leq n \leq L\}$, g and h are regularizers, which vary from one work to another, and $\lambda, \mu > 0$ are some constants. In particular, the addition of g and h is to make (42) more well posed through exploitation of the problem natures. Also, for the same reason, we incorporate the unit simplex constraints on S .

In the literature, one can find a plethora of NMF-based blind HU algorithms—each work may use different g, h , modified constraints for simpler implementations (e.g., no constraints on A), and a different optimization algorithm. Our intention here is not to give an extensive coverage of all these developments. Instead, we are interested in several representative NMF-based blind HU formulations, where we will see connections between NMF, CG, and SR. A summary of those formulations is shown in Table 1.

Although we see many choices with the regularizers g and h , the philosophies behind the choices follow a few core principles. For the endmember regularizer g , the principle can be traced back to VolMin in CG. A classical example is minimum volume constrained NMF (MVC-NMF) [69]

$$\min_{A \geq 0, S \in \mathcal{S}^L} \|Y - AS\|_F^2 + \lambda \cdot (\text{vol}(\mathcal{B}))^2, \quad (43)$$

where $\text{vol}(\mathcal{B})$ is the simplex volume corresponding to A , in which $b_i = C^i(a_i - d)$ for all i ; cf. the section "Convex Geometry." MVC-NMF is essentially a variation of the VolMin formulation [see (27)] in the noisy case, with endmember nonnegativity incorporated. As mentioned before, $\text{vol}(\mathcal{B})$ is nonconvex. Iterated constrained endmember (ICE) [70] and sparsity promoting ICE (SPICE) [73] avoid this issue by replacing $(\text{vol}(\mathcal{B}))^2$ with a

[TABLE 1] A SUMMARY OF SOME NMF FORMULATIONS.

ALGORITHM	$g(A)$	$h(S)$	OPTIMIZATION SCHEMES AND REMARKS
MVC-NMF [69]	$\text{vol}^2(\mathbf{C}^t(\mathbf{A} - \mathbf{d}\mathbf{1}^T))$	0	AO + ONE-STEP PROJECTED GRADIENT
ICE [70]	$\sum_{i=1}^{N-1} \sum_{j=i+1}^N \ \mathbf{a}_i - \mathbf{a}_j\ _2^2$	0	AO; UNCONSTRAINED \mathbf{A}
DL [60]	0	$\ \mathbf{S}\ _{1,1}$	AO + ONE-STEP PROJECTED GRADIENT FOR \mathbf{A} ; $\mathbf{S} \geq \mathbf{0}$
$L_{1/2}$ -NMF [71]	0	$\ \mathbf{S}\ _{1/2,1/2}^2$	AO + MULTIPLICATIVE UPDATE
APS [72]	0	$\sum_{n=1}^L \sum_{j \in \mathcal{N}(n)} \ \mathbf{s}[n] - \mathbf{s}[j]\ _1$, WHERE $\mathcal{N}(n)$ IS THE NEIGHBORHOOD PIXEL INDEX SET OF PIXEL n .	AO + ONE-STEP PROJECTED SUBGRADIENT
SPICE [73]	$\sum_{i=1}^{N-1} \sum_{j=i+1}^N \ \mathbf{a}_i - \mathbf{a}_j\ _2^2$	$\sum_{i=1}^N \gamma_i \ \mathbf{s}^i\ _1$	AO; UNCONSTRAINED \mathbf{A} ; ITERATIVELY REWEIGHTED γ_i VIA $\gamma_i := 1/\ \mathbf{S}^{(k-1)}\ _{1,1}$, $1 \leq i \leq N$
CoNMF [74]	$\sum_{i=1}^N \ \mathbf{a}_i - \boldsymbol{\mu}_i\ _2^2$	$\sum_{i=1}^N \ \mathbf{s}^i\ _p^p$, $0 < p \leq 1$	AO + ONE-STEP MAJORIZATION MINIMIZATION; UNCONSTRAINED \mathbf{A}

convex surrogate, specifically, $g(\mathbf{A}) = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \|\mathbf{a}_i - \mathbf{a}_j\|_2^2$, which is the sum of differences between vertices. A similar idea is also adopted in collaborative NMF (CoNMF) [74]; see Table 1.

As for the abundance regularizer h , the design principle usually follows that of sparsity. A good showcasing example, curiously, lies in dictionary learning (DL) [60]

$$\min_{A \geq 0, S \geq 0} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \mu \cdot \|\mathbf{S}\|_{1,1}; \quad (44)$$

note that $\|\mathbf{S}\|_{1,1} = \sum_{n=1}^L \sum_{i=1}^N |s_i[n]|$. The original idea of (44) is to learn the dictionary \mathbf{A} by joint dictionary and sparse signal optimization; cf. the section “Dictionary-Based Semiblind HU” and, in particular, (38). However, (44) can also be seen as an NMF with sparsity-promoting regularization. Following the same spirit, $L_{1/2}$ -NMF [71] uses a nonconvex, but stronger sparsity-promoting regularizer based on the $\ell_{1/2}$ quasinorm. Apart from sparsity, exploitation of spatial contextual information via TV regularization may also be used [72].

The aforementioned connection between DL and NMF provides an additional insight. In DL, the dictionary size is often set to be large, and should be larger than the true number of endmembers; the number of endmembers is instead determined by the row sparsity of \mathbf{S} , i.e., $\|\mathbf{S}\|_{\text{row}-0}$. From an NMF-based blind HU perspective, this means that we can use row sparsity to provide joint endmember number, endmember and abundance estimation. More formally, consider a blind version of the MMV (38)

$$\min_{A \geq 0, S \in S^L} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda \cdot g(\mathbf{A}) + \mu \cdot \|\mathbf{S}\|_{\text{row}-0}, \quad (45)$$

where the number of columns of \mathbf{A} , given by N , is now chosen to be a number greater than the true number of endmembers (say, by overestimating the latter), and we use $\|\mathbf{S}\|_{\text{row}-0}$ to represent the endmember number. SPICE is arguably the first algorithm that explores such opportunity [73]. In SPICE, the abundance regularizer can be expressed as $h(\mathbf{S}) = \sum_{i=1}^N \gamma_i \|\mathbf{s}^i\|_1$ for some weights $\{\gamma_i\}$ that are iteratively updated; this regularizer is a convex surrogate of $\|\mathbf{S}\|_{\text{row}-0}$. CoNMF also aims at row sparsity, using a nonconvex surrogate $h(\mathbf{S}) = \sum_{i=1}^K \|\mathbf{s}^i\|_p^p$, $0 < p \leq 1$ [74].

We should also discuss optimization in NMF-based blind HU. Most NMF-based blind HU algorithms follow a two-block AO strategy, although their implementation details exhibit many differences. Two-block AO optimizes (42) w.r.t. either \mathbf{A} or \mathbf{S} alternately. Specifically, it generates a sequence of iterates $\{(\mathbf{A}^{(k)}, \mathbf{S}^{(k)})\}_k$ via

$$\mathbf{A}^{(k)} = \arg \min_{A \geq 0} \|\mathbf{Y} - \mathbf{A}\mathbf{S}^{(k-1)}\|_F^2 + \lambda \cdot g(\mathbf{A}) \quad (46a)$$

$$\mathbf{S}^{(k)} = \arg \min_{S \in S^L} \|\mathbf{Y} - \mathbf{A}^{(k)}\mathbf{S}\|_F^2 + \mu \cdot h(\mathbf{S}). \quad (46b)$$

Note that if g and h are convex, then (46a)–(46b) are convex and hence can usually be solved efficiently. Moreover, every limit point of $\{(\mathbf{A}^{(k)}, \mathbf{S}^{(k)})\}_k$ is a stationary point of (42) under some fairly mild assumptions [75], [76]. For practical reasons, most algorithms use cheap but inexact updates for (46a) and (46b), e.g., multiplicative update [71], one-step projected gradient or subgradient update [60], [69], [72], and one-step majorization minimization [74]. Convergence to a stationary point of these inexact AO methods has still to be thoroughly analyzed. However, by numerical experience, many NMF-based blind HU algorithms work well under appropriate settings (e.g., using reasonable initializations that can be obtained, e.g., with VCA or N-FINDR).

To summarize, NMF is a versatile approach that has connections to both CG and SR. It leads to a fundamentally hard optimization problem, although practical solutions based on two-block AO usually offer good performance by experience. Also, we should highlight that the more exciting developments of NMF-based blind HU lie in extensions to scenarios such as nonlinear HU [77], EV [78], and multispectral and hyperspectral data fusion [79]. Such extensions may not be easily achieved in other approaches.

CONCLUSIONS

This article provided a tutorial review on blind HU techniques using a fundamental SP perspective. Four major blind HU approaches—pure pixel search, convex geometry, sparse regression, and NMF—have been studied. We briefly compare their advantages and drawbacks. Pure pixel search and VolMax are very simple but require the pure pixel assumption; VolMin is resistant

to lack of pure pixels but still has limitations when data are too heavily mixed; sparse regression holds great potential in unmixing heavily mixed data but one should be aware of its reliance on dictionaries; NMF is a very flexible formulation for blind HU but leads us to a hard optimization problem to solve. Also, real hyperspectral data can be quite elusive at times, where we may be faced with issues such as outliers, modeling errors, and uncertainty in the number of endmembers. Their subsequent effects on the aforementioned approaches could be substantial. On the other hand, the need for meeting these challenges also makes HU continue to be a vibrant and active field of research.

ACKNOWLEDGMENTS

This work was supported in part by the Hong Kong Research Grants Council General Research Fund Project CUHK 415509; Portuguese Science and Technology Foundation, Project PEst-OE/EEI/LA0008/2013; Labex CIMI during visits of J. Bioucas-Dias at the University of Toulouse and National Science Council (R.O.C.) under grant NSC 102-2221-E-007-035-MY2.

AUTHORS

Wing-Kin Ma (wkma@ieee.org) is an associate professor with the Department of Electronic Engineering, the Chinese University of Hong Kong.

José M. Bioucas-Dias (bioucas@lx.it.pt) is an associate professor in the Department of Electronic Engineering, Instituto Superior Técnico, Technical University of Lisbon, Portugal.

Tsung-Han Chan (Th.chan@adsc.com.sg) is a research scientist at the Advanced Digital Sciences Centre, Singapore.

Nicolas Gillis (nicolas.gillis@umons.ac.be) is an assistant professor in the Department of Mathematics and Operational Research, Faculté Polytechnique, Université de Mons, Belgium.

Paul Gader (pgader@cise.ufl.edu) is a professor with the Department of Computer and Information Science and Engineering, University of Florida.

Antonio J. Plaza (aplaza@unex.es) is an associate professor with the Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain.

ArulMurugan Ambikapathi (a.arulmurugan@gmail.com) is a postdoctoral research fellow at the Institute of Communications Engineering and Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan.

Chong-Yung Chi (cychi@ee.nthu.edu.tw) is a professor with the Institute of Communications Engineering and Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan.

REFERENCES

- [1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Mag.*, vol. 19, no. 1, pp. 44–57, 2002.
- [2] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [3] M. D. Craig, "Unsupervised unmixing of remotely sensed images," in *Proc. Australasian Remote Sensing Conf.*, Perth, Australia, 1990, pp. 324–330.
- [4] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, no. 3, pp. 542–552, May 1994.

- [5] J. W. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in *Proc. Summary 4th Annual JPL Airborne Geoscience Workshop*, Dec. 1993, vol. 1, pp. 11–14.

- [6] J. Boardman, F. Kruse, and R. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Proc. Summary JPL Airborne Earth Science Workshop*, Pasadena, CA, Dec. 1995, vol. 1, pp. 23–26.

- [7] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proc. SPIE Conf. Imaging Spectrometry*, Pasadena, CA, Oct. 1999, pp. 266–275.

- [8] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4355–4368, 2009.

- [9] B.-C. Gao, M. J. Montes, C. O. Davis, and A. F. Goetz, "Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean," *Remote Sensing Environ.*, vol. 113, pp. S17–S24, Sept. 2009.

- [10] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. Bermudez, S. McLaughlin, and A. Hero, "Nonlinear unmixing of hyperspectral images," *IEEE Signal Processing Mag.*, vol. 31, no. 1, pp. 82–94, Jan. 2014.

- [11] A. Zare and K. C. Ho, "Endmember variability in hyperspectral analysis," *IEEE Signal Processing Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.

- [12] CVX Research, Inc. (2012, Sept.). CVX: MATLAB software for disciplined convex programming, version 2.0 beta. [Online]. Available: <http://cvxr.com/cvx>

- [13] D. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.

- [14] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Reykjavik, Iceland, 2010, pp. 1–4.

- [15] R. Heylen, D. Burazerovi, and P. Scheunders, "Fully constrained least squares spectral unmixing by simplex projection," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4112–4122, Nov. 2011.

- [16] J. Nascimento and J. Bioucas-Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 1, pp. 175–187, 2005.

- [17] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Hoboken, NJ: Wiley, 2002.

- [18] U. Araújo, B. Saldanha, R. Galvão, T. Yoneyama, H. Chame, and V. Visani, "The successive projections algorithm for variable selection in spectroscopic multicomponent analysis," *Chemometr. Intell. Lab. Syst.*, vol. 57, no. 2, pp. 65–73, 2001.

- [19] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1232–1249, 2003.

- [20] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4177–4193, 2011.

- [21] G. Golub, "Numerical methods for solving linear least squares problems," *Numerische Mathematik*, vol. 7, no. 3, pp. 206–216, June 1965.

- [22] S. Arora, R. Ge, Y. Halpern, D. Mimno, A. Moitra, D. Sontag, Y. Wu, and M. Zhu, "A practical algorithm for topic modeling with provable guarantees," in *Proc. Int. Conf. Machine Learning*, vol. 28, no. 2, 2013, pp. 280–288.

- [23] N. Gillis and S. Vavasis, "Fast and robust recursive algorithms for separable nonnegative matrix factorization," [Online]. Available: <http://arxiv.org/abs/1208.1237>

- [24] A. Ambikapathi, T.-H. Chan, C.-Y. Chi, and K. Keizer, "Hyperspectral data geometry-based estimation of number of endmembers using p-norm-based pure pixel identification algorithm," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 5, pp. 2753–2769, 2013.

- [25] J. Nascimento and J. Bioucas-Dias, "Vertex component analysis: A fast algorithm to extract endmembers spectra from hyperspectral data," in *Pattern Recognition and Image Analysis: 1st Iberian Conf.*, 2003, pp. 626–635.

- [26] J. Nascimento and J. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.

- [27] C.-I. Chang, W. Xiong, H.-M. Chen, and J.-W. Chai, "Maximum orthogonal subspace projection approach to estimating the number of spectral signal sources in hyperspectral imagery," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 3, pp. 504–520, 2011.

- [28] J. Imbrie, "Vector analysis of heavy-mineral data," *Geological Soc. Amer. Bull.*, vol. 75, pp. 1131–1156, Nov. 1964.

- [29] R. Ehrlich and W. E. Full, "Sorting out geology—Unmixing mixtures," in *Use and Abuse of Statistical Methods in the Earth Sciences*, W. Size, Ed. New York: Oxford Univ. Press, 1987, pp. 33–46.

- [30] J. E. Klovan and A. T. Miesch, "Extended CABFAC and QMODEL computer programs for Q-mode factor analysis of compositional data," *Comput. Geosci.*, vol. 1, no. 3, pp. 161–178, 1976.
- [31] W. E. Full, R. Ehrlich, and J. E. Klovan, "Extended QMODEL—Objective definition of external endmembers in the analysis of mixtures," *Math. Geology*, vol. 13, no. 4, pp. 331–344, 1981.
- [32] A. Perczel, M. Hollósi, G. Tusnády, and G. D. Fasman, "Convex constraint analysis: A natural deconvolution of circular dichroism curves of proteins," *Protein Eng.*, vol. 4, no. 6, pp. 669–679, 1991.
- [33] W. Naanaa and J.-M. Nuzillard, "Blind source separation of positive and partially correlated data," *Signal Process.*, vol. 85, no. 9, pp. 1711–1722, 2005.
- [34] T.-H. Chan, W.-K. Ma, C.-Y. Chi, and Y. Wang, "A convex analysis framework for blind separation of non-negative sources," *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 5120–5134, 2008.
- [35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [36] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4418–4432, 2009.
- [37] A. Çivril and M. Magdon-Ismail, "On selecting a maximum volume sub-matrix of a matrix and related problems," *Theor. Comput. Sci.*, vol. 410, no. 47, pp. 4801–4811, 2009.
- [38] C.-C. Wu, S. Chu, and C.-I. Chang, "Sequential N-FINDR algorithms," in *Proc. SPIE*, vol. 7086, 2008, p. 70860C.
- [39] R. Clark, G. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. Sutley. (2007). USGS digital spectral library splib06a: U.S. Geological Survey, Digital Data Series 231. [Online]. Available: <http://speclab.cr.usgs.gov/spectral.lib06>.
- [40] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *Proc. IEEE IGARSS*, Aug. 2008, pp. 250–253.
- [41] J. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. IEEE WHISPERS*, Aug. 2009, pp. 1–4.
- [42] J. Nascimento and J. Bioucas-Dias, "Hyperspectral unmixing based on mixtures of Dirichlet components," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 3, pp. 863–878, 2012.
- [43] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "Chance-constrained robust minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4194–4209, 2011.
- [44] M. B. Lopes, J. C. Wolff, J. Bioucas-Dias, and M. Figueiredo, "NIR hyperspectral unmixing based on a minimum volume criterion for fast and accurate chemical characterisation of counterfeit tablets," *Anal. Chem.*, vol. 82, no. 4, pp. 1462–1469, 2010.
- [45] P. Honeine and C. Richard, "Geometric unmixing of large hyperspectral images: A barycentric coordinate approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 6, pp. 2185–2195, 2012.
- [46] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.
- [47] T.-H. Chan, A. Ambikapathi, W.-K. Ma, and C.-Y. Chi, "Robust affine set fitting and fast simplex volume max-min for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 7, pp. 3982–3997, 2013.
- [48] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–121, 2007.
- [49] D. L. Donoho and M. Elad, "Optimally sparse representation in general (non-orthogonal) dictionaries via l_1 minimization," *Proc. Natl. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003.
- [50] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011.
- [51] J. B. Greer, "Sparse demixing of hyperspectral images," *IEEE Trans. Image Processing*, vol. 21, no. 1, pp. 219–228, 2012.
- [52] D. L. Donoho and J. Tanner, "Sparse nonnegative solution of underdetermined linear equations by linear programming," *Proc. Natl. Acad. Sci.*, vol. 102, no. 27, pp. 9446–9451, 2005.
- [53] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Processing*, vol. 54, no. 12, pp. 4634–4643, 2006.
- [54] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sensing*, to be published.
- [55] Y. C. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Trans. Inform. Theory*, vol. 56, no. 1, pp. 505–519, 2010.
- [56] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "MUSIC-CSR: Hyperspectral unmixing via multiple signal classification and collaborative sparse regression," *IEEE Trans. Geosci. Remote Sensing*, to be published.
- [57] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 11, pp. 4484–4502, 2012.
- [58] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.
- [59] J. M. Kim, O. K. Lee, and J.-C. Ye, "Compressive MUSIC: Revisiting the link between compressive sensing and array signal processing," *IEEE Trans. Inform. Theory*, vol. 58, no. 1, pp. 278–301, 2012.
- [60] A. Charles, B. Olshausen, and C. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 5, pp. 963–978, 2011.
- [61] A. Castrodad, Z. Xing, J. B. Greer, E. Bosch, L. Carin, and G. Sapiro, "Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4263–4281, 2011.
- [62] E. Esser, M. Moller, S. Osher, G. Sapiro, and J. Xin, "A convex model for non-negative matrix factorization and dimensionality reduction on physical space," *IEEE Trans. Image Processing*, vol. 21, no. 7, pp. 3239–3252, 2012.
- [63] X. Fu, W.-K. Ma, T.-H. Chan, J. M. Bioucas-Dias, and M.-D. Iordache, "Greedy algorithms for pure pixels identification in hyperspectral unmixing: A multiple-measurement vector viewpoint," in *Proc. EUSIPCO 2013*, to be published.
- [64] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.
- [65] D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [66] S. A. Vavasis, "On the complexity of nonnegative matrix factorization," *SIAM J. Optim.*, vol. 20, no. 3, pp. 1364–1377, 2009.
- [67] N. Gillis and R. Plemmons, "Sparse nonnegative matrix underapproximation and its application to hyperspectral image analysis," *Lin. Alg. Appl.*, vol. 438, no. 10, pp. 3991–4007, 2013.
- [68] S. Arora, R. Ge, R. Kannan, and A. Moitra, "Computing a nonnegative matrix factorization—Provably," in *Proc. 44th Symp. Theory of Computing*, ACM, 2012, pp. 145–162.
- [69] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sensing*, vol. 45, no. 3, pp. 765–777, 2007.
- [70] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "ICE: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 10, pp. 2085–2095, 2004.
- [71] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4282–4297, 2011.
- [72] A. Zymnis, S.-J. Kim, J. Skaf, M. Parente, and S. Boyd, "Hyperspectral image unmixing via alternating projected subgradients," in *Proc. Asilomar Conf. Signals, Systems and Computers*, 2007, pp. 1164–1168.
- [73] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 446–450, 2007.
- [74] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Collaborative nonnegative matrix factorization for remotely sensed hyperspectral unmixing," in *Proc. IEEE IGARSS*, 2012, pp. 3078–3081.
- [75] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Oper. Res. Lett.*, vol. 26, no. 3, pp. 127–136, 2000.
- [76] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, 2001.
- [77] P. Gader, D. Dranishnikov, A. Zare, and J. Chanussot, "A sparsity promoting bilinear unmixing model," in *Proc. IEEE WHISPERS*, 2012.
- [78] A. Zare and P. Gader, "PCE: Piecewise convex endmember detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 6, pp. 2620–2632, 2010.
- [79] N. Yokoya, T. Yair, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 2, pp. 528–537, 2012.

[Nicolas Dobigeon, Jean-Yves Tournet, Cédric Richard,
José Carlos M. Bermudez, Stephen McLaughlin, and Alfred O. Hero]

Nonlinear Unmixing of Hyperspectral Images



[Models and algorithms]

When considering the problem of unmixing hyperspectral images, most of the literature in the geoscience and image processing areas relies on the widely used linear mixing model (LMM). However, the LMM may be not valid, and other nonlinear models need to be considered, for instance, when there are multiscattering effects or intimate interactions. Consequently, over the last few years, several significant contributions have been proposed to overcome the limitations inherent in the LMM. In this article, we present an overview of recent advances in nonlinear unmixing modeling.

Digital Object Identifier 10.1109/MSP.2013.2279274

Date of publication: 5 December 2013

MOTIVATION FOR NONLINEAR MODELS

Spectral unmixing (SU) is widely used for analyzing hyperspectral data arising in areas such as remote sensing, planetary science chemometrics, materials science, and other areas of microspectroscopy. SU provides a comprehensive and quantitative mapping of the elementary materials that are present in the acquired data. More precisely, SU can identify the spectral signatures of these materials (usually called *endmembers*) and can estimate their relative contributions (known as *abundances*) to the measured spectra. Similar to other blind source separation tasks, the SU problem is naturally ill posed and admits a wide range of admissible solutions. As a consequence, SU is a challenging problem that has received considerable attention in the remote sensing, signal, and image processing communities [1]. Hyperspectral data analysis can be supervised, when the endmembers are known, or unsupervised, when they are unknown. Irrespective of the case,

most SU approaches require the definition of the mixing model underlying the observations. A mixing model describes in an analytical fashion how the endmembers combine to form the mixed spectrum measured by the sensor. The abundances parametrize the model. Given the mixing model, SU boils down to estimating the inverse of this formation process to infer the quantities of interest, specifically the endmembers and/or the abundances, from the collected spectra. Unfortunately, defining the direct observation model that links these meaningful quantities to the measured data is a nontrivial issue, and requires a thorough understanding of complex physical phenomena. A model based on radiative transfer (RT) could accurately describe the light scattering by the materials in the observed scene [2] but would lead to very complex unmixing problems. Fortunately, invoking simplifying assumptions can lead to exploitable mixing models.

When the mixing scale is macroscopic and each photon reaching the sensor has interacted with just one material, the measured spectrum $\mathbf{y}_p \in \mathbb{R}^L$ in the p th pixel can be accurately described by the LMM

$$\mathbf{y}_p = \sum_{r=1}^R a_{r,p} \mathbf{m}_r + \mathbf{n}_p, \quad (1)$$

where L is the number of spectral bands, R is the number of endmembers present in the image, \mathbf{m}_r is the spectral signatures of the r th endmember, $a_{r,p}$ is the abundance of the r th material in the p th pixel and \mathbf{n}_p is an additive term associated with the measurement noise and the modeling error. The abundances can be interpreted as the relative areas occupied by the materials in a given image pixel [3]. Thus it is natural to consider additional constraints regarding the abundance coefficients $a_{r,p}$

$$\begin{cases} a_{r,p} \geq 0, & \forall p, \forall r \\ \sum_{r=1}^R a_{r,p} = 1, & \forall p. \end{cases} \quad (2)$$

In that case, SU can be formulated as a constrained blind source separation problem, or constrained linear regression, depending on the prior knowledge available regarding the endmember spectra.

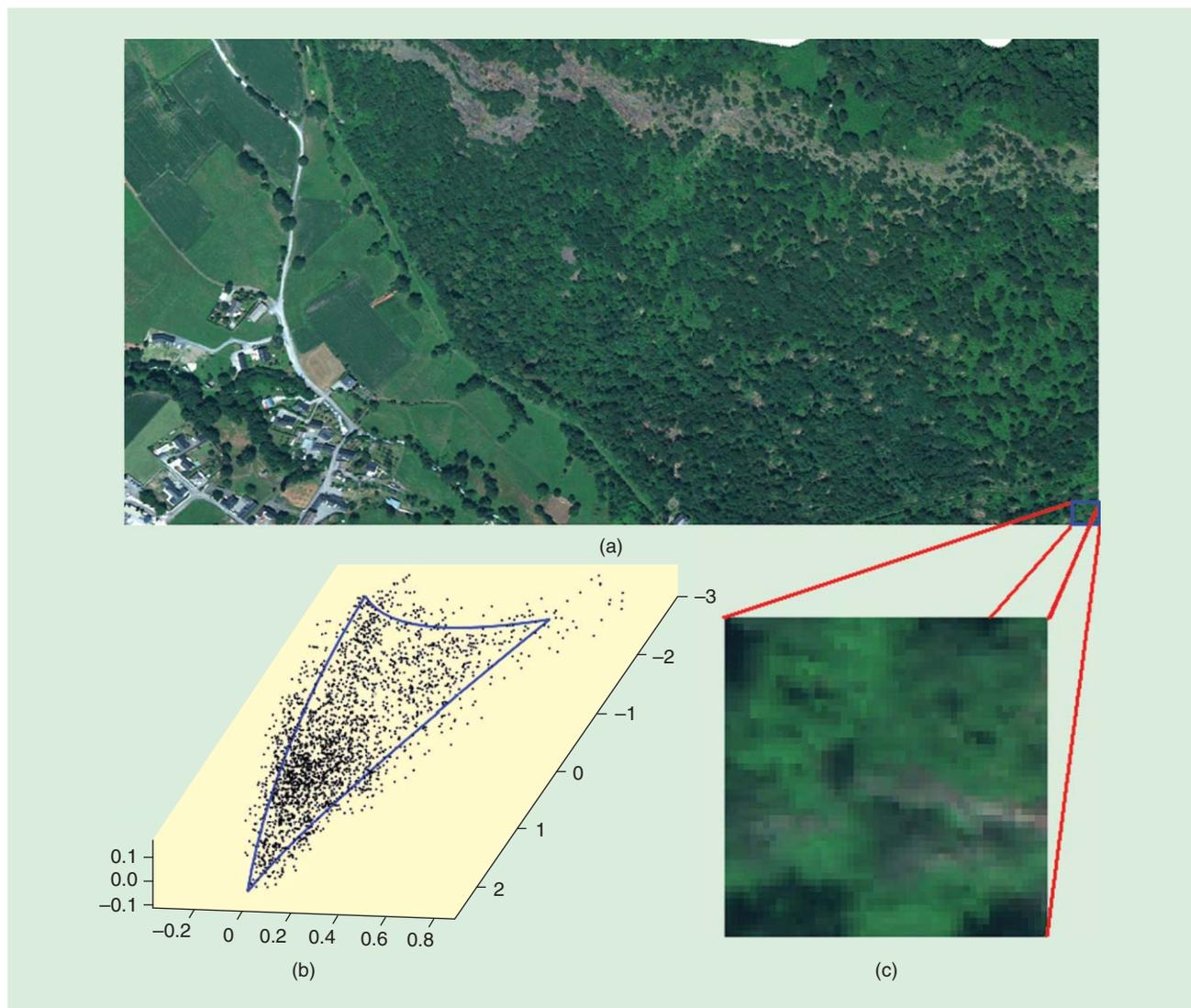
Due to the relative simplicity of the model and the straightforward interpretation of the analysis results, LMM-based unmixing strategies predominate in the literature. All of these techniques have been shown to be very useful whenever the LMM represents a good approximation to the actual mixing. There are, however, practical situations in which the LMM is not a suitable approximation [1]. As an illustrative example, consider a real hyperspectral image, composed of $L = 160$ spectral bands from the visible to near infrared, acquired in 2010 by the airborne Hypspec hyperspectral sensor over Villelongue, France. This image, with a spatial resolution of 0.5 m, is represented in Figure 1(a). From primary inspection and prior knowledge coming from available ground truth, the 50×50 pixel region of interest depicted in Figure 1(c) is known to be composed of mainly $R = 3$ macroscopic components (oak tree, chestnut tree, and an additional nonplanted-tree component). When considering the LMM to model the interactions between

these $R = 3$ components, all the observed pixels should lie in a two-dimensional linear subspace, that can be easily identified by a standard principal component analysis (PCA). Conversely, if nonlinear effects are present in the considered scene, the observed data may belong to a two-dimensional nonlinear manifold. In that case, more complex nonlinear dimension reduction procedures need to be considered to accurately represent the data. The accuracy of these dimension reduction procedures in representing the data set into a two-dimensional linear or nonlinear subspace can be evaluated thanks to the average reconstruction error (ARE), defined as

$$\text{ARE} = \sqrt{\frac{1}{LP} \sum_{p=1}^P \|\mathbf{y}_p - \hat{\mathbf{y}}_p\|^2}, \quad (3)$$

where \mathbf{y}_n are the observed pixels and $\hat{\mathbf{y}}_n$ the corresponding estimates, and $P = 2,500$ is the number of pixels. Here we contrast two approaches, a locally linear Gaussian process latent variable model (LL-GPLVM) introduced in [4] and PCA. When using PCA to represent the data, the obtained ARE is 8.4×10^{-3} while using the LL-GPLVM, the ARE is reduced to 7.9×10^{-3} . This demonstrates that the investigated data set should be preferably represented in a nonlinear subspace, as clearly demonstrated in Figure 1(b), where the nonlinear simplex identified by the fully constrained LL-GPLVM has been represented as blue lines. For the studied hyperspectral image, the nonlinearity seems to be weak, which is often the case for most real applications. However, the models and algorithms presented in this article are also interested for more severe nonlinearities, for instance encountered when analyzing mineral data set.

Consequently, more complex mixing models need to be considered to cope with nonlinear interactions. These models are expected to capture important nonlinear effects that are inherent characteristics of hyperspectral images in several applications. They have proven essential to unveil meaningful information for the geoscience community [5]–[10]. Several approximations to the RT theory have been proposed, such as Hapke's bidirectional model [3]. Unfortunately, these models require highly nonlinear and integral formulations that hinder practical implementations of unmixing techniques. To overcome these difficulties, several physics-based approximations of Hapke's model have been proposed, mainly in the spectroscopy literature (e.g., see [3]). However, despite their wide interest, these approximations still remain difficult to apply for automated hyperspectral imaging. In particular, for such models, there is no unsupervised nonlinear unmixing algorithm able to jointly extract the endmembers from the data and estimate their relative proportions in the pixels. Meanwhile, several approximate but exploitable non-LMMs have been recently proposed in the remote sensing and image processing literatures. Some of them are similarly motivated by physical arguments, such as the class of bilinear models introduced later. Others exploit a more flexible nonlinear mathematical model to improve unmixing performance. Developing effective unmixing algorithms based on non-LMMs represents a challenge for the signal and image processing community.



[FIG1] (a) Real hyperspectral Madonna data acquired by the Hypspec hyperspectral scanner over Villelongue, France. (b) The representation of the $P = 2,500$ pixels (black dots) of the data and boundaries of the estimated nonlinear simplex (blue lines). (c) The region of interest shown in true colors.

Supervised and unsupervised algorithms need to be designed to cope with nonlinear transformations that can be partially or totally unknown. Solving the nonlinear unmixing problem requires innovative approaches to existing signal processing techniques.

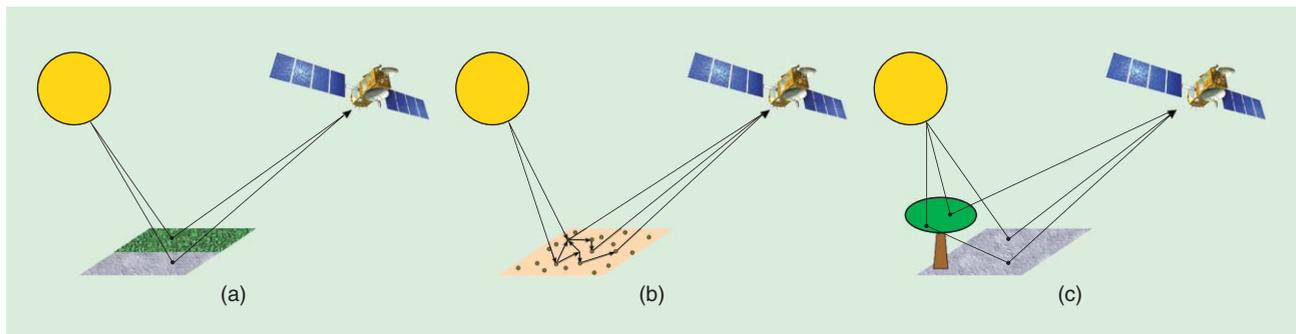
More than ten years after Keshava and Mustard's comprehensive review article on spectral unmixing [11], this article provides an updated review focusing on nonlinear unmixing techniques introduced in the past decade. In [11], the problem on nonlinear mixtures was thoroughly addressed but, at that time, very few algorithmic solutions were available. Capitalizing on almost one decade of advances in solving the linear unmixing problem, scientists from the signal and image processing communities have developed, and continue to do so, automated tools to extract endmembers from nonlinear mixtures, and to quantify their proportions in nonlinearly mixed pixels.

NONLINEAR MODELS

In [1], it is explained that linear mixtures are reasonable when two assumptions are wholly fulfilled. First the mixing process must occur at a macroscopic scale [12]. Second, the photons that reach the sensor must interact with only one material, as is the case in checkerboard type scenes [13]. An illustration of this model is depicted in Figure 2(a) for a scene composed of two materials. When one of these two assumptions does not hold, different nonlinear effects may occur. Two families of nonlinear models are described in what follows.

INTIMATE MIXTURES

The first assumption for linear mixtures is a macroscopic mixing scale. However, there are common situations when interactions occur at a microscopic level. The spatial scales involved are typically smaller than the path length followed by



[FIG2] (a) The LMM: the imaged pixel is composed of two materials. (b) Intimate mixture: the imaged pixel is composed of a microscopic mixture of several constituents. (c) Bilinear model: the imaged pixel is composed of two endmembers: tree and soil. In addition to the individual contribution of each material, bilinear interactions between the tree and the soil reach the sensor.

the photons. The materials are said to be intimately mixed [3]. Such mixtures have been observed and studied for some time, e.g., for imaged scenes composed of sand or mineral mixtures [14]. They have been advocated for analyzing mixtures observed in laboratory [15]. Based on RT theory, several theoretical frameworks have been derived to accurately describe the interactions suffered by the light when encountering a surface composed of particles.

An illustration of these interactions is represented in Figure 2(b). Probably the most popular approaches dealing with intimate mixtures are those of Hapke in [3] since they involve meaningful and interpretable quantities that have physical significance. Based on these concepts, several simplified non-LMMs have been proposed to relate the measurements to some physical characteristics of the endmembers and to their corresponding abundances (that are associated with the relative mass fractions for intimate mixtures). In [2], the author derives an analytical model to express the measured reflectances as a function of parameters intrinsic to the mixtures, e.g., the mass fractions, the characteristics of the individual particles (density, size) and the single-scattering albedo. Other popular approximating models include the discrete-dipole approximation [16] and the Shkuratov's model [17] (interested readers are invited to consult [3] or the more signal processing-oriented papers [18] and [19]). However these models also strongly depend on parameters inherent to the experiment since it requires the perfect knowledge of the geometric positioning of the sensor with respect to the observed sample. This dependency upon external parameters makes the inversion (i.e., the estimation of the mass fractions from the collected spectra) very difficult to implement and, obviously, even more challenging in a unsupervised scenario, i.e., when the spectral signatures of the materials are unknown and need to be also recovered.

More generally, it is worth noting that the first requirement of having a macroscopic mixing scale is intrinsically related to the definition of the endmembers. Indeed, defining a pure material requires specification of the spatial or spectral resolution, which is application dependent. Consider a simple scene composed of three materials A , B , and C . It is natural to expect retrieval of these components individually when analyzing the scene. However, in other circumstances, one may be interested

in the material components themselves, for instance, A_1, A_2, B_1, B_2, C_1 , and C_2 if we assume that each material is composed of two constituents. In that case, pairs of subcomponents combine and, by performing unmixing, one might also be interested in recovering each of these six components. Conversely, it may be well known that the material A can never be present in the observed scene without the material B . In such a case, unmixing would consist of identifying the couple $A + B$ and C , without distinguishing the subcomponent A from the subcomponent B . This issue is frequently encountered in automated spectral unmixing. To circumvent this difficulty in defining the mixture scale, it makes sense to associate pure components with individual instances whose resolutions have the same order of magnitude than the sensor resolution. For example, a patch of sand of spatially homogeneous composition can be considered as a unique pure component. In that case, most of the interactions occurring in most of the scenes of interest can be reasonably assumed to occur at a macroscopic level, at least when analyzing airborne and spaceborne remotely sensed images.

BILINEAR MODELS

Another type of nonlinear interaction occurs at a macroscopic scale, in particular in so-called multilayered configurations. One may encounter this nonlinear model when the light scattered by a given material reflects off other materials before reaching the sensor. This is often the case for scenes acquired over forested areas, where there may be many interactions between the ground and the canopy. An archetypal example of this kind of scene is shown in Figure 2(c).

Several models have been proposed to analytically describe these interactions. They consist of including powers of products of reflectance. However they are usually employed such that interactions of orders greater than two are neglected. The resulting models are known as the family of the bi-LMMs. Mathematically, for most of these bilinear models, the observed spectrum $y_p \in \mathbb{R}^L$ in L spectral bands for the i th pixel is approximated by the following expansion:

$$y_p = \sum_{r=1}^R a_{r,p} \mathbf{m}_r + \sum_{i=1}^{R-1} \sum_{j=i+1}^R \beta_{i,j,p} \mathbf{m}_i \odot \mathbf{m}_j + \mathbf{n}_p, \quad (4)$$

where \odot stands for the termwise (Hadamard) product

$$\mathbf{m}_i \odot \mathbf{m}_j = \begin{pmatrix} m_{1,i} \\ \dots \\ m_{L,i} \end{pmatrix} \odot \begin{pmatrix} m_{1,j} \\ \dots \\ m_{L,j} \end{pmatrix} = \begin{pmatrix} m_{1,i}m_{1,j} \\ \dots \\ m_{L,i}m_{L,j} \end{pmatrix}.$$

In the right-hand side of (4), the first term, also found in (1), summarizes the linear contribution in the mixture while the second term models nonlinear interactions between the materials. The coefficient $\beta_{i,j,p}$ adjusts the amount of nonlinearities between the components \mathbf{m}_i and \mathbf{m}_j in the p th pixel. Several alternatives for imposing constraints on these nonlinear coefficients have been suggested. Similarly to [10], Nascimento and Dias assume in [20] that the (linear) abundance and nonlinearity coefficients obey

$$\begin{cases} a_{r,p} \geq 0, & \forall p, \forall r \\ \beta_{i,j,p} \geq 0, & \forall p, \forall i \neq j \\ \sum_{r=1}^R a_{r,p} + \sum_{i=1}^{R-1} \sum_{j=i+1}^R \beta_{i,j,p} = 1. \end{cases} \quad (5)$$

It is worth noting that, from (5), this Nascimento model (NM), also used in [21], can be interpreted as an LMM with additional virtual endmembers. Indeed, considering $\mathbf{m}_i \odot \mathbf{m}_j$ as a pure component spectral signature with corresponding abundance $\beta_{i,j,p}$, the model in (5) can be rewritten

$$\mathbf{y}_p = \sum_{s=1}^{\tilde{R}} \tilde{a}_{s,p} \tilde{\mathbf{m}}_s + \mathbf{n}_p$$

with the positivity and additivity constraints in (2) where

$$\begin{cases} \tilde{a}_{s,p} \triangleq a_{r,p}, & \tilde{\mathbf{m}}_s \triangleq \mathbf{m}_r & s = 1, \dots, R \\ \tilde{a}_{s,p} \triangleq \beta_{i,j,p}, & \tilde{\mathbf{m}}_s \triangleq \mathbf{m}_i \odot \mathbf{m}_j & s = R+1, \dots, \tilde{R} \end{cases}$$

and $\tilde{R} = (1/2)R(R+1)$. This NM reduces to the LMM when $\tilde{a}_{s,p} = 0$ for $s = R+1, \dots, \tilde{R}$.

Conversely, in [9], Fan and his coauthors have fixed the nonlinearity coefficients as functions of the (linear) abundance coefficients themselves: $\beta_{i,j,p} = a_{i,p}a_{j,p}$ ($i \neq j$). The resulting model, called the Fan model (FM) in what follows, is fully described by the mixing equation

$$\mathbf{y}_p = \sum_{r=1}^R a_{r,p} \mathbf{m}_r + \sum_{i=1}^{R-1} \sum_{j=i+1}^R a_{i,p} a_{j,p} \mathbf{m}_i \odot \mathbf{m}_j + \mathbf{n}_p \quad (6)$$

subject to the constraints in (2). One argument to explain the direct relation between the abundances and the nonlinearity coefficients is the following: if the i th endmember is absent in the p th pixel, then $a_{i,p} = 0$ and there are no interactions between \mathbf{m}_i and the other materials \mathbf{m}_j ($j \neq i$). More generally, it is quite natural to assume that the quantity of nonlinear interactions in a given pixel between two materials is directly related to the quantity of each material present in that pixel. However, it is clear that this model does not generalize the LMM, which can be a restrictive property.

More recently, to alleviate this issue, the generalized bilinear model (GBM) has been proposed in [22] by setting $\beta_{i,j,p} = \gamma_{i,j,p} a_{i,p} a_{j,p}$

$$\mathbf{y}_p = \sum_{r=1}^R a_{r,p} \mathbf{m}_r + \sum_{i=1}^{R-1} \sum_{j=i+1}^R \gamma_{i,j,p} a_{i,p} a_{j,p} \mathbf{m}_i \odot \mathbf{m}_j + \mathbf{n}_p, \quad (7)$$

where the interaction coefficient $\gamma_{i,j,p} \in (0,1)$ quantifies the nonlinear interaction between the spectral components \mathbf{m}_i and \mathbf{m}_j . This model has the same interesting characteristic as the FM: the amount of nonlinear interactions is governed by the presence of the endmembers that linearly interact. In particular, if an endmember is absent in a pixel, there is no nonlinear interaction supporting this endmember. However, it also has the significant advantage of generalizing both the LMM when $\gamma_{i,j,p} = 0$ and the FM when $\gamma_{i,j,p} = 1$. Having $\gamma_{i,j,p} > 0$ indicate that only constructive interactions are considered.

For illustration, synthetic mixtures of $R = 3$ spectral components have been randomly generated according to the LMM, NM, FM, and GBM. The resulting data set are represented in the space spanned by the three principal eigenvectors (associated with the three largest eigenvalues of the sample covariance matrix of the data) identified by a principal component analysis in Figure 3. These plots illustrate an interesting property for the considered data set: the spectral signatures of the pure components are still extremal points, i.e., vertices of the clusters, in the cases of FM and GBM mixtures contrary to the NM. In other words, geometrical endmember extraction algorithms (EEAs) and, in particular, those that are looking for the simplex of largest volume (see [23] for details), may still be valid for the FM and the GBM under the assumption of weak nonlinear interactions.

All these bilinear models only include between-component interactions $\mathbf{m}_i \odot \mathbf{m}_j$ with $i \neq j$ but no within-component interactions $\mathbf{m}_i \odot \mathbf{m}_i$. Finally, in [24], the authors derived a non-LMM using a RT model applied to a simple canyonlike urban scene. Successive approximations and simplifying assumptions lead to the following linear-quadratic mixing model (LQM)

$$\mathbf{y}_p = \sum_{r=1}^R a_{r,p} \mathbf{m}_r + \sum_{i=1}^R \sum_{j=i}^R \beta_{i,j,p} \mathbf{m}_i \odot \mathbf{m}_j + \mathbf{n}_p \quad (8)$$

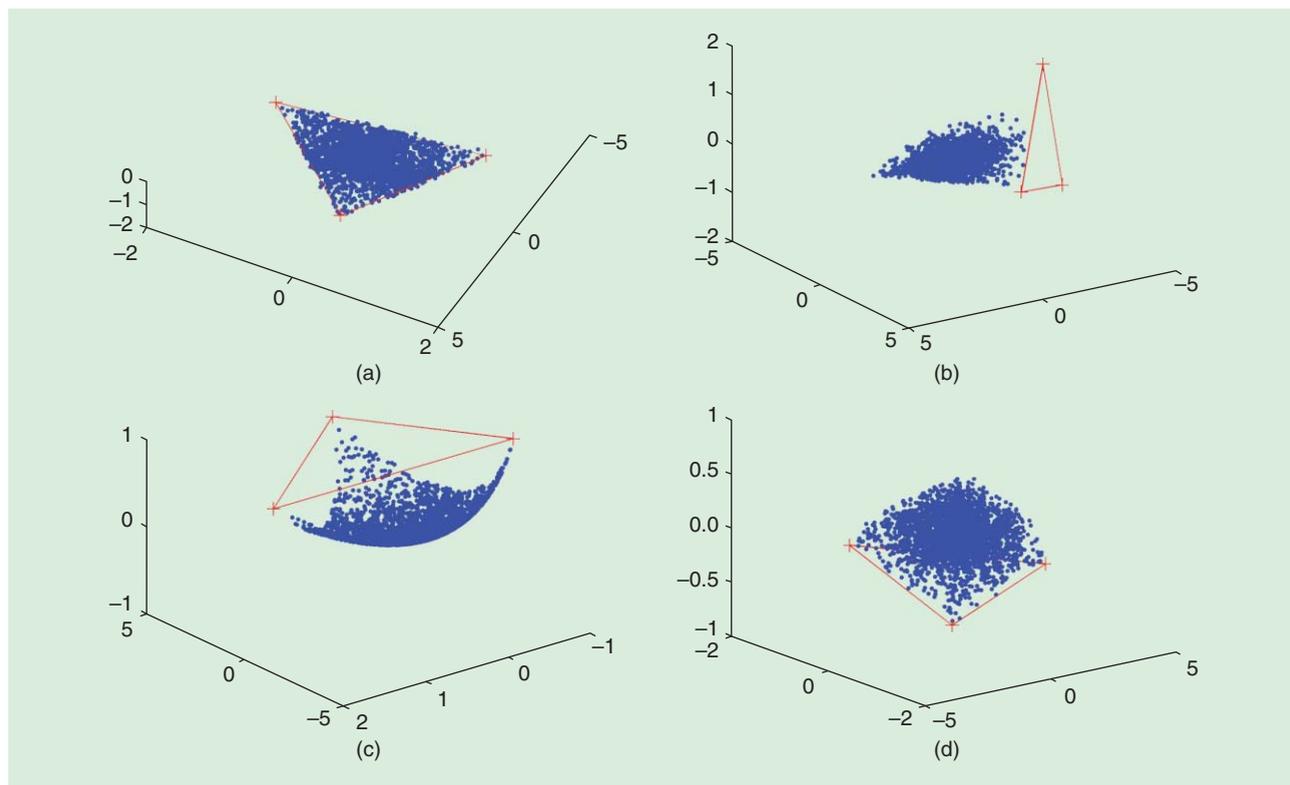
with the positivity and additivity constraints in (2) and $\beta_{i,j,p} \in (0,1)$. This model is similar to the general formulation of the bilinear models in (4), with the noticeable difference that the nonlinear contribution includes quadratic terms $\mathbf{m}_i \odot \mathbf{m}_i$. This contribution also shows that it is quite legitimate to include the termwise products $\mathbf{m}_i \odot \mathbf{m}_j$ as additional components of the standard linear contribution, which is the core of the bilinear models described in this section.

OTHER APPROXIMATING PHYSICS-BASED MODELS

To describe both macroscopic and microscopic mixtures, [25] introduces a dual model composed of two terms

$$\mathbf{y}_p = \sum_{r=1}^R a_{r,p} \mathbf{m}_r + a_{R+1,p} \mathcal{R} \left(\sum_{r=1}^R f_{r,p} \mathbf{w}_r \right) + \mathbf{n}_p.$$

The first term is similar to the one encountered in LMM and comes from the macroscopic mixing process. The second one, considered as an additional endmember with abundance $a_{R+1,p}$, describes the intimate mixture by the average single-scattering albedo [2] expressed in the reflective domain by the mapping $\mathcal{R}(\cdot)$.



[FIG3] Clusters of observations generated according to (a) the LMM, (b) NM, (c) FM, and (d) GBM (blue) and the corresponding endmembers (red crosses).

Altmann et al. have proposed in [26] an approximating model able to describe a wide class of nonlinearities. This model is obtained by performing a second-order expansion of the nonlinearity defining the mixture. More precisely, the p th observed pixel spectrum is defined as a nonlinear transformation $\mathbf{g}_p(\cdot)$ of a linear mixture of the endmember spectra

$$\mathbf{y}_p = \mathbf{g}_p\left(\sum_{r=1}^R a_{r,p} \mathbf{m}_r\right) + \mathbf{n}_p, \quad (9)$$

where the nonlinear function \mathbf{g}_p is defined as a second-order polynomial nonlinearity parameterized by the unique nonlinearity parameter b_p

$$\mathbf{g}_p: (0, 1)^L \rightarrow \mathbb{R}^L$$

$$\mathbf{x} \mapsto [x_1 + b_p x_1^2, \dots, x_L + b_p x_L^2]^T. \quad (10)$$

This model can be rewritten

$$\mathbf{y}_p = \mathbf{M}\mathbf{a}_p + b_p(\mathbf{M}\mathbf{a}_p) \odot (\mathbf{M}\mathbf{a}_p) + \mathbf{n}_p,$$

where $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_R]$ and $\mathbf{a}_p = [a_{1,p}, \dots, a_{R,p}]^T$. The parameter b_p tunes the amount of nonlinearity present in the p th pixel of the image and this model reduces to the standard LMM when $b_p = 0$. It can be easily shown that this polynomial postnonlinear model (PPNM) includes bilinear terms $\mathbf{m}_i \odot \mathbf{m}_j$ ($i \neq j$) similar to those defining the FM, NM and GBM, as well as quadratic terms $\mathbf{m}_i \odot \mathbf{m}_i$ similar to the LQM in (8). This PPNM has been shown to be sufficiently flexible to describe most of the bilinear models introduced in this section [26].

LIMITATION OF A PIXEL-WISE NONLINEAR SU

Having reviewed the above physics-based models, an important remark must be made. It is important to note that these models do not take into account spatial interactions from materials present in the neighborhood of the targeted pixel. It means that these bilinear models only consider scattering effects in a given pixel induced by components that are present in this specific pixel. This is a strong simplifying assumption that allows the model parameters (abundance and nonlinear coefficients) to be estimated pixel-by-pixel in the inversion step. Note, however, that the problem of taking adjacency effects into account, i.e., nonlinear interactions coming from spectral interference caused by atmospheric scattering, has been addressed in an unmixing context in [27].

NONLINEAR UNMIXING ALGORITHMS

Significant promising approaches have been proposed to nonlinearly unmix hyperspectral data. A wide class of nonlinear unmixing algorithms rely explicitly on a nonlinear physics-based parametric model, as detailed earlier. Others do not require definition of the mixing model and rely on very mild assumptions regarding the nonlinearities. For these two classes of approaches, unmixing algorithms have been considered under two different scenarios, namely supervised or unsupervised, depending on the available prior knowledge on the endmembers. When the endmembers are known, supervised algorithms reduce to estimating the abundance coefficients in a single supervised inversion step. In this case, the pure spectral signatures present in the scene must have been previously identified. For instance, they use prior

ON THE USE OF GEOMETRICAL LMM-BASED EEAs TO IDENTIFY NONLINEARLY MIXED ENDMEMBERS

The first automated spectral unmixing algorithms, proposed in the 1990s, were based on geometrical concepts and were designed to identify endmembers as pure pixels (see [1] and [23] for comprehensive reviews of geometrical linear unmixing methods). It is worth noting that this class of algorithms does not explicitly rely on the assumption of pixels coming

from linear mixtures. They only search for endmembers as extremal points in the hyperspectral data set. Provided there are pure pixels in the analyzed image, this might indicate that some of these geometrical approaches can be still valid for nonlinear mixtures that preserve this property, such as the GBM and the FM as illustrated in Figure 3.

information or suboptimal linear EEA. Indeed, as previously noted, when considering weakly nonlinearly mixed data, the LMM-based EEA may produce good endmember estimates when there are pure pixels in the data set (see “On the Use of Geometrical LMM-Based EEAs to Identify Nonlinearly Mixed Endmembers”). In contrast, an unsupervised unmixing algorithm jointly estimates the endmembers and the abundances. Thus the unmixing problem becomes even more challenging, since a blind source separation problem must be solved.

MODEL-BASED PARAMETRIC NONLINEAR UNMIXING ALGORITHMS

Given a nonlinear parametric model, SU can be formulated as a constrained nonlinear regression or a nonlinear source separation problem, depending on whether the endmember spectral signatures are known or not. When dealing with intimate mixtures, some authors have proposed converting the measured reflectance into a single scattering albedo average; since this obeys a linear mixture, the mass fractions associated with each endmember can be estimated using a standard linear unmixing algorithm. This is the approach adopted in [15] and [18] for known and unknown endmembers, respectively. To avoid the functional inversion of the reflectance measurements into the single scattering albedo, a common approach is to use neural networks (NNs) to learn this nonlinear function. This is the strategy followed by Guilfoyle et al. in [28], for which several improvements have been proposed in [29] to reduce the computationally intensive learning step. In these NN-based approaches, the endmembers are assumed to be known a priori, and are required to train the NN. Other NN-based algorithms have been studied in [30]–[33].

For the bilinear models introduced previously, supervised nonlinear optimization methods have been developed based on the assumption that the endmember matrix \mathbf{M} is known. When the observed pixel spectrum y_p is related to the parameters of interest θ_p (a vector containing the abundance coefficients as well as any other nonlinearity parameters) through the function $\varphi(\mathbf{M}, \cdot)$, unmixing the pixel y_p consists of solving the following minimization problem:

$$\hat{\theta}_p = \underset{\theta}{\operatorname{argmin}} \|y_p - \varphi(\mathbf{M}, \theta)\|_2^2. \quad (11)$$

This problem raises two major issues: 1) the nonlinearity of the criterion resulting from the underlying nonlinear model $\varphi(\cdot)$ and 2) the constraints that have to be satisfied by the parameter

vector θ . Since the NM can be interpreted as an LMM with additional virtual endmembers, estimation of the parameters can be conducted with a linear optimization method as in [20]. In [9], [34] dedicated to FM and GBM, the authors propose to linearize the objective criterion via a first-order Taylor series expansion of $\varphi(\cdot)$. Then, the fully constrained least square (FCLS) algorithm of [35] can be used to estimate parameter vector θ . An alternative algorithmic scheme proposed in [34] consists of resorting to a gradient descent method, where the step-size parameter is adjusted by a constrained line search procedure enforcing the constraints inherent to the mixing model. Another strategy initially introduced in [22] for the GBM is based on Monte Carlo approximations, developed in a fully Bayesian statistical framework. The Bayesian setting has the great advantage of providing a convenient way to include the parameter constraints within the estimation problem, by defining appropriate priors for the parameters. This strategy has been also considered to unmix the PPNM [26].

When the spectral signatures \mathbf{M} involved in these bilinear models need also to be identified in addition to the abundances and nonlinearity parameters, more ambitious unmixing algorithms need to be designed. In [36], the authors differentiate the NM to implement updating rules that generalize the sparsity promoting iterated constrained endmember (SPICE) algorithm introduced in [37] for the LMM. Conversely, NMF-based iterative algorithms have been advocated in [38] for the GBM defined in (7), and in [24] for the LQM described in (8). More recently, an unsupervised version of the Bayesian PPNM-based unmixing algorithm initially introduced in [26] has been investigated in [39].

Adopting a geometrical point of view, Heylen and Scheunders propose in [40] an integral formulation to compute geodesic distances on the nonlinear manifold induced by the GBM. The underlying idea is to derive an EEA that identifies the simplex of maximum volume contained in the manifold defined by the GBM-mixed pixels.

MODEL-FREE NONLINEAR UNMIXING ALGORITHMS

When the nonlinearity defining the mixing is unknown, the SU problem becomes even more challenging. In such cases, when the endmember matrix \mathbf{M} is fixed a priori, a classification approach can be adopted to estimate the abundance coefficients, which can be solved using support vector machines [41], [42]. Conversely, when the endmember signatures are not known, a geometrical-based unmixing technique can be used, based on

graph-based approximate geodesic distances [43] or manifold learning techniques [44], [45]. Another promising approach is to use nonparametric methods based on reproducing kernels [46]–[51] or on Gaussian processes [4] to approximate the unknown nonlinearity. These two later techniques are described below.

Nonlinear algorithms operating in reproducing kernel Hilbert spaces (RKHS) have received considerable interest in the machine-learning community, and have proved their efficiency in solving nonlinear problems. Kernel-based methods have been widely considered for detection and classification in hyperspectral images. Surprisingly, nonlinear unmixing approaches operating in RKHS have been investigated in a less in-depth way. The algorithms derived in [46] and [47] were mainly obtained by replacing each inner product between endmember spectra in the cost functions to be optimized by a kernel function. This can be viewed as a nonlinear distortion map applied to the spectral signature of each material, independently of their interactions. This principle can be extremely efficient in solving detection and classification problems as a proper distortion can increase the detectability or separability of some patterns. It is, however, of little physical interest in solving the unmixing problem because the nonlinear nature of the mixtures is not only governed by individual spectral distortions, but also by nonlinear interactions between the materials. In [48], a new kernel-based paradigm was proposed to take the nonlinear interactions of the endmembers into account, when these endmembers are assumed to be a priori known. It solves the optimization problem

$$\min_{\psi_{\theta} \in \mathcal{H}} \sum_{\ell=1}^L [y_{\ell,p} - \psi_{\theta}(\mathbf{m}_{\lambda_{\ell}})]^2 + \mu \|\psi_{\theta}\|_{\mathcal{H}}^2, \quad (12)$$

where $\mathbf{m}_{\lambda_{\ell}}$ is the vector of the endmember signatures at the ℓ th frequency band, particularly, $\mathbf{m}_{\lambda_{\ell}} = [m_{\ell,1}, \dots, m_{\ell,R}]^T$, with \mathcal{H} a given functional space, and μ a positive parameter that controls the tradeoff between regularity of the function $\psi_{\theta}(\cdot)$ and fitting. Again, θ is a vector containing the abundance coefficients as well as any other nonlinearity parameters. It is interesting to note that (12) is the functional counterpart to (11), where $\psi_{\theta}(\cdot)$ defines the nonlinear interactions between the endmembers assumed to be known in [48]. Clearly, this strategy may fail if the functional space \mathcal{H} is not chosen appropriately. A successful strategy is to define \mathcal{H} as an RKHS to exploit the so-called kernel trick. Let $\kappa(\cdot, \cdot)$ be the reproducing kernel of \mathcal{H} . The RKHS \mathcal{H} must be carefully selected via its kernel to make it flexible enough to capture wide classes of nonlinear relationships, and to reliably interpret a variety of experimental measurements. To extract the mixing ratios of the endmembers, the authors in [48] focus their attention on partially linear models, resulting in the so-called K-HYPE SU algorithm. More precisely, the function $\psi_{\theta}(\cdot)$ in (12) is defined by an LMM parameterized by the abundance vector \mathbf{a} , combined with a nonparametric term,

$$\psi_{\theta}(\mathbf{m}_{\lambda_{\ell}}) = \mathbf{a}^T \mathbf{m}_{\lambda_{\ell}} + \psi_{\text{nlm}}(\mathbf{m}_{\lambda_{\ell}}) \quad (13)$$

possibly subject to the constraints in (2), where ψ_{nlm} can be any real-valued function of an RKHS denoted by \mathcal{H}_{nlm} . This

model generalizes the standard LMM, and mimics the PPNM when \mathcal{H}_{nlm} is defined to be the space of polynomial functions of degree two. Remember that the latter is induced by the polynomial kernel $\kappa(\mathbf{m}_{\lambda_{\ell}}, \mathbf{m}_{\lambda_{\ell}}) = (\mathbf{m}_{\lambda_{\ell}}^T \mathbf{m}_{\lambda_{\ell}})^q$ of degree $q = 2$. More complex interaction mechanisms can be considered by simply changing $\kappa(\mathbf{m}_{\lambda_{\ell}}, \mathbf{m}_{\lambda_{\ell}})$. By virtue of the reproducing kernel machinery, the problem can still be solved in the framework of (12).

Another strategy introduced in [4] considers a kernel-based approach for unsupervised nonlinear SU based on a nonlinear dimensionality reduction using a Gaussian process latent variable model (GPLVM). In this work, the authors have used a particular form of kernel that extends the generalized bilinear model in (7). The algorithm proposed in [4] is unsupervised in the sense that the endmembers contained in the image and the mixing model are not known. Only the number of endmembers is assumed to be known. As a consequence, the parameters to be estimated are the kernel parameters, the endmember spectra and the abundances for all image pixels. The main advantage of GPLVMs is their capacity to accurately model many different nonlinearities. GPLVMs construct a smooth mapping from the space of fractional abundances to the space of observed mixed pixels that preserves dissimilarities. This strategy has been also considered in [51] by Nguyen et al., who solve the so-called pre-image problem [52] studied in the machine-learning community. In the SU context, it means that pixels that are spectrally different have different latent variables and thus different abundance vectors. However, preserving local distances is also interesting: spectrally close pixels are expected to have similar abundance vectors and thus similar latent variables. Several approaches have been proposed to preserve similarities, including back-constraints and locally linear embedding.

For illustration, a small set of experiments has been conducted to evaluate some of the model-based and model-free algorithms introduced above. First, four synthetic images of size 50×50 have been generated by mixing $R = 3$ endmember spectra (i.e., green grass, olive green paint, and galvanized steel metal) extracted from the spectral libraries provided with the ENVI software [53]. These four images have been generated according to the standard LMM (1), GBM (7), FM (6), and PPNM (9), respectively. For each image, the abundance coefficient vectors $\mathbf{a}_p \triangleq [a_{1,p}, \dots, a_{3,p}]$ ($p = 1, \dots, 2,500$) have been randomly and uniformly generated in the admissible set defined by the constraints (2). We have also considered the more challenging scenario defined by the assumption that there is no pure pixel (by imposing $a_{r,p} < 0.9, \forall r, \forall p$). The nonlinearity coefficients are uniformly drawn in the set $[0, 1]$ for the GBM. The PPNM-parameters $b_p, p = 1 \dots, P$ have been generated uniformly in the set $[-0.3, 0.3]$. For both scenario (i.e., with or without pure pixels), all images have been corrupted by an additive independent and identically distributed (i.i.d) Gaussian noise of variance $\sigma^2 = 10^{-4}$, which corresponds to an average signal-to-noise ratio (SNR) of 20 dB (note that the usual SNR for most of the spectro-imagers are not below 30 dB). Various unmixing strategies have been implemented to recover

the endmember signatures and then estimate the abundance coefficients. For supervised unmixing, the N-FINDR algorithm [54] and its nonlinear geodesic-based counterpart [43] have been used to extract the endmembers from linear and nonlinear mixtures, respectively. Then, dedicated model-based strategies were used to recover the abundance fractions. The fully constrained least square (FCLS) algorithm [35] was used for linear mixtures. Gradient-based algorithms (GBAs) were used for nonlinear mixtures. The GBAs are detailed in [55], [34], and [9] for the PPNM, GBM, and FM, respectively. For comparison with supervised unmixing, and to evaluate the impact of having no pure pixels in these images, joint estimations of endmembers and abundances was implemented using the Markov chain Monte Carlo techniques detailed in [56] and [39] for the LMM and PPNM images, respectively. Finally, the model-free supervised K-HYPE algorithm detailed in [48] was also coupled with the nonlinear EEA in [43]. The performance of these unmixing strategies has been evaluated in term of abundance estimation error measured by

$$\text{RMSE} = \sqrt{\frac{1}{RP} \sum_{n=1}^N \| \mathbf{a}_p - \hat{\mathbf{a}}_p \|^2},$$

where \mathbf{a}_p is the n th actual abundance vector and $\hat{\mathbf{a}}_p$ its corresponding estimate. The results are reported in Table 1. These results clearly show that the prior knowledge of the actual mixing model underlying the observations is a clear advantage for abundance estimation. However, in the absence of such knowledge, using an inappropriate model-based algorithm may lead to poor unmixing results. In such cases, as advocated before, PPNM seems to be sufficiently flexible to provide reasonable estimates, whatever the mixing model may be. Otherwise, one may prefer to resort to model-free-based strategy such as K-HYPE.

DETECTING NONLINEAR MIXTURES

The consideration of nonlinear effects in hyperspectral images can provide more accurate results in terms of endmember and abundance identification. However, working with nonlinear models generally requires a higher computational complexity than approaches based on the LMM. Thus, unmixing linearly mixed pixels using nonlinear models should be avoided. Consequently, it

is of interest to devise techniques to detect nonlinearities in the mixing process before applying any unmixing method. Linearly mixed pixels can then be unmixed using linear unmixing techniques, leaving the application of more involved nonlinear unmixing methods to situations where they are really necessary. This section describes approaches that have been recently proposed to detect nonlinear mixing in hyperspectral images.

DETECTION USING A PPNM

One interesting approach for nonlinearity detection is to assume a parametric non-LMM that can model different nonlinearities between the endmembers and the observations. A model that has been successfully applied to this end is the PPNM (9) studied in [26] and [55]. PPNM assumes the postnonlinear mixing described in (9) with the polynomial nonlinearity g_p defined in (10). Hence, the nonlinearity is characterized by the parameter b_p for each pixel in the scene. This parameter can be estimated in conjunction with the abundance vector \mathbf{a}_p and the noise variance σ^2 . Denote as $s^2(\mathbf{a}_p, b_p, \sigma^2)$ the variance of the maximum likelihood estimator \hat{b}_p of b . Using the properties of the maximum likelihood estimator, it makes sense to approximate the distribution of \hat{b}_p by the following Gaussian distribution:

$$\hat{b}_p \sim \mathcal{N}(b_p, s^2(\mathbf{a}_p, b_p, \sigma^2)).$$

The nonlinearity detection problem can be formulated as the binary hypothesis testing problem

$$\begin{cases} \mathcal{H}_0 : y_p \text{ is distributed according to the LMM (1)} \\ \mathcal{H}_1 : y_p \text{ is distributed according to the PPNM (9)}. \end{cases} \quad (14)$$

Hypothesis \mathcal{H}_0 is characterized by $b_p = 0$, whereas nonlinear models (\mathcal{H}_1) correspond to $b_p \neq 0$. Then, (14) can be rewritten as

$$\begin{cases} \mathcal{H}_0 : \hat{b}_p \sim \mathcal{N}(0, s_0^2) \\ \mathcal{H}_1 : \hat{b}_p \sim \mathcal{N}(b_p, s_1^2), \end{cases} \quad (15)$$

where $s_0^2 = s^2(\mathbf{a}_p, 0, \sigma^2)$ and $s_1^2 = s^2(\mathbf{a}_p, b_p, \sigma^2)$ with $b_p \neq 0$. Detection can be performed using the generalized likelihood ratio test. This test accepts \mathcal{H}_1 (respectively \mathcal{H}_0) if the ratio $T \triangleq \hat{b}_p^2 / s_0^2$ is greater (respectively lower) than a threshold η . As shown in [55], the statistic T is approximately normally distributed under the two hypotheses. Consequently, the threshold η

[TABLE 1] ABUNDANCE RMSEs ($\times 10^{-2}$) FOR VARIOUS LINEAR/NONLINEAR UNMIXING SCENARIOS.

		MIXING MODELS: WITH PURE PIXELS				MIXING MODELS: WITHOUT PURE PIXELS				
		LMM	PPNM	GBM	FM	LMM	PPNM	GBM	FM	
MODEL-BASED ALGORITHM	LMM	N-FINDR + FCLS	1.42	14.1	7.71	13.4	3.78	13.2	6.83	9.53
		UNSUPERVISED MCMC	0.64	12.4	5.71	8.14	0.66	10.9	4.21	3.92
	PPNM	GEODESIC + GBA	1.52	10.3	6.04	12.1	4.18	6.04	4.13	3.74
		UNSUPERVISED MCMC	0.39	0.73	1.32	2.14	0.37	0.81	1.38	2.25
	GBM	GEODESIC + GBA	2.78	14.3	6.01	13.0	4.18	11.1	5.02	1.45
	FM	GEODESIC + GBA	13.4	21.8	9.90	3.40	12.2	18.1	7.17	4.97
	GEODESIC + K-HYPE	2.43	9.71	5.23	11.3	2.44	5.92	3.18	2.58	

can be explicitly related to the probability of false alarm (PFA) and the probability of detection (PD), i.e., the power of the test. However, this detection strategy assumes the prior knowledge of the variances s_0^2 and s_1^2 . In practical applications, Altmann et al. have proposed to modify the previous test strategy as follows [55]:

$$\hat{T} = \frac{\hat{b}_p^2}{\hat{s}_0^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta^*, \quad (16)$$

where \hat{s}_0^2 can be calculated as

$$\hat{s}_0^2 = \text{CCRLB}(0; \hat{\mathbf{a}}_p, \hat{\sigma}^2). \quad (17)$$

In (17), CCRLB is the constrained Cramér–Rao lower-bound [57] on estimates of the parameter vector $\boldsymbol{\theta} = [\mathbf{a}_p^T, b_p, \sigma^2]^T$ under \mathcal{H}_0 , and $(\hat{\mathbf{a}}_p, \hat{\sigma}^2)$ is the MLE of (\mathbf{a}_p, σ^2) . The performance of the resulting test is illustrated in Figure 4, which shows the pixels detected as linear (red crosses) and nonlinear (blue dots) when generated according to various mixing models (LMM, FM, GBM, and PPNM).

ROBUST MODEL-FREE DETECTION

The detector discussed in the previous section assumes a specific non-LMM under the alternative hypothesis. However, there are situations where the actual mixing does not obey any available model. It is also possible that there is insufficient information to opt for any existing nonlinearity model. In these cases, it is interesting to address the problem of determining whether an

observed pixel is a linear function of endmembers or results from a generic nonlinear mixing.

One may consider the LMM (1) and the hyperplane \mathcal{P} defined by

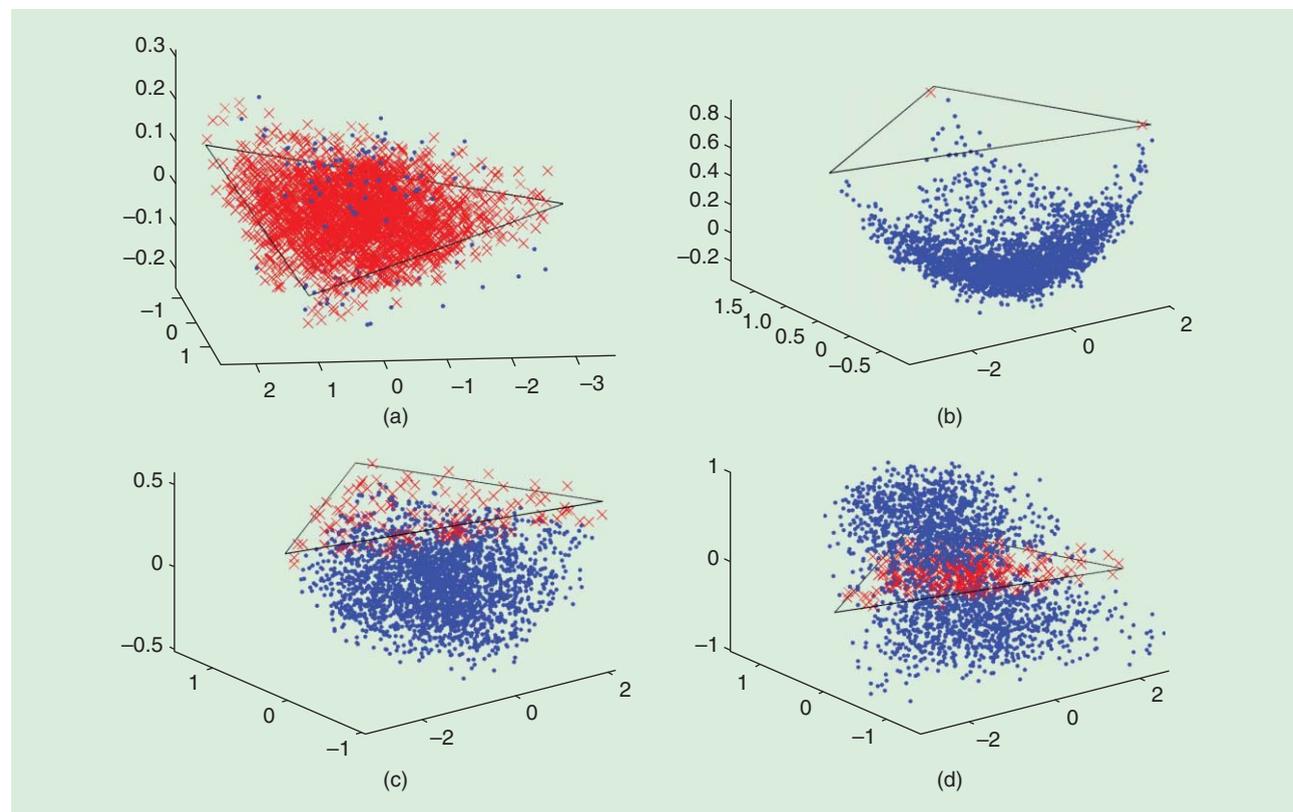
$$\mathcal{P} : \left\{ \mathbf{z}_p \mid \mathbf{z}_p = \mathbf{M}\mathbf{a}_p, \sum_{r=1}^R a_{r,p} = 1 \right\}. \quad (18)$$

In the noise-free case, the hyperplane \mathcal{P} lies in an $(R - 1)$ -dimensional subspace embedding all observations distributed according to the LMM. On the other hand, consider the general non-LMM

$$\mathbf{y}_p = \mathbf{M}\mathbf{a}_p + \boldsymbol{\mu}_p + \mathbf{n}_p, \quad (19)$$

where $\boldsymbol{\mu}_p$ is an $L \times 1$ deterministic vector that does not belong to \mathcal{P} , i.e., $\boldsymbol{\mu}_p \notin \mathcal{P}$ and \mathbf{a}_p satisfies the constraints (2). Note that a similar non-LMM coupled with a group-sparse constraint on $\boldsymbol{\mu}_p$ has been explicitly adopted in [58] and [59] to make more robust the unmixing of hyperspectral pixels. In (19), $\boldsymbol{\mu}_p$ can be a nonlinear function of the endmember matrix \mathbf{M} and/or the abundance vector \mathbf{a}_p and should be denoted as $\boldsymbol{\mu}_p(\mathbf{M}, \mathbf{a}_p)$ [60]. However, the arguments \mathbf{M} and \mathbf{a}_p are omitted here for brevity. Given an observation vector \mathbf{y}_p , the detection of nonlinear mixtures can be formulated as the following binary hypothesis testing problem:

$$\begin{cases} \mathcal{H}_0 : \mathbf{y}_p \text{ is distributed according to the LMM (1)} \\ \mathcal{H}_1 : \mathbf{y}_p \text{ is distributed according to the model (19).} \end{cases}$$



[FIG4] Pixels detected as linear (red crosses) and nonlinear (blue dotted) for the four subimages generated according to the (a) LMM, (b) FM, (c) GBM, and (d) PPNM. Black lines depict the simplex corresponding to the noise-free LMM.

Using the statistical properties of the noise \mathbf{n}_p , we obtain $E[y_p | \mathcal{H}_0] = \mathbf{M}\mathbf{a}_p \in \mathcal{P}$, whereas $E[y_p | \mathcal{H}_1] = \mathbf{M}\mathbf{a}_p + \boldsymbol{\mu}_p \notin \mathcal{P}$. As a consequence, it makes sense to consider the squared Euclidean distance

$$\delta^2(y_p) = \min_{z_p \in \mathcal{P}} \|y_p - z_p\|^2 \quad (20)$$

between the observed pixel y_p and the hyperplane \mathcal{P} to decide which hypothesis (\mathcal{H}_0 or \mathcal{H}_1) is true.

As shown in [60], the test statistic $\delta^2(y_p)$ is distributed according to χ^2 distribution under the two hypotheses \mathcal{H}_0 and \mathcal{H}_1 . The parameters of this distribution depend on the known matrix \mathbf{M} , the noise variance σ^2 and the nonlinearity vector $\boldsymbol{\mu}_p$. If σ^2 is known, the distribution of $\delta^2(y_p)$ is perfectly known under \mathcal{H}_0 and partially known under \mathcal{H}_1 . In this case, one may employ a statistical test that does not depend on $\boldsymbol{\mu}_p$. This test accepts \mathcal{H}_1 (respectively \mathcal{H}_0) if the ratio $T \triangleq \delta^2(y_p)/\sigma^2$ is greater (respectively lower) than a threshold η . As in the PPNM-based detection procedure, the threshold η can be related to the PFA and PD through closed-form expressions. In particular, it is interesting to note that the PD is intrinsically related to a non-Euclidean norm of the residual component $\boldsymbol{\mu}_p$ (see [60, Eq. (11)]), which is unfortunately unknown in most practical applications. If the noise variance σ^2 is unknown, which is the case in most practical applications, one can replace σ^2 with an estimate $\hat{\sigma}^2$, leading to

$$T^* \triangleq \frac{\delta^2(y_p)}{\hat{\sigma}^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \quad (21)$$

where η is the threshold computed as previously indicated. The PFA and PD of the test (21) are then explicitly obtained using cumulative distribution functions of the χ^2 distribution. It was shown in [60] that the better the estimation of σ^2 , the closer the distributions of T and T^* and thus the closer the performances of the two corresponding tests. Several techniques can be used to estimate σ^2 . For instance, $\hat{\sigma}^2$ has been estimated in [60] through an eigenanalysis of the sample covariance matrix of a set of pixels assumed to share the same variance. The value of $\hat{\sigma}^2$ was determined as the average of the smallest eigenvalues of the sample covariance matrix. The accuracy of the estimator is a function of the number of eigenvalues considered. It was shown in [60] that a PFA smaller (respectively larger) than P_{FA}^* is obtained if $\hat{\sigma}^2 > \sigma^2$ (respectively $\hat{\sigma}^2 < \sigma^2$).

CONCLUSIONS AND OPEN CHALLENGES

To overcome the intrinsic limitations of the LMM, several recent contributions have been made for modeling of the physical processes that underlie hyperspectral observations. Some models attempt to account for between-material interactions affecting photons before they reach the spectro-imager. Based on these models, several parametric algorithms have been proposed to solve the resulting nonlinear unmixing problem. Another class of unmixing algorithms attempts to avoid the use of any rigid nonlinear model by using nonparametric

machine-learning-inspired techniques. The price to pay for handling nonlinear interactions induced by multiple scattering effects or intimate mixtures is the computational complexity and a possible degradation of unmixing performance when processing large hyperspectral images. To overcome these difficulties, one possible strategy consists of detecting pixels subjected to nonlinear mixtures in a preprocessing step. The pixels detected as linearly mixed can then benefit from the huge and reliable literature dedicated to the linear unmixing problem. The remaining pixels (detected as nonlinear) can then be the subject of particular attention.

This article has described development methods in nonlinear mixing for hyperspectral imaging. Several important challenges remain. First of all, better integration of algorithmic approaches and physical models have the potential to greatly improve nonlinear unmixing performance. By fully accounting for complex RT effects, such as scattering, dispersion, and beam interaction depth, a physical model can guide the choice of simplified mathematical and statistical models. Preliminary results have been recently communicated in [61], based on in situ measurements coupled with simulation tools (e.g., ray-tracing techniques). A second challenge is to develop unmixing models that take heterogeneity of the medium into account. Heterogeneous regions consist of combinations of linear, weakly nonlinear, and strongly nonlinear pixels. The detection strategies detailed above might be one solution to tackle this problem since they are able to locate the areas where a nonlinear model may outperform a linear model and vice versa. Another approach adopted in [58] and [59], which works well when there are only a few nonlinear subregions, consists of using a statistical outlier approach to identify the nonlinear pixels. Moreover, as any nonlinear blind source separation problem, deriving flexible unsupervised unmixing algorithms is still a major challenge, especially if one wants to go one step further than a crude pixel-by-pixel analysis by exploiting spatial information inherent to these images. Finally, we observe that the presence of nonlinearity in the observed spectra is closely related to the number R of endmembers, which is usually unknown. For example, in analogy to kernelization in machine learning, after nonlinear transformation, a nonlinear mixture of R components can often be represented as a linear mixture of \hat{R} endmembers, with $\hat{R} > R$. Recent advances in manifold learning and dimensionality estimation are promising approaches to the nonlinear unmixing problem.

ACKNOWLEDGMENTS

Part of this work has been funded by the Hypanema ANR Project n°ANR-12-BS03-003 and the MADONNA project supported by INP Toulouse, France. Some results obtained in this article result from fruitful discussions during the CIMI Workshop on Optimization and Statistics in Image Processing (Toulouse, France, 24–28 June 2013). The authors are grateful to Yoann Altmann and Abderrahim Halimi (University of Toulouse) for sharing their numerous MATLAB codes and for stimulating discussions.

AUTHORS

Nicolas Dobigeon (Nicolas.Dobigeon@enseeiht.fr) received the M.Sc. and Ph.D. degrees in signal processing from the National Polytechnic Institute of Toulouse in 2004 and 2007, respectively. In 2008, he joined the University of Toulouse (INP-ENSEEIH), where he is currently an associate professor. He conducts his research within the Signal and Communications Group of the IRIT Laboratory, and he is also an affiliated faculty member of the Telecommunications for Space and Aeronautics Cooperative Laboratory. His recent research activities have been focused on statistical signal and image processing, with a particular interest in Bayesian inverse problems for applications to remote sensing, biomedical imaging, and genomics.

Jean-Yves Tourneret (Jean-Yves.Tourneret@enseeiht.fr) received the Ph.D. degree from the National Polytechnic Institute of Toulouse in 1992. He is currently a professor with the University of Toulouse. His research activities center around statistical signal and image processing with a particular interest to Bayesian and Markov chain Monte Carlo methods. He has been involved in the organization of several conferences including EUSIPCO'02 (program chair), ICASSP'06 (plenaries), SSP'12 (international liaisons), and SSP'14 (special sessions). He has been a member of the Signal Processing Theory and Methods Committee of the IEEE Signal Processing Society (2001–2007, 2010–present). He was an associate editor of *IEEE Transactions on Signal Processing* (2008–2011).

Cédric Richard (Cedric.Richard@unice.fr) is a professor at the University of Nice Sophia-Antipolis, France, and a junior member of the Institut Universitaire de France. His current research interests include statistical signal processing and machine learning. He was the general chair of the 2011 IEEE Statistical Signal Processing Workshop held in Nice, France. From 2006 to 2010, he was an associate editor of *IEEE Transactions on Signal Processing*. Currently, he is an associate editor of Elsevier's *Signal Processing* and of *IEEE Signal Processing Letters*. He is a member of the IEEE Signal Processing Society as well as the Signal Processing Theory and Methods Technical Committee and the Machine Learning for Signal Processing Technical Committee.

José Carlos M. Bermudez (j.bermudez@ieee.org) received the Ph.D. degree in electrical engineering from Concordia University, Canada, in 1985. In 1985, he joined the Federal University of Santa Catarina, Florianópolis, Brazil, where he is currently a professor of electrical engineering. His recent research interests are in adaptive filtering and statistical signal processing and their applications. He was an associate editor of *IEEE Transactions on Signal Processing* (1994–1996, 1999–2001) and of the *EURASIP Journal of Advances on Signal Processing* (2006–2010). He was a member of the Signal Processing Theory and Methods Technical Committee of the IEEE Signal Processing Society (1998–2004).

Stephen McLaughlin (S.McLaughlin@hw.ac.uk) received his Ph.D. degree from the University of Edinburgh, Scotland, in

1990. From 1991 until 2001, he held a Royal Society University Research Fellowship to study nonlinear signal processing techniques at Edinburgh University, where he was awarded a Chair in Electronic Communication Systems in 2002. In 2011, he joined Heriot-Watt University as a professor of signal processing and head of the School of Engineering and Physical Sciences. He is a Fellow of the IEEE, the Royal Academy of Engineering, and the Royal Society of Edinburgh.

Alfred O. Hero (hero@umich.edu) received the Ph.D. degree from Princeton University (1984) in electrical engineering. Since 1984, he has been with the University of Michigan, Ann Arbor, where he is the R. Jamison and Betty Williams Professor of Engineering. His recent research interests are in statistical signal processing, machine learning, and the analysis of high-dimensional spatiotemporal data. He received the IEEE Signal Processing Society Meritorious Service Award (1998) and the IEEE Third Millennium Medal (2000). He was president of the IEEE Signal Processing Society (2006–2008) and was on the Board of Directors of the IEEE (2009–2011), where he was the director of Division IX (Signals and Applications).

REFERENCES

- [1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [2] B. W. Hapke, "Bidirectional reflectance spectroscopy: I. Theory," *J. Geophys. Res.*, vol. 86, no. B4, pp. 3039–3054, Apr. 1981.
- [3] B. Hapke, *Theory of Reflectance and Emission Spectroscopy*. Cambridge, U.K.: Cambridge Univ. Press, 1993.
- [4] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Nonlinear spectral unmixing of hyperspectral images using Gaussian processes," *IEEE Trans. Signal Processing*, vol. 61, no. 10, pp. 2442–2453, May 2013.
- [5] T. Ray and B. Murray, "Nonlinear spectral mixing in desert vegetation," *Remote Sensing Environ.*, vol. 55, pp. 59–64, 1996.
- [6] J.-P. Combe, P. Launeau, V. Carrère, D. Despan, V. Mède, L. Barillé, and C. Sotin, "Mapping microphytobenthos biomass by non-linear inversion of visible-infrared hyperspectral images," *Remote Sensing Environ.*, vol. 98, no. 4, pp. 371–387, 2005.
- [7] W. Liu and E. Y. Wu, "Comparison of non-linear mixture models: Sub-pixel classification," *Remote Sensing Environ.*, vol. 18, no. 2, pp. 145–154, Jan. 2005.
- [8] K. Arai, "Nonlinear mixture model of mixed pixels in remote sensing satellite images based on Monte Carlo simulation," *Adv. Space Res.*, vol. 41, no. 11, pp. 1715–1723, 2008.
- [9] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Int. J. Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, June 2009.
- [10] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. V. der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sensing Environ.*, vol. 113, pp. 1183–1193, Feb. 2009.
- [11] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [12] R. B. Singer and T. B. McCord, "Mars: Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance," in *Proc. 10th Lunar and Planetary Science Conf.*, Mar. 1979, pp. 1835–1848.
- [13] R. N. Clark and T. L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," *J. Geophys. Res.*, vol. 89, no. 7, pp. 6329–6340, 1984.
- [14] D. B. Nash and J. E. Conel, "Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite, and ilmenite," *J. Geophys. Res.*, vol. 79, no. 11, pp. 1615–1621, Apr. 1974.
- [15] J. F. Mustard and C. M. Pieters, "Photometric phase functions of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra," *J. Geophys. Res.*, vol. 94, no. B10, pp. 13619–13634, Oct. 1989.
- [16] B. T. Draine, "The discrete-dipole approximation and its application to interstellar graphite grains," *Astrophys. J.*, vol. 333, pp. 848–872, Oct. 1988.

- [17] Y. Shkuratov, L. Starukhina, H. Hoffmann, and G. Arnold, "A model of spectral albedo of particulate surfaces: Implication to optical properties of the moon," *Icarus*, vol. 137, no. 2, pp. 235–246, 1999.
- [18] J. M. P. Nascimento and J. M. Bioucas-Dias, "Unmixing hyperspectral intimate mixtures," in *Proc. SPIE Image and Signal Processing Remote Sensing XVI*, vol. 74830, L. Bruzzone, Ed. SPIE: Bellingham, WA, Oct. 2010, p. 78300C.
- [19] R. Close, P. Gader, A. Zare, J. Wilson, and D. Dranishnikov, "Endmember extraction using the physics-based multi-mixture pixel model," in *Proc. SPIE Imaging Spectrometry XVII*, vol. 8515, S. S. Shen and P. E. Lewis, Eds. SPIE: San Diego, CA, Aug. 2012, pp. 85 150L–1–85 150L–14.
- [20] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image and Signal Processing Remote Sensing XV*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds. SPIE: Bellingham, WA, vol. 7477, no. 1, p. 747701, 2009.
- [21] N. Raksuntorn and Q. Du, "Nonlinear spectral mixture analysis for hyperspectral imagery in an unknown environment," *IEEE Geosci. Remote Sensing Lett.*, vol. 7, no. 4, pp. 836–840, 2010.
- [22] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.
- [23] W.-K. Ma, J. M. Bioucas-Dias, P. Gader, T.-H. Chan, N. Gillis, A. Plaza, A. Ambikapathi, and C.-Y. Chi, "A signal processing perspective on hyperspectral unmixing," *IEEE Signal Processing Mag.*, vol. 31, no. 1, pp. 67–81, 2013.
- [24] I. Meganem, P. Déliot, X. Briottet, Y. Deville, and S. Hosseini, "Linear-quadratic mixing model for reflectances in urban environments," *IEEE Trans. Geosci. Remote Sensing*, to be published.
- [25] R. Close, P. Gader, J. Wilson, and A. Zare, "Using physics-based macroscopic and microscopic mixture models for hyperspectral pixel unmixing," in *Proc. SPIE Algorithms and Technologies Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, vol. 8390, S. S. Shen and P. E. Lewis, Eds. SPIE: Baltimore, MA, Apr. 2012, pp. 83 901L–83 901L–13.
- [26] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a post-nonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Processing*, vol. 21, no. 6, pp. 3017–3025, June 2012.
- [27] D. Burazerovic, R. Heylen, B. Geens, S. Sterckx, and P. Scheunders, "Detecting the adjacency effect in hyperspectral imagery with spectral unmixing techniques," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1070–1078, June 2013.
- [28] K. J. Guilfoyle, M. L. Althouse, and C.-I. Chang, "A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, no. 8, pp. 2314–2318, Aug. 2001.
- [29] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using radial basis functions and orthogonal least squares," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, Vancouver, Canada, July 2011, pp. 1151–1154.
- [30] J. Plaza, P. Martinez, R. Perez, and A. Plaza, "Nonlinear neural network mixture models for fractional abundance estimation in AVIRIS hyperspectral images," in *Proc. XIII NASA/Jet Propulsion Laboratory Airborne Earth Science Workshop*, Pasadena, CA, Mar. 2004, pp. 1–12.
- [31] J. Plaza, A. Plaza, R. Perez, and P. Martinez, "Automated generation of semi-labeled training samples for nonlinear neural network-based abundance estimation in hyperspectral data," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, 2005, pp. 345–350.
- [32] J. Plaza, A. Plaza, R. Pérez, and P. Martinez, "Joint linear/nonlinear spectral unmixing of hyperspectral image data," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, 2007, pp. 4037–4040.
- [33] J. Plaza, A. Plaza, R. Perez, and P. Martinez, "On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images," *Pattern Recognit.*, vol. 42, no. 11, pp. 3032–3045, Nov. 2009.
- [34] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unmixing hyperspectral images using the generalized bilinear model," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, Vancouver, Canada, July 2011, pp. 1886–1889.
- [35] D. C. Heinz and C.-I. Chang, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geoscience and Remote Sensing*, vol. 29, no. 3, pp. 529–545, Mar. 2001.
- [36] P. Gader, D. Dranishnikov, A. Zare, and J. Chanussot, "A sparsity promoting bilinear unmixing model," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Processing: Evolution Remote Sensing (WHISPERS)*, Shanghai, China, June 2012.
- [37] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection in hyperspectral imagery," *IEEE Geosci. Remote Sensing Lett.*, vol. 4, no. 3, pp. 446–450, July 2007.
- [38] N. Yokoya, J. Chanussot, and A. Iwasaki, "Generalized bilinear model based nonlinear unmixing using semi-nonnegative matrix factorization," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, Munich, Germany, 2012, pp. 1365–1368.
- [39] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a Hamiltonian Monte Carlo algorithm," submitted for publication.
- [40] R. Heylen and P. Scheunders, "Calculation of geodesic distances in nonlinear mixing models: Application to the generalized bilinear model," *IEEE Geosci. Remote Sensing Lett.*, vol. 9, no. 4, pp. 644–648, July 2012.
- [41] J. Plaza, A. Plaza, P. Martinez, and R. Perez, "Nonlinear mixture models for analyzing laboratory simulated-forest hyperspectral data," in *Proc. SPIE Image and Signal Processing Remote Sensing IX*, vol. 5238, pp. 480–487, 2004.
- [42] P.-X. Li, B. Wu, and L. Zhang, "Abundance estimation from hyperspectral image based on probabilistic outputs of multi-class support vector machines," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, 2005, pp. 4315–4318.
- [43] R. Heylen, D. Burazerovic, and P. Scheunders, "Non-linear spectral unmixing by geodesic simplex volume maximization," *IEEE J. Select. Topics Signal Process.*, vol. 5, no. 3, pp. 534–542, June 2011.
- [44] H. Nguyen, C. Richard, P. Honeine, and C. Theys, "Hyperspectral image unmixing using manifold learning methods. Derivations and comparative tests," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, Munich, Germany, July 2012, pp. 3086–3089.
- [45] G. Licciardi, X. Ceamanos, S. Doute, and J. Chanussot, "Unsupervised nonlinear spectral unmixing by means of NLPCA applied to hyperspectral imagery," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, Munich, Germany, July 2012, pp. 1369–1372.
- [46] J. Broadwater, R. Chellappa, A. Banerjee, and P. Burlina, "Kernel fully constrained least squares abundance estimates," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, Barcelona, Spain, July 2007, pp. 4041–4044.
- [47] J. Broadwater and A. Banerjee, "A comparison of kernel functions for intimate mixture models," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Aug. 2009, pp. 1–4.
- [48] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Processing*, vol. 60, no. 2, pp. 480–492, Jan. 2013.
- [49] J. Chen, C. Richard, and P. Honeine, "Nonlinear estimation of material abundances in hyperspectral images with ℓ_1 -norm spatial regularization," *IEEE Trans. Geosci. Remote Sensing*, to be published.
- [50] X. Li, J. Cui, and L. Zhao, "Blind nonlinear hyperspectral unmixing based on constrained kernel nonnegative matrix factorization," *Signal, Image Video Process.*, pp. 1–13, Oct. 2012.
- [51] N. H. Nguyen, J. Chen, C. Richard, P. Honeine, and C. Theys, "Supervised nonlinear unmixing of hyperspectral images using a pre-image methods," in *New Concepts in Imaging: Optical and Statistical Models*, EAS Publ. Ser., vol. 59, D. Mary, C. Theys, and C. Aime, Eds. Les Ulis, France: EDP Sciences, 2013, pp. 417–437.
- [52] P. Honeine and C. Richard, "Preimage problem in kernel-based machine learning," *IEEE Signal Processing Mag.*, vol. 28, no. 2, pp. 77–88, Mar. 2011.
- [53] RSI, *ENVI User's Guide Version 4.0*. Boulder, CO: RSI, 2003.
- [54] M. Winter, "Fast autonomous spectral end-member determination in hyperspectral data," in *Proc. 13th Int. Conf. Applied Geologic Remote Sensing*, Vancouver, Apr. 1999, vol. 2, pp. 337–344.
- [55] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinearity detection in hyperspectral images using a polynomial post-nonlinear mixing model," *IEEE Trans. Image Processing*, vol. 22, no. 4, pp. 1267–1276, Apr. 2013.
- [56] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.
- [57] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1285–1301, Nov. 1990.
- [58] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Residual component analysis of hyperspectral images—Application to joint nonlinear unmixing and nonlinearity detection," *IEEE Trans. Image Processing*, submitted for publication.
- [59] N. Dobigeon and C. Févotte, "Robust nonnegative matrix factorization for nonlinear unmixing of hyperspectral images," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Gainesville, FL, June 2013.
- [60] Y. Altmann, N. Dobigeon, J.-Y. Tourneret, and J. C. M. Bermudez, "A robust test for nonlinear mixture detection in hyperspectral images," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, June 2013, pp. 2149–2153.
- [61] L. Tits, W. Delabastita, B. Somers, J. Farifteh, and P. Coppin, "First results of quantifying nonlinear mixing effects in heterogeneous forests: A modeling approach," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing (IGARSS)*, 2012, pp. 7185–7188.

[Alina Zare and K.C. Ho]

Endmember Variability in Hyperspectral Analysis

[Addressing spectral variability
during spectral unmixing]

Variable illumination and environmental, atmospheric, and temporal conditions cause the measured spectral signature for a material to vary within hyperspectral imagery. By ignoring these variations, errors are introduced and propagated throughout hyperspectral image analysis. To develop accurate spectral unmixing and endmember estimation methods, a number of approaches that account for spectral variability have been developed. This article motivates and provides a review for methods that account for spectral variability during

hyperspectral unmixing and endmember estimation and a discussion on topics for future work in this area.

INTRODUCTION

A wide range of applications including planetary exploration, environmental monitoring, and target detection have been tackled using hyperspectral image analysis. In all of these applications, a prominent area of study is in the development of spectral unmixing and endmember estimation techniques [1]. Endmembers are the spectral signatures, the radiance or reflectance values over hundreds of contiguous spectral bands, of the pure, constituent materials in a hyperspectral scene. Given a hyperspectral image, endmember estimation is the task of

Digital Object Identifier 10.1109/MSP.2013.2279177

Date of publication: 5 December 2013

extracting the spectral signatures for all of the materials located in the scene. Spectral unmixing estimates the proportions of each endmember at every spatial location. In the overwhelming majority of hyperspectral unmixing and endmember estimation algorithms, pixels in a hyperspectral image are modeled as convex combinations of endmembers

$$\mathbf{x}_n = \sum_{m=1}^M p_{mn} \mathbf{e}_m + \boldsymbol{\epsilon}_n \quad n = 1, \dots, N \quad (1)$$

such that $p_{mn} \geq 0$ and $\sum_{m=1}^M p_{mn} = 1$, and where \mathbf{x}_n is a $D \times 1$ vector containing the spectral signature of the n th pixel in a hyperspectral image, N is the number of pixels in the image, M is the number of endmembers (or materials) found in the scene, p_{mn} is the proportion of endmember m in pixel n , \mathbf{e}_m is the $D \times 1$ vector containing the spectral signature of the m th endmember, $\boldsymbol{\epsilon}_n$ is an error term, and D is the number of spectral bands (i.e., the dimensionality) of the hyperspectral data [1]. Although used extensively in the literature, the linear mixing model in (1) lacks the ability to represent the spectral variability of the endmembers in a scene. Instead, endmember spectral signatures are represented as single points in a high-dimensional space.

The spectral signature for a material varies within hyperspectral data collections due to a number of reasons including environmental, atmospheric, and temporal factors. A material may also have intrinsic spectral variability. One major source of spectral variability results from variation due to illumination conditions [2], [3]. Changes in illumination can result from variation in topography and surface roughness leading to varying levels of shadowed and brightly lit regions. Illumination is also dependent on solar elevation, solar azimuth, and local incidence angle on the material of interest [2]. Differences in the architecture of plant canopies, changes in the distribution of leaf orientation in vegetated regions, or varying building structure and layout in urban areas cause differing illumination levels and areas of shade [4]. When considering minerals, changes in grain size and texture affect illumination. In this case, smaller grain sizes allow for stronger and more uniform backscattered energy resulting in shallower absorption features and higher reflectance values [5]. If accurate digital terrain-elevation models and photometric information are known for an area, then some of the effects of illumination may be able to be removed. In general, however, this information is unavailable for a scene and, if it were known, would require significant computationally intensive preprocessing.

Another significant source of spectral variation results from changing atmospheric conditions. In particular, the

levels of atmospheric gases and aerosols such as water vapor, oxygen, ozone, carbon monoxide, and carbon dioxide can absorb and scatter radiation and, as an example, cause a significant impact on the measured spectral signature of a material [6]. Many atmospheric gases have strong absorption features or scattering characteristics in a number of wavelengths throughout the electromagnetic spectrum. These absorption and scattering features modify measured spectral signatures as they affect the downward and upward transmittance of radiation from the sun to the surface being measured and, then, from the ground surface to the hyperspectral sensor. In a common spectral radiance model, the downward transmittance of radiation to surface is a combination of the direct solar radiation from the sun to the ground surface; the skylight also known as the *diffuse transmittance*, which is the solar radiation scattered by atmospheric gases and aerosols and redirected toward the

ground surface being measured; and light due to multiple, repeated reflections and scattering from neighboring surfaces and the atmosphere. Then, the total upward transmittance is a combination of the light reflected by the ground surface; light reflected by the ground surface, and rescattered by atmospheric gases and aerosols, and sunlight scattered by the atmosphere and redirected toward the sensor

without reaching a ground surface [7]. A number of approaches have been developed that attempt to remove the atmospheric effects from hyperspectral data; however, some of the spectral variability due to atmospheric conditions may not be eliminated using these approaches. For example, as discussed by Gao et al. [8], many of these approaches may not accurately account for nitrogen dioxide levels in the atmosphere, which can be extremely high in urban areas, or measured radiance from a pixel may be modified by the radiance of neighboring pixels due to scattering of solar radiation by atmospheric molecules.

Although spectral variability due to these sources is expected, unmixing and endmember estimation methods generally do not account for spectral variability. As such, errors resulting from inaccurate endmember representation will be propagated throughout analysis. The most prominent effects from inaccurate endmember representations are resulting errors in estimated proportion values, termed *proportion indeterminacy*, or the use of too many endmembers to represent a spectrum [4]. To avoid these errors and to represent spectral variability during analysis, a number of spectral unmixing and endmember estimation algorithms that incorporate spectral variability have been developed in the literature [9]. As shown by results presented by Garcia-Haro et al. [4], accounting for endmember variability can result in a significant improvement

METHODS THAT ACCOUNT FOR ENDMEMBER SPECTRAL VARIABILITY CAN BE ORGANIZED INTO TWO GENERAL CATEGORIES BASED ON THE VARIABLE ENDMEMBER REPRESENTATION USED: ENDMEMBERS AS SETS AND ENDMEMBERS AS STATISTICAL DISTRIBUTIONS.

[TABLE 1] AN OVERVIEW OF SOME BENEFITS AND CHALLENGES TO ENDMEMBER VARIABILITY REPRESENTATIONS.

	ENDMEMBERS AS SETS	ENDMEMBERS AS STATISTICAL DISTRIBUTIONS
ADVANTAGES	MAKES USE OF A NONPARAMETRIC REPRESENTATION THAT DOES NOT REQUIRE THE ASSUMPTION OF ANY PARTICULAR DISTRIBUTION ON SPECTRAL VALUES. THE APPROACH CAN BE EASILY CONSTRAINED TO ONLY PHYSICALLY POSSIBLE SPECTRAL SIGNATURES.	MAKES USE OF AN EFFICIENT, COMPACT, AND CONTINUOUS REPRESENTATION THAT CAN ACCOUNT FOR A RANGE OF SPECTRAL VALUES THAT MAY NOT HAVE BEEN PREVIOUSLY MEASURED FOR INCLUSION IN A DISCRETE SPECTRAL LIBRARY.
CHALLENGES	ACCURACY IS LIMITED BY THE ENDMEMBER REPRESENTATIVES THAT HAVE BEEN MEASURED OR EXTRACTED FROM INPUT DATA FOR INCLUSION DURING SPECTRAL UNMIXING.	ACCURACY OF RESULTS IS DEPENDENT ON THE SELECTION OF AN ACCURATE PARAMETRIC FORM AND PARAMETER VALUES. DEPENDING ON THESE SELECTIONS, THIS APPROACH MAY RESULT IN REPRESENTATIONS THAT ALLOW FOR ENDMEMBER VARIATIONS THAT ARE NOT PHYSICALLY MEANINGFUL.
TYPICAL METHODS	<p>MESMA, MELSUM, BSMA, AND AUTOMCU: <i>PROS:</i> DIRECT AND STRAIGHTFORWARD TO APPLY. <i>CONS:</i> COMPUTATIONALLY INEFFICIENT.</p> <p>ENDMEMBER BUNDLES: <i>PROS:</i> DESIGNED TO QUANTIFY PROPORTION INDETERMINANCY. <i>CONS:</i> ACCURACY DEPENDENT ON CHOSEN ENDMEMBER "SEEDS."</p> <p>BAND WEIGHTING AND TRANSFORMATION: <i>PROS:</i> COMPUTATIONALLY EFFICIENT; MAXIMIZED BETWEEN ENDMEMBER VARIANCE TO REDUCE UNMIXING CONFUSION. <i>CONS:</i> NOISE CORRELATION AFFECTED AFTER TRANSFORMATION.</p> <p>SPARSE AND LOCAL UNMIXING: <i>PROS:</i> EXPLICIT SPECTRAL LIBRARY NOT REQUIRED A PRIORI; DATA-POINT OR REGION SPECIFIC ENDMEMBER SETS ALLOWED. <i>CONS:</i> ENFORCING SPARSITY AND SPATIAL ASSUMPTIONS THAT MAY BE INVALID (E.G., EDGES).</p>	<p>BAYESIAN SOURCE SEPARATION: <i>PROS:</i> JOINT ESTIMATION OF ENDMEMBERS AND PROPORTIONS; ENDMEMBERS NOT LIMITED BY SPECTRAL LIBRARY; <i>CONS:</i> FULL SPECIFICATIONS OF ENDMEMBER DISTRIBUTIONS NEEDED; PHYSICALLY UNREALISTIC ENDMEMBERS MAY BE ALLOWED.</p> <p>NORMAL COMPOSITIONAL MODEL: <i>PROS:</i> EFFICIENT USE OF GAUSSIAN DISTRIBUTION. <i>CONS:</i> PHYSICALLY UNREALISTIC ENDMEMBERS INCLUDED; COVARIANCE BETWEEN BANDS NOT ADDRESSED.</p> <p>BETA COMPOSITIONAL MODEL: <i>PROS:</i> REFLECTANCE VALUES CONSTRAINED TO PHYSICALLY MEANINGFUL RANGE; ABLE TO REPRESENT SKEW IN ENDMEMBER DISTRIBUTIONS. <i>CONS:</i> DETAILED SPECIFICATIONS OF ENDMEMBER DISTRIBUTIONS REQUIRED.</p> <p>METHODS OF HIGHER MOMENTS: <i>PROS:</i> SIMPLE FORM TO ADDRESS ANY ENDMEMBER DISTRIBUTION. <i>CONS:</i> ACCURACY LIMITED BY NUMBER OF MOMENTS USED; ENDMEMBER MOMENTS NEEDED IN ADVANCE.</p>

in proportion estimation and signal fitting. Previous work by Somers et al. [9] has reviewed a number of methods in using the endmembers as sets approach to account for spectral variability. This article extends the review for the latest developments of this approach and includes the other approach of endmembers as statistical distributions that was not discussed in [9]. The objective here is to provide a comprehensive overview of the two approaches with an algorithmic and signal processing viewpoint.

Methods that account for endmember spectral variability can be organized into two general categories based on the variable endmember representation used: endmembers as sets (which can also be seen as a linear mixing model approach) and endmembers as statistical distributions (which can be seen as a stochastic mixing model approach). The former has a longer history and the latter is more recent. Both of these representations come with associated advantages and challenges; some of these are listed in Table 1. In the following, a review of these methods is provided.

ENDMEMBERS AS SETS

One approach to account for spectral variability is to represent each endmember of a material with a set or "bundle" of spectra. Generally speaking, with this representation, a pixel will be modeled as a convex combination of one or more representatives selected from different sets of spectral signatures.

KNOWN SPECTRAL LIBRARY

MULTIPLE ENDMEMBER ENDMEMBER MIXTURE ANALYSIS AND VARIANTS

Several methods for spectral unmixing estimate proportion values by exhaustively searching a given spectral library for endmembers whose corresponding estimated proportion values satisfy some criteria. The most prominent of these approaches is the multiple endmember spectral mixture analysis (MESMA) algorithm [10]. MESMA estimates proportions for an input pixel by searching the endmembers for which proportion values are found that satisfy three conditions: 1) the root mean square (rms) error between the input pixel and its reconstruction using endmembers and proportions is below a prescribed threshold, $\text{rms}(\mathbf{x}_n, \{\mathbf{e}_m\}_{m=1}^M) < t$, where

$$\text{rms}(\mathbf{x}_n, \{\mathbf{e}_m\}_{m=1}^M) = \min_{p_{mn}} \left(\sqrt{\frac{1}{D} \left\| \mathbf{x}_n - \sum_{m=1}^M p_{mn} \mathbf{e}_m \right\|_2^2} \right) \quad (2)$$

and $\{\mathbf{e}_m\}_{m=1}^M$ is the collection of M endmembers used to unmix the pixel, 2) the rms error for contiguous spectral bands is below a prescribed threshold, and 3) the estimated proportion values are within some prescribed range. This range may be fixed to $[0 \leq p_{mn} \leq 1]$ or, to account for some error in endmember or proportion values allow for values slightly outside of this range (e.g., $[-0.01 \leq p_{mn} \leq 1.01]$). This approach allows each input pixel from a hyperspectral scene to be represented using a unique

collection of endmember spectra from the spectral library. Therefore, even if each input pixel is restricted to be represented by, e.g., no more than three endmembers, the full scene can be mapped using many more.

When the spectral library is large, there will be a considerable number of elements to choose from. In such a case, MESMA may result in difficulties such as finding an overabundance of viable solutions or extreme computational inefficiency by exhaustively searching over all possible combinations of endmembers. A number of extensions to MESMA attempt to mitigate these difficulties such as defining methods for pruning endmembers before applying MESMA.

There are a number of variants based on MESMA. The multiple-endmember linear spectral unmixing (MELSUM) method relaxes the criteria to identify viable proportions where non-negative values are sufficient [5]. The Bayesian spectral mixture analysis (BSMA) method obtains the final proportion value of a material in a pixel through a weighted sum of the proportion values found in all combinations, where the weights are proportional to the probabilities of each endmember deduced from the spectral library [11]. The automated Monte Carlo unmixing (AutoMCU) method randomly selects endmembers from the spectral library to unmix the hyperspectral scene over several times [12], rather than doing an exhaustive search. The proportion values of each pixel are summarized using their mean and standard deviation associated with each material from the several runs. In addition to improving computational efficiency, AutoMCU is able to quantify explicitly the proportion indeterminacy.

MESMA, MELSUM, BSMA, and AutoMCU do not inherently address spectral variability without an appropriate spectral library. The spectral library must have more than one spectra to represent each material, and they are grouped together to form the set $E_m, m = 1, 2, \dots, M$, where M is the number of materials. Then, during unmixing, one or none of the spectra from each set are selected to estimate proportion values, i.e., $e_m \in E_m$. As such, when an appropriate spectral library is provided, each pixel can be unmixed using a unique spectral variation of each material.

During unmixing, these approaches often also account for variability due to illumination by including a “shade” endmember [3]. The shade effect often corresponds to multiplication factor in the hyperspectral image. In the linear mixing model, however, the shade effect can be accounted for through normalization of the image data or the addition of a photometric shade endmember (i.e., the origin) because they are mathematically equivalent.

ENDMEMBER BUNDLES

The endmember bundles unmixing approach was designed to explicitly address spectral variability and quantify the associated

proportion indeterminacy [13]. It estimates the minimum and maximum proportion values for the m th material of pixel x_n by minimizing (3) and (4), respectively,

$$p_{\min, mn} = \min_{p_{1m}, \dots, p_{Tm}, s} \sum_{i=1}^T \chi_{m,i} p_{in} + ws \quad (3)$$

$$p_{\max, mn} = \max_{p_{1m}, \dots, p_{Tm}, s} 1 - \left[\sum_{i=1}^T (1 - \chi_{m,i}) p_{in} + ws \right], \quad (4)$$

where

$$\chi_{m,i} = \begin{cases} 1 & \text{if } e_i \in E_m; \\ 0 & \text{if } e_i \notin E_m. \end{cases} \quad (5)$$

The minimization is over the nonnegative proportion values p_{in} that must satisfy $x_n = \sum_{i=1}^T p_{in} e_i + sd$ and $\sum_{i=1}^T p_{in} = 1$. T is the total number of elements across all endmember sets, e_i denotes the i th spectral signature across the whole library, E_m is the endmember set composed of all endmember spectra of the m th material, and $d = x_n - (1/T) \sum_{i=1}^T e_i$. The variables $s \in [0, \infty)$ and $w \in [0, \infty)$ are used to minimize the residual error. The parameter w is set to a very large value (tending toward ∞) to drive the estimated s value toward zero. The optimization is conducted using Dikin's affine algorithm method [13]. Given these min and max proportion values,

the mean proportion value can also be computed, $p_{\text{mean}, mn} = (p_{\min, mn} + p_{\max, mn})/2$.

BAND SELECTION, WEIGHTING, AND TRANSFORMATION

Another approach to address spectral variability is through band selection or weighting [9] such that the wavelengths with minimum spectral variability are the ones that are primarily used for spectral unmixing. Extending this concept, one could find spectral transformations that transform the input hyperspectral data into a space that minimizes the effect of spectral variability. In particular, the Fisher discriminant approach (FDA) for spectral unmixing learns a transformation for the spectral signature elements to minimize the scatter within endmember sets and maximize the scatter among them prior to the estimation of proportion values [14] to avoid unmixing confusion [15]. The goal of FDA is to estimate the transformation projection matrix, \hat{W}

$$\hat{W} = \arg \max_w \frac{|W^T S_b W|}{|W^T S_w W|}, \quad (6)$$

where S_b and S_w are between- and within-class dispersion matrices [16]. Each pixel is transformed by \hat{W} before estimating proportion values, and (1) becomes

$$\hat{W} x_n = \sum_{m=1}^M p_{mn} \hat{W} e_m + \hat{W} \epsilon_n \quad n = 1, \dots, N. \quad (7)$$

An effective transformation matrix will make the transformed elements within the same set nearly identical to each

THE ENDMEMBER BUNDLES UNMIXING APPROACH WAS DESIGNED TO EXPLICITLY ADDRESS SPECTRAL VARIABILITY AND QUANTIFY THE ASSOCIATED PROPORTION INDETERMINACY.

other and, thus, any one of the transformed elements from each of the endmember sets can be used to estimate the corresponding proportion values for that material. This linear transformation will affect the noise distribution and, to address this, the noise correlation can be accounted for through the use of a weighting matrix when estimating the proportion values.

SUPPORT VECTOR MACHINE UNMIXING

A number of approaches using support vector machines (SVMs) have been developed for spectral unmixing while addressing spectral variability [17], [18]. An SVM is a commonly used supervised two-class classifier. Given a training set, $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$ where \mathbf{x}_i is the i th data point and $y_i \in \{0, 1\}$ is the desired class label, an SVM learns a hyperplane that separates the two classes [16]. In the context of hyperspectral unmixing, the two SVM classes are the favorable class containing pixels from mixing endmembers at specific proportion values versus the unfavorable class of those coming from other proportion choices. To elaborate further [17], the first step for SVM unmixing is to discretize the solution space of unmixed proportions to a finite number of candidates. For example, if two materials are under consideration, with a proportion resolution of 0.1 for each material, nine proportion solution candidates $[(.9, .1), (.8, .2), \dots, (.1, .9)]$ are generated. For each candidate, synthesized pixel data are created by drawing elements from the endmember sets and mixing them according to the proportion candidate. Next, a number of SVMs are trained, one for each solution candidate, by labeling the synthesized data of the selected candidate as one and the rest as zero. Given an unknown pixel for unmixing, it is evaluated through all of the SVM classifiers and the SVM that gives the largest classification margin will identify the corresponding candidate as the unmixed proportion solution.

One advantage of SVM unmixing is that spectral variability is automatically taken care of when creating synthesized data for SVM training. Due to the discretization of proportion values, however, this approach produces a finite and possibly limited number of abundance proportion choices only, which might be not acceptable in some applications.

UNKNOWN SPECTRAL LIBRARY

Spectral libraries can be obtained using laboratory measurements of materials of interest or manual endmember identification from the imagery under study or data previously acquired. Often, spectral libraries or the expertise needed to identify spectral signatures of various materials is unavailable. Thus, approaches to autonomously estimate endmember sets and perform spectral unmixing of input data are needed.

AUTOMATED ENDMEMBER BUNDLES

An earlier work [13] proposes a semiautomatic endmember set estimation technique. It begins with manually selected “endmember seeds” and grows an endmember set by identifying the data pixels that have high correlation coefficients with the seeds, provided that spectral reflectance values are constrained to the physically meaningful range of zero to one.

In the subsequent work [19], Somers et al. developed a fully automated approach for building endmember sets by repeatedly applying a standard endmember extraction method, such as vertex component analysis (VCA) [20], to a randomly selected portion of the input data. The endmember sets are obtained by grouping all of the estimated endmembers into M clusters using the K -means clustering algorithm. After extracting the endmember sets, any of the previously described spectral unmixing methods can be applied.

SPARSE UNMIXING

The automated endmember bundles (AEB) technique treats endmember set estimation independently from proportion estimation in unmixing. However, many endmember estimation approaches are paired with spectral unmixing steps such as with most sparse unmixing approaches. Various sparsity approaches can also be categorized in terms of the specific sparsity constraint employed such as l_0 -norm constraints or inclusion of l_1 regularization terms. In the sparse unmixing approach for endmember variability presented by Castrodad et al. [21], endmember set estimation is conducted in conjunction with unmixing on a training data set containing many examples of a material. In particular, given N_m pixels composed purely of material m that is denoted by the $D \times N_m$ matrix \mathbf{X}_m , the dictionary elements representing the endmember set of material m , represented by the columns of the matrix $\tilde{\mathbf{E}}_m$, are obtained by minimizing the objective function

$$R(\tilde{\mathbf{E}}_m, \mathbf{P}_m) = \|\mathbf{X}_m - \tilde{\mathbf{E}}_m \mathbf{P}_m\|_F^2 + \lambda_S \|\mathbf{P}_m \mathbf{1}\|_1, \quad (8)$$

where each column of \mathbf{P}_m represents the proportion values contributed from different columns of $\tilde{\mathbf{E}}_m$, $\mathbf{1}$ is a length N_m column vector of unity, and λ_S is a fixed regularization parameter. The N_m pixels in \mathbf{X}_m are given from some training data, such as hand-selected from a ground-truthed data set. The second term in the objective function promotes more sparseness as λ_S increases. The function is minimized through alternating optimization by solving for updated $\tilde{\mathbf{E}}_m$ and \mathbf{P}_m values, iteratively. After minimization of (8), the endmember set of material m is constructed by maintaining the columns in $\tilde{\mathbf{E}}_m$ that are being used to describe the training set, $E_m = \{\tilde{\mathbf{e}}_j : j \ni \sum_{n=1}^{N_m} p_{jn} > 0\}$, where \ni denotes “such that,” p_{jn} is the (j, n) element of \mathbf{P}_m . The minimization is conducted using a Gauss–Seidel type iteration as discussed in [21].

ONE ADVANTAGE OF SVM UNMIXING IS THAT SPECTRAL VARIABILITY IS AUTOMATICALLY TAKEN CARE OF WHEN CREATING SYNTHESIZED DATA FOR SVM TRAINING.

After the dictionary elements are obtained, spectral unmixing can be conducted on previously unseen hyperspectral data \mathbf{x}_n . Let \mathbf{E}_m be the $D \times |E_m|$ matrix whose columns are the elements of the dictionary E_m obtained above and where $|E_m|$ is the number of elements in the m th dictionary. Also, let \mathbf{p}_{mn} be an $|E_m| \times 1$ vector that represents the proportion values of the elements in E_m . The unmixing is achieved through the minimization of the following objective:

$$\left\| \mathbf{x}_n - \sum_{m=1}^M \mathbf{E}_m \mathbf{p}_{mn} \right\|_2^2 + \lambda_S \sum_{m=1}^M \|\mathbf{p}_{mn}\|_1 + \lambda_G \sum_{m=1}^M \left\| \mathbf{p}_{mn} - \sum_{i \in \mathcal{R}(n)} w_{ni} \mathbf{p}_{mi} \right\|_2^2, \quad (9)$$

where $\mathcal{R}(n)$ is the neighborhood defined around the n th data point, $w_{ni} = 1/Z_i \exp\{-\|\mathbf{x}_n - \mathbf{x}_i\|_2^2/\sigma^2\}$, Z_i is a pixel-dependent normalization constant such that $\sum_i w_{ni} = 1$, σ^2 is a fixed constant, and λ_G is a fixed regularization parameter. The second term in the objective function promotes sparseness. The third term can be viewed as a spatial smoothing term in which pixels in the same neighborhood are encouraged to have similar proportion values. Increasing σ^2 enlarges the neighborhood region and larger λ_G enhances more spatial smoothing.

LOCAL UNMIXING

Local unmixing (LU) extracts endmember sets with elements identified in spatial neighborhoods across the input hyperspectral image [22], [23]. In the technique developed by Goenaga et al. [23], an input hyperspectral image is divided into equally sized tiles. Local endmembers are extracted manually or using some endmember extraction method. The extracted local endmembers are clustered using spectral angle criteria to form the endmember sets. Proportion values are obtained by minimizing $\left\| \mathbf{x}_n - \sum_{i=1}^T \mathbf{p}_{in} \mathbf{e}_i \right\|_2^2$ subject to $p_{in} > 0$ and $\sum_i p_{in} \leq 1$. The proportion of a material is computed by the sum of the proportion values associated with its clustered elements, $p_{mn} = \sum_{j \ni \mathbf{e}_j \in E_m} p_{jn}$.

One interesting point of discussion is that LU approaches assume pixels that are spatially close are likely to be composed of the same materials. This assumption could be viewed as a balance between 1) the traditional hyperspectral unmixing algorithms that consider all pixels in a scene to be composed of the same endmember elements and 2) the pixel-to-pixel independent spectral variability approaches such as MESMA, AutoMCU, and MELSUM, where every input pixel in a scene could be composed of a completely unique collection of elements from across the endmember sets.

Sparse unmixing applied to a dictionary consisting of several variants of each endmember can be used to address spectral variability. In terms of LU, we can impose neighboring pixels with

AN ALTERNATIVE TO THE SET-BASED APPROACH TO ADDRESS SPECTRAL VARIABILITY IN A MATERIAL IS THE USE OF A MULTIVARIATE STATISTICAL DISTRIBUTION.

the same dictionary elements in their sparse representation. In the spectral-variability sparse unmixing approach presented above, in addition to the sparsity constraint, pixels in predefined neighborhoods are encouraged to be composed of the same materials at roughly the same proportion levels. The predefined

neighborhoods are not restricted to spatial neighborhoods and could be neighborhoods-based spectral similarity. Similar ideas have been applied using the piece-wise convex unmixing methods, which identify groups of pixels (not necessarily spatial neighbors) that are constrained to have same endmember elements during unmixing [24].

These additional constraints based on spatial or spectral neighborhoods impose additional structures that further confine the solution space during spectral unmixing. However, the applicability of each of these additional constraints is certainly data- and application dependent.

ENDMEMBERS AS DISTRIBUTIONS

An alternative to the set-based approach to address spectral variability in a material is the use of a multivariate statistical distribution. When endmembers are represented as statistical distributions, then, a sample from these distributions can be viewed as a possible variation on the spectral signature of the material being represented

$$\mathbf{e}_m \sim \mathcal{F}(\cdot | \boldsymbol{\theta}_m), \quad (10)$$

where \mathcal{F} is the multivariate statistical distribution used to represent an endmember, and $\boldsymbol{\theta}_m$ are parameters of the distribution associated with the m th material. Under this approach, hyperspectral pixels are random variables distributed according to the stochastic mixture model defined by the convex combination of the endmembers, $\mathbf{x}_n = \sum_{m=1}^M p_{mn} \mathbf{e}_m$, where \mathbf{e}_m is a random variable distributed according to $\mathcal{F}(\mathbf{e}_m | \boldsymbol{\theta}_m)$. There is no additive noise in this model as variation in the \mathbf{x} value are accounted for through the stochastic model and the variability of the endmember distributions.

A number of methods make use of a Bayesian approach for endmember estimation and spectral unmixing. When the distributions of the endmembers are completely specified, joint estimation of endmembers and proportions, such as Bayesian source separation and nonnegative matrix factorization, can also provide endmember variations in spectral unmixing [25], [26]. In particular, [25] assumes a Gamma distribution as the prior for each value of the endmembers. Such an approach requires the exact knowledge of the endmember distributions, which may not be easy to obtain in practice. In the normal compositional model (NCM) methods described below, the models imposed on the endmember distributions do not require complete specification and the unknown distribution parameters are jointly estimated together with proportions during unmixing.

NORMAL COMPOSITIONAL MODEL

By far, the most prominent statistical distribution for endmembers is the NCM [27] and, as a result,

$$x_n \sim \mathcal{N} \left(\cdot \left| \sum_{m=1}^M p_{mn} \boldsymbol{\mu}_m, \sum_{m=1}^M p_{mn}^2 \boldsymbol{\Sigma}_m \right. \right), \quad (11)$$

in which $\mathcal{F}(e_m | \theta_m) = \mathcal{N}(e_m | \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$, where $\boldsymbol{\mu}_m$ are endmember mean values and $\boldsymbol{\Sigma}_m$ are the endmember covariances. A number of methods for spectral unmixing and endmember estimation have been developed using the NCM.

Stein [27] proposed an expectation-maximization algorithm to iteratively update proportion values for every input pixel as well as the mean and variance parameter values for each endmember distribution using the NCM, until convergence is reached.

Under the condition that the mean values of the endmembers are known, Eches et al. [28] developed a Markov chain Monte Carlo (MCMC) sampling approach for estimating proportion values and the endmember distribution covariance using the NCM. The hierarchical Bayesian framework is

$$x_n | \mathbf{p}_n, \sigma^2 \sim \mathcal{N} \left(\cdot \left| \sum_{m=1}^M p_{mn} \boldsymbol{\mu}_m, \sum_{m=1}^M p_{mn}^2 \sigma^2 \mathbf{I} \right. \right) \quad (12)$$

$$\mathbf{p}_n \sim \mathcal{D}(\cdot | \mathbf{1}) \quad (13)$$

$$\sigma^2 | \delta \sim \mathcal{IG}(\nu, \delta), \quad (14)$$

where $\mathcal{D}(\cdot | \mathbf{1})$ is Dirichlet distribution such that all parameter values are equal to one (i.e., a uniform distribution over the set of proportion values that satisfy the nonnegativity and sum-to-one constraints), $\mathcal{IG}(\nu, \delta)$ is an inverse-gamma distribution with parameter values ν and δ in which, for the implementation presented in [28], $\nu = 1$ and δ is assigned a given noninformative Jeffreys prior, and \mathbf{I} is the identity matrix. To estimate the unknown proportion $\mathbf{p}_n = [p_{1n}, p_{2n}, \dots, p_{Mn}]^T, \forall n = 1, \dots, N$, the endmember variance σ^2 and the parameter δ , a Metropolis-within-Gibbs sampler is used. In [28], endmember covariances are assumed to be diagonal and may not be identical.

Zare et al. [24] also developed an MCMC sampling approach under the NCM. However, in contrast to the method by Eches et al., it assumes endmember covariance values are available and the endmember means and proportion values are to be estimated. The hierarchical Bayesian framework defined for a single set of endmember distributions is

$$x_n | \boldsymbol{\Pi}, \mathbf{p}_n \sim \mathcal{N} \left(\cdot \left| \sum_{m=1}^M p_{mn} \boldsymbol{\mu}_m, \sum_{m=1}^M p_{mn}^2 \boldsymbol{\Sigma}_m \right. \right) \quad (15)$$

$$\boldsymbol{\mu}_m \sim \mathcal{N}(\cdot | \mathbf{m}, \mathbf{C}) \quad (16)$$

$$\mathbf{p}_n \sim \mathcal{D}(\cdot | \mathbf{1}), \quad (17)$$

where $\boldsymbol{\Pi} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_M]$ is the collection of endmember means, $\mathbf{p}_n = [p_{1n}, p_{2n}, \dots, p_{Mn}]^T$ is the vector of proportion values, and \mathbf{m} and \mathbf{C} are hyperparameters defining the prior distribution on the endmember means. The \mathbf{m} and \mathbf{C} values are parameters to be fixed (such as setting them to the mean and covariance data). In [24], additional hyperpriors are applied to \mathbf{m} and \mathbf{C} such that $\mathbf{m} \sim \mathcal{N}(\cdot | (1/N) \sum_{n=1}^N x_n, \mathbf{V})$ and $\mathbf{C} \sim IW(\Psi, t)$ to allow for multiple sets of endmember distributions in a piece-wise convex mixing model. Using the piece-wise convex approach, input data pixels are partitioned into sets that are composed of mixtures of the same materials. The partitions are estimated by incorporating a Dirichlet process prior from which partitions of the input data set are sampled. To estimate the endmember means and

proportion values, a Metropolis-within-Gibbs sampler is proposed.

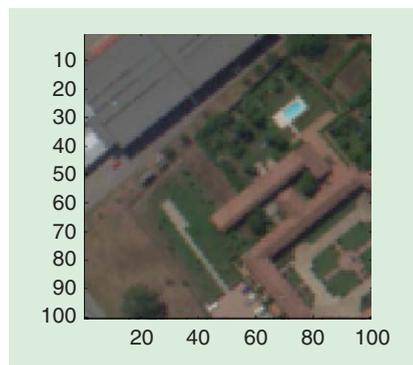
BETA COMPOSITIONAL MODEL

The majority of investigation with endmembers as distributions has been conducted using the NCM. Yet, investigation into alternative compositional models, such as the beta compositional model (BCM) has been conducted [29]. Under the BCM, the input data points are random variables distributed according to a convex combination of beta random variables. The motivation for the use of the beta is that the values are constrained to the range from zero to one, which is appropriate and a physically meaningful range for endmember reflectance values. The spectral unmixing method developed in [29], based on the BCM, assumes known endmember parameter values and estimates proportion values for input pixels using an approximation to the BCM.

METHODS OF HIGHER MOMENTS

In the approach presented by Bosdogianni et al. [30], rather than defining a fixed parametric form for each endmember distribution, given known values for the first and second moments of each endmember distribution, proportion values of an input data set are estimated by minimizing the squared difference between the first and second moments of the

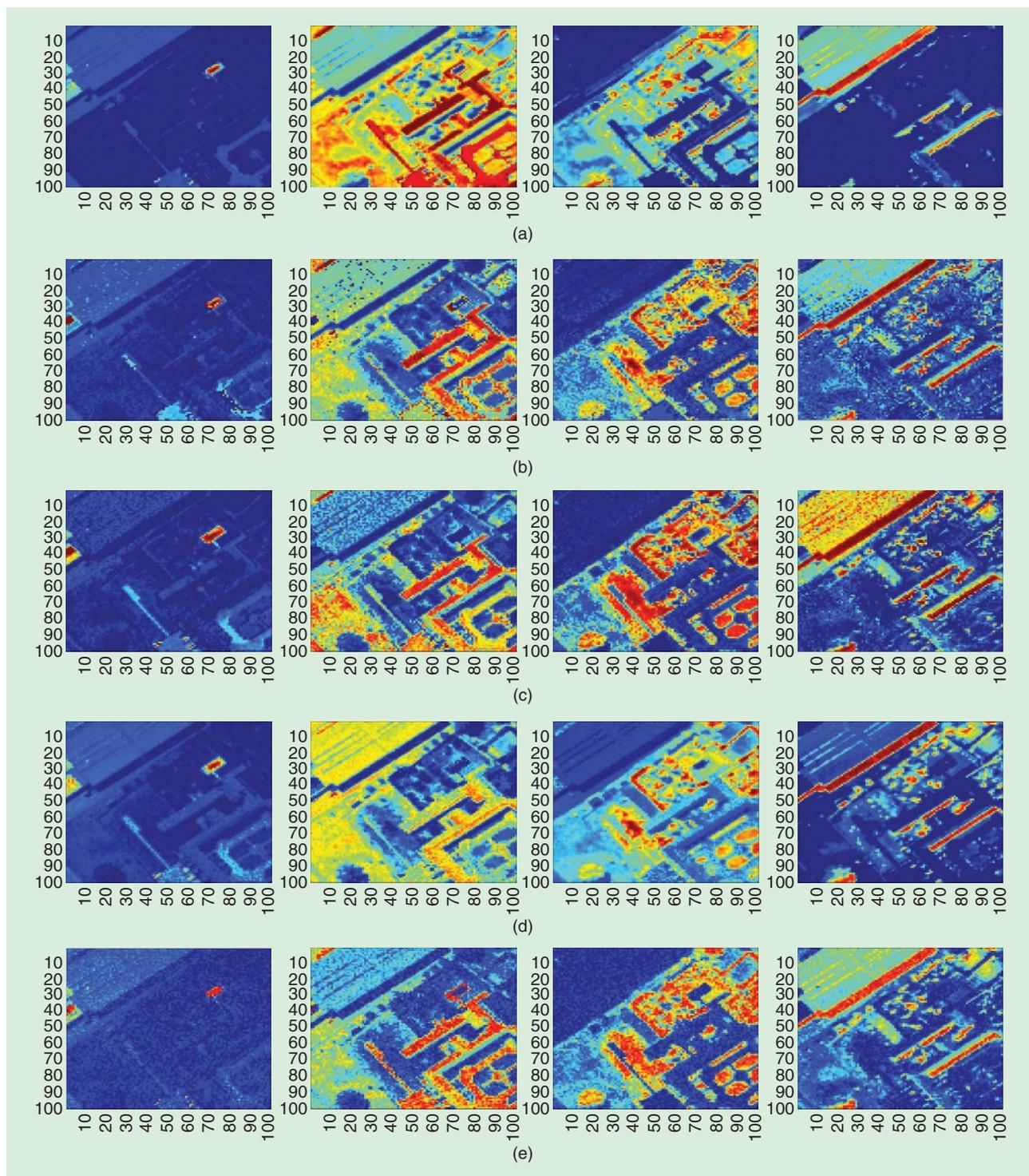
IN GENERAL, ENDMEMBERS AS A DISTRIBUTIONS APPROACH COULD HAVE A COMPUTATIONAL ADVANTAGE OVER ENDMEMBERS AS SETS IF THE ENDMEMBER SETS ARE LARGE.



[FIG1] RGB image (bands 56, 29, and 12) of the Pavia subscene.

estimated convex combination of the endmember values and those of the input data. Given that statistics of the full input data set are used to estimate proportion values, pixel-specific proportion values are not estimated but, instead, test-site wide values. An advantage to this approach is that the full

parametric form for each endmember distribution does not need to be specified and, instead, only the first and second moments (equivalent to mean and covariance) of each endmember distribution are needed. Using only the first two moments follows a symmetric distribution as in the NCM.



[FIG2] The proportion maps of the Pavia subimage from VCA + FCLS: (a) AEB, (b) LU, (c) NCM-1 [28], (d) NCM-2 [24], and (e) of water; buildings and dirt; vegetation; and shadow.

However, presumably, this approach could be extended to include higher-order moments.

RESULTS

A sampling of five methods, one from typical unmixing method without accounting for spectral variability, and two from each representation of spectral variability are used

- 1) VCA+FCLS
- 2) AEB by Somers et al. [19]
- 3) LU approach by Goenaga et al. [23]
- 4) NCM approach by Eches et al. [28] NCM-1
- 5) NCM approach by Zare et al. [24] NCM-2.

These algorithms were applied to a 100×100 pixel subimage collected by the Reflective Optics System Imaging Spectrometer (ROSIS) over an urban area of Pavia in northern Italy on 8 July 2002. The RGB image of the scene is illustrated in Figure 1. The image contains spectra from vegetation and man-made objects.

All the methods were set with four endmember sets or distributions ($M = 4$). Note that the performance could vary when a different value of M is used, too-large M splits a material across multiple proportion values, and too-small M groups distinct materials. The results are shown in Figure 2.

Each of these methods requires a number of parameter settings that are encumbered with various tradeoffs between running time and accuracy. In VCA+FCLS, VCA is used for endmember extraction and fully constrained least-squares (FCLS) [31] is applied for proportion estimation. It does not assume spectral variability and is served as a reference for comparison. The AEB estimated the four endmember sets by repeatedly applying the VCA algorithm with 3,000 iterations to 5,000 randomly sampled pixels from the scene. The proportion values were obtained using MESMA [19]. LU obtained endmembers on 33×33 -sized tiles across the image using VCA [20]. After which, the resulting local endmembers were clustered into four endmember sets from which proportion values were estimated [23]. NCM-1 generated the means of the four endmembers using VCA and per-pixel proportion values were estimated using 1,000 iterations of MCMC-sampling. NCM-2 used isotropic diagonal covariance matrices with values of 0.01, and the corresponding MCMC sampling algorithm was iterated 200 times.

The results are shown in Figure 2. As a measure of improvement in spectral unmixing, we used two approaches: 1) the average per-pixel squared residual (signal fitting) error for each pixel and 2) the median of the proportion maps over the five methods as an ideal output to compute the residual unmixing error. The average squared residual errors were

- VCA+FCLS: 0.403
- AEB: 0.009
- LU: 0.010

- NCM-1: 0.191
- NCM-2: 0.108.

The percentage in mean-square residual errors of the five methods were

- VCA+FCLS: 32.3%
- AEB: 5.9%
- LU: 9.1%
- NCM-1: 10.6%
- NCM-2: 5.2%.

The performance of the four methods with spectral variability is consistent and the amount of error reduction over VCA+FCLS is significant. Obviously, the cost of better results from spectral variability unmixing is the increase in computation time, which is proportional to the number of endmembers M . In general, endmembers as a distributions approach could have a computational advantage over endmembers as sets if the endmember sets are large.

**AN OPEN PROBLEM REMAINS
IN HOW TO DETERMINE WHETHER
A PARTICULAR SPECTRUM IS
A VARIATION ON AN ENDMEMBER
OR A MIXED PIXEL WITH HIGH
ABUNDANCE OF AN ENDMEMBER.**

SUMMARY AND FUTURE WORK

This article presented an overview of methods that address spectral variability for hyperspectral unmixing and endmember estimation. Although significant progress has been made in this area, there are several open lines of research. In particular, as discussed earlier, many of the current approaches rely on the availability of known spectral libraries, and results are highly dependent on the availability of appropriate endmember spectra. Given an extremely large spectral library, investigation into automated pruning and subselection of the library for maintaining scene-appropriate endmembers is needed. The applicability of these libraries depends upon items such as the inclusion of the materials found in the scene with the exclusion of spectrally similar confusers.

Automated parameter setting for all of these approaches is another subject for study. In particular, an appropriate selection of the number of endmembers can have a large impact on the quality of results. Many of the existing methods rely on manual selection through trial and error or broad assumptions (i.e., fewer numbers of endmembers is better).

Additional investigation of methods to leverage spatial information can be conducted. Methods employed in the image processing and computer vision communities make extensive use of color and texture information for scene understanding. An area of study is in how these or similar approaches can be transferred to hyperspectral image analysis. The challenge is being able to effectively balance spatial information without loss of the subpixel information.

In addition to linear mixing model as in (1), nonlinear mixing (such as the bilinear model [1]) can be used to account for different phenomena such as multiple reflections or intimate mixtures of materials, resulting in more accurate endmember and proportion estimates. Interesting future work will include incorporating spectral variability into nonlinear mixing models and identifying when nonlinear or linear models with or without spectral variability are needed.

Additionally, future investigation into models that limit variability to physically meaningful values should be conducted. Also, all of the current methods in this area assume independent variation between the bands of an endmember. In practice, neighboring bands are often highly correlated and extension of these models to make use of full covariances to account for this correlation is needed. These models are also currently limited to representing all endmembers with distributions of the same parametric form. It is possible that differing materials may be best represented with different forms. The extension of these approaches to mixed-form models is an interesting subject for research.

Finally, future work remains in developing evaluation metrics or measured data sets with proportion-level ground-truth to allow for evaluation of methods that address spectral variability. An open problem remains in how to determine whether a particular spectrum is a variation on an endmember or a mixed pixel with high abundance of an endmember.

AUTHORS

Alina Zare (zare@missouri.edu) is an assistant professor in the Electrical and Computer Engineering Department, University of Missouri. Her research interests include machine learning, Bayesian methods, SONAR analysis, image analysis, and hyperspectral image analysis.

K.C. Ho (hod@missouri.edu) is a professor in the Electrical and Computer Engineering Department of the University of Missouri. His research interests are in sensor array processing, source localization, and remote sensing. He is a Fellow of the IEEE.

REFERENCES

- [1] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Select. Topics Appl. Earth Observ.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [2] J. Adams, D. Sabol, V. Kapos, R. Filho, D. Roberts, M. Smith, and A. Gillespie, "Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon," *Remote Sensing Environ.*, vol. 52, no. 2, pp. 137–154, May 1995.
- [3] P. Dennison, K. Halligan, and D. Roberts, "A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper," *Remote Sensing Environ.*, vol. 93, pp. 359–367, Nov. 2004.
- [4] F. Garcia-Haro, S. Sommer, and T. Kemper, "A new tool for variable multiple endmember spectral mixture analysis (VMESMA)," *Int. J. Remote Sensing*, vol. 26, no. 10, pp. 2135–2162, May 2005.
- [5] J.-Ph. Combe, S. Le Mouelic, C. Sotin, A. Gendrin, J. F. Mustard, L. Le Deit, P. Launeau, J.-P. Bibring, B. Gondet, Y. Langevin, and P. Pinet, "Analysis of OMEGA/Mars Express data hyperspectral data using a multiple-endmember linear spectral unmixing model (MELSUM): Methodology and first results," *Planet. Space Sci.*, vol. 56, pp. 951–978, Dec. 2008.
- [6] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 6, pp. 2706–2717, Nov. 1999.
- [7] J. Nascimento and J. Bioucas-Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 1, pp. 175–187, Jan. 2005.
- [8] B. Gao, M. Montes, C. Davis, and A. Goetz, "Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean," *Remote Sensing Environ.*, vol. 113, Suppl. 1, pp. S17–S24, Sept. 2009.

- [9] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sensing Environ.*, vol. 115, no. 7, pp. 1603–1616, July 2011.
- [10] D. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. O. Green, "Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models," *Remote Sensing Environ.*, vol. 65, pp. 267–279, Sept. 1998.
- [11] C. Song, "Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability?" *Remote Sensing Environ.*, vol. 95, no. 2, pp. 248–263, Mar. 2005.
- [12] G. Asner, M. Bustamante, and A. Townsend, "Scale dependence of biophysical structure in deforested areas bordering the Tapajs national forest, Central Amazon," *Remote Sensing Environ.*, vol. 87, pp. 507–520, Mar. 2003.
- [13] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, no. 2, pp. 1083–1093, Mar. 2000.
- [14] J. Jin, B. Wang, and L. Zhang, "A novel approach based on Fisher discriminant null space for decomposition of mixed pixels in hyperspectral imagery," *IEEE Geosci. Remote Sensing Lett.*, vol. 7, no. 4, pp. 699–703, Oct. 2010.
- [15] P. Dennison and D. Roberts, "Multiple endmember spectral mixture analysis using endmember average (MESMA)," *Remote Sensing Environ.*, vol. 87, nos. 2–3, pp. 123–135, Oct. 2003.
- [16] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006.
- [17] F. Mianji and Y. Zhang, "SVM-based unmixing-to-classification conversion for hyperspectral abundance quantification," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4318–4327, Nov. 2011.
- [18] F. Bovolo, L. Bruzzone, and L. Carlini, "A novel technique for subpixel image classification based on support vector machine," *IEEE Trans. Image Processing*, vol. 19, no. 11, pp. 2983–2999, Nov. 2010.
- [19] B. Somers, M. Zortea, A. Plaza, and G. Asner, "Automated extraction of image-based endmember bundles for improved spectral unmixing," *IEEE J. Select. Topics Appl. Earth Observ.*, vol. 5, no. 2, pp. 396–408, Apr. 2012.
- [20] J. Nascimento and J. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
- [21] A. Castrodad, Z. Xing, J. Greer, E. Bosch, L. Carin, and G. Sapiro, "Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4263–4281, Nov. 2011.
- [22] K. Canham, A. Schlamm, A. Ziemann, B. Basener, and D. Messinger, "Spatially adaptive hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4248–4262, Nov. 2011.
- [23] M. Goenaga, M. Torres-Madronero, M. Velez-Reyes, S. Van Bloem, and J. Chinaea, "Unmixing analysis of a time series of Hyperion images over the Gunica dry forest in Puerto Rico," *IEEE J. Select. Topics Appl. Earth Observ.*, vol. 6, no. 2, pp. 329–338, Apr. 2013.
- [24] A. Zare, P. Gader, and G. Casella, "Sampling piecewise convex unmixing and endmember extraction," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 3, pp. 1655–1665, Mar. 2013.
- [25] S. Moussaoui, D. Brie, A. Mohammad-Djafari, and C. Carteret, "Separation of non-negative mixture of non-negative sources using a Bayesian approach and MCMC sampling," *IEEE Trans. Signal Processing*, vol. 54, no. 11, pp. 4133–4145, Nov. 2006.
- [26] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.
- [27] D. Stein, "Application of the normal compositional model to the analysis of hyperspectral imagery," in *Proc. IEEE Workshop Advances in Techniques for Analysis of Remotely Sensed Data*, Oct. 2003, pp. 44–51.
- [28] O. Eches, N. Dobigeon, C. Mailhes, and J. Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model. Application to hyperspectral imagery," *IEEE Trans. Image Processing*, vol. 19, no. 6, pp. 1403–1413, June 2010.
- [29] A. Zare, P. Gader, D. Dranishnikov, and T. Glenn, "Spectral unmixing using the beta compositional model," in *Proc. IEEE Workshop. Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Gainesville, FL, June 2013.
- [30] P. Bosdogianni, M. Petrou, and J. Kittler, "Mixture models with higher order moments," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, no. 2, pp. 341–353, Mar. 1997.
- [31] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 529–545, Mar. 2001.



[Gonzalo R. Arce, David J. Brady, Lawrence Carin, Henry Arguello, and David S. Kittle]

Compressive Coded Aperture Spectral Imaging



[An introduction]

Imaging spectroscopy involves the sensing of a large amount of spatial information across a multitude of wavelengths. Conventional approaches to hyperspectral sensing scan adjacent zones of the underlying spectral scene and merge the results to construct a spectral data cube. Push broom spectral imaging sensors, for instance, capture a spectral cube with one focal plane array (FPA) measurement per spatial line of the scene [1], [2]. Spectrometers based on optical bandpass filters sequentially scan the scene by tuning the bandpass filters in steps. The disadvantage of these techniques is that they require scanning a number of zones linearly in proportion to the desired spatial and spectral resolution. This article surveys

compressive coded aperture spectral imagers, also known as *coded aperture snapshot spectral imagers (CASSI)* [1], [3], [4], which naturally embody the principles of compressive sensing (CS) [5], [6]. The remarkable advantage of CASSI is that the entire data cube is sensed with just a few FPA measurements and, in some cases, with as little as a single FPA shot.

INTRODUCTION

CS dictates that one can recover spectral scenes from far fewer measurements than that required by conventional linear scanning spectral sensors. To make this possible, CS relies on two principles: sparsity, which characterizes the spectral scenes of interest, and incoherence, which shapes the sensing structure [5], [7]. Sparsity indicates that spectral images found in nature can be concisely represented in some basis Ψ with just a small

Digital Object Identifier 10.1109/MSP.2013.2278763

Date of publication: 5 December 2013

number of coefficients. This is indeed the case in spectral imaging where natural scenes exhibit correlation among adjacent pixels and also across spectral bands [2]. Incoherence refers to the structure of the sampling waveforms used in CS that, unlike the signals of interest, have a dense representation in the basis Ψ [7]. The remarkable discovery behind CS is that it is possible to design sensing protocols capable of capturing the essential information content in sparse signals with just a small number of compressive measurements. The sensing modality simply correlates incoming signals with a small number of fixed waveforms that satisfy the incoherence principle. The signals of interest are then accurately reconstructed from the small number of compressive measurements by numerical optimization [5], [6], [8]–[11]. In CASSI, the random projections occur naturally as the result of the optical dispersion phenomena affecting coded aperture light fields as they transverse a prism before these are integrated by the imaging detector.

Our intent in this article is to overview the fundamental optical phenomena behind compressive spectral imaging sensors, present the key mathematical concepts embodying the sensing and reconstruction mechanisms, and describe the optimization framework used to design optimal coded apertures in a number of applications, including hyperspectral image reconstruction, spectral selectivity, and superresolution. The article describes many practical aspects of the instrumentation, including calibration, discretization models, parameter design, and physical limitations, and it illustrates results with real data and imagery. A fascinating aspect of compressive spectral sensing is that it draws from various disciplines within optics, signal processing, and probability theory. Our survey is to highlight the

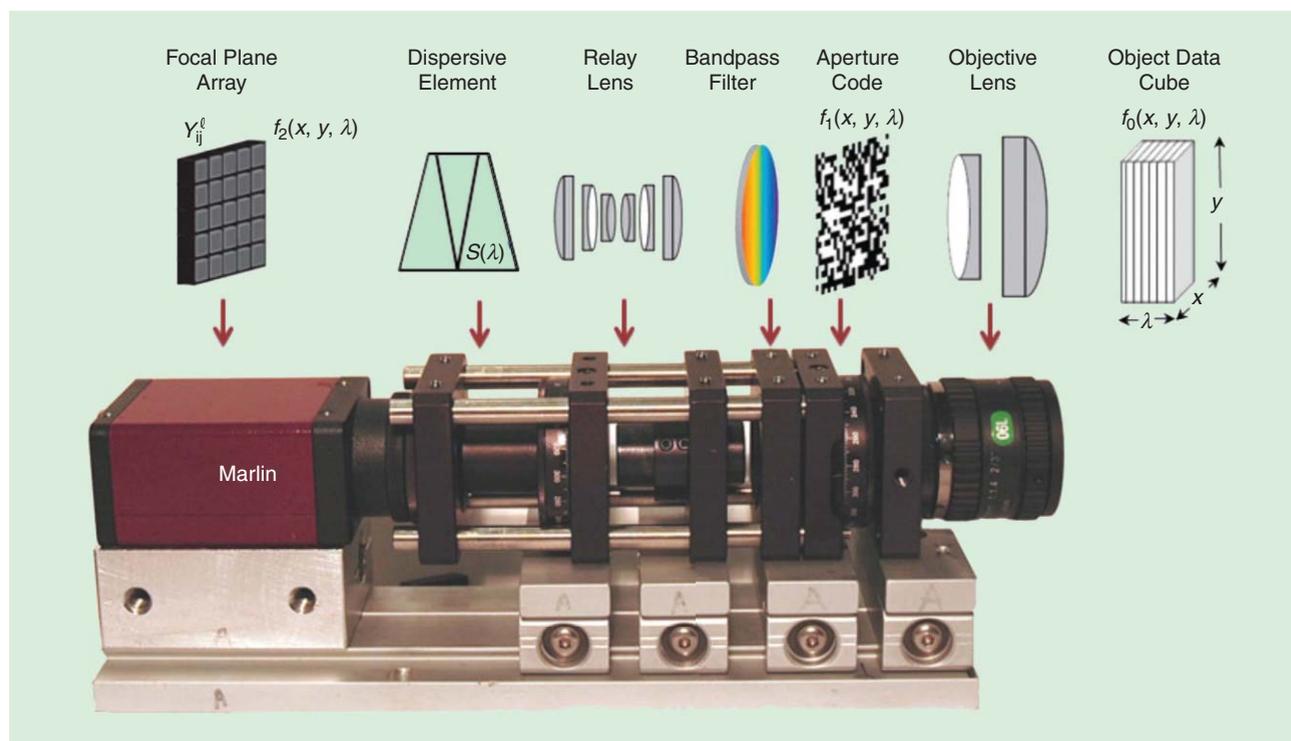
rich interaction among these fields of study as they come together in the discovery of novel compressive spectral sensors. Our treatment explains as plainly as possible four fundamental aspects of CASSI:

- 1) the sensing problem
- 2) coded aperture optimization
- 3) reconstruction algorithms
- 4) computational spectral imaging.

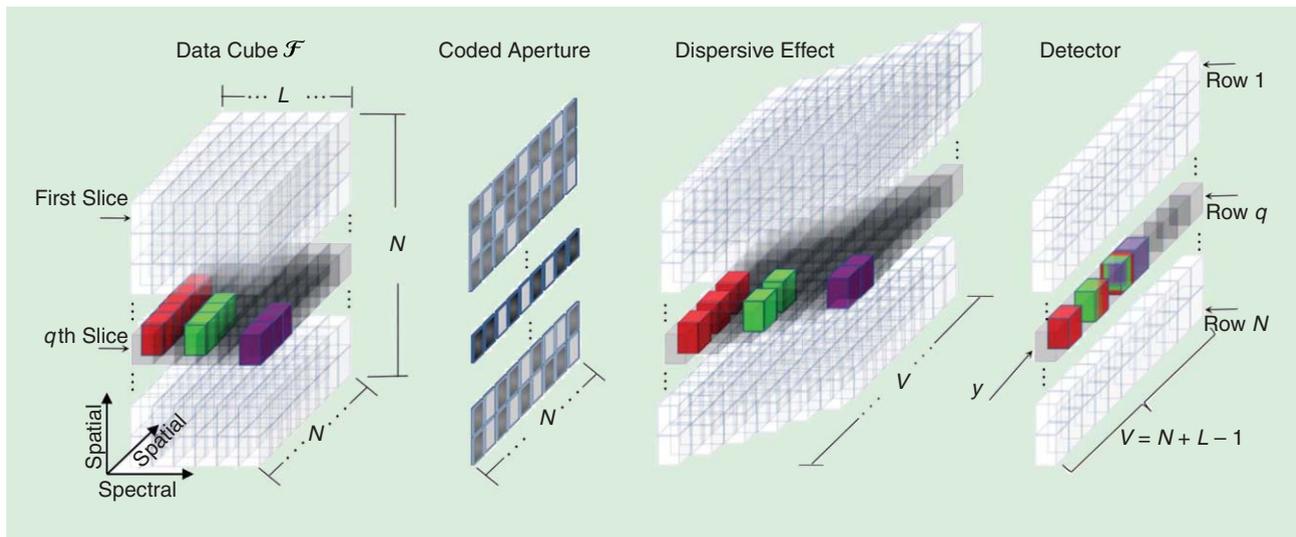
It should be noted that compressive spectral imaging is of interest in many fields and advances in this area are growing rapidly. Consequently, it is not possible to review all work in the field, and we focus on CASSI type compressive optical sensing. The topics in this tutorial are complemented by a companion article in this issue [12], where other approaches and methods are described.

THE SPECTRAL SENSING PROBLEM

The sensing physical phenomena in CASSI is strikingly simple, yet it adheres to the incoherence principles required in CS. In its simplest form, CASSI measurements are realized optically by a coded aperture, a dispersive element such as a prism, and an FPA detector [1], [3]. The coding is applied to the (spatial-spectral) image source density $f_0(x, y, \lambda)$ by means of a coded aperture $T(x, y)$ as realized by the CASSI system depicted in Figure 1, where (x, y) are the spatial coordinates and λ is the wavelength [3]. The resulting coded field $f_1(x, y, \lambda)$ is subsequently modified by a dispersive element before it impinges onto the FPA detector. The compressive measurements across the FPA are realized by the integration of the dispersed field $f_2(x, y, \lambda)$ over the detector's spectral range sensitivity.



[FIG1] Compressive CASSI sensor components. (Image courtesy of David J. Brady and David S. Kittle.)



[FIG2] An illustration of the spectral optical flow in CASSI. The q th slice of the data cube \mathcal{F} with 11 nonzero voxels is coded by a row of the coded aperture and dispersed by the prism. The detector captures the intensity y by integrating the coded light.

The sensing mechanism is illustrated by the discretized model shown in Figure 2, where the spectral data cube \mathcal{F} having L spectral bands and $N \times N$ spatial pixels is first amplitude modulated by a pixelated $N \times N$ coded aperture T .

In this case, T is a *block-* or *unblock-*coded aperture such that the energy along the spectral coordinate of the data cube is first punched out when a block-coded aperture element is encountered. As the coded field transverse the prism, it is then spatially sheared along one spatial axis. In essence, each coded image plane at a fixed wavelength is shifted along the x-axis where the amount of shifting increases with the wavelength coordinate index. Finally, the coded and dispersed field is “collapsed” in the spectral dimension by the integration of the energy impinging on each detector element over its spectral range sensitivity. The integrated field is then measured by the FPA detector elements.

Several properties of the sensing phenomena model should be pointed out [13], [14]. First, note that the $N \times N$ spatial dimensions of the spectral data cube are mapped to an array of $V \times N$ FPA measurements, where $V = N + L - 1$. This is due to the dispersion of the optical field as it transverse the prism. Second, observe that the optical coding across rows of the FPA measurements are mutually independent. That is, the aperture coding affecting one row slice of the data cube is independent from the coding affecting other row slices in the data cube. These characteristics of the sensing phenomena are important, as they shape subsequent signal processing algorithms. Third, note how the sequence of optical transformations altogether end up in a set of compressive linear measurements in the FPA detector. As Figure 2 illustrates, each FPA shot captures a massive set of compressive measurements.

The discretized output at the detector can thus be modeled as [3]

$$Y_{j\ell} = \sum_{k=0}^{L-1} \mathcal{F}_{j(\ell+k)(k)} T_{j(\ell+k)} + \omega_{j\ell}, \quad (1)$$

where $Y_{j\ell}$ is the intensity measured at the j, ℓ position of the detector whose dimensions are $N \times (N + L - 1)$, L is the number of spectral bands, $T_{j\ell}$ is the binary coded aperture, and $\omega_{j\ell}$ is the noise of the system. In essence, (1) sums each of the spectral image slices that have been coded and spatially shifted in proportion to the wavelength index k . Notice in (1) that each discrete spectral band is defined such that a continuous region of the analog spectrum span one pixel in the detector.

Assume that the bandpass filter of the instrument limits the spectral components between λ_1 and λ_2 . If the pixel width of the detector and of the coded aperture are both equal to Δ , then the number of resolvable bands L is limited by $L = \alpha((\lambda_2 - \lambda_1)/(\Delta))$, where for a given wavelength λ , $\alpha\lambda$ corresponds to the dispersion induced by the prism, i.e., to the displacement of light in the focal plane along the x-axis.

The spectral resolution is limited by Δ/α . The horizontal and vertical spatial resolutions are limited by Δ , and the number of spatially resolvable pixels of the underlying scene is $N \times N$.

For spectrally rich scenes or very detailed spatial scenes, a single-shot FPA measurement is not sufficient, and additional shots are required, each with a distinct coded aperture that remains fixed during the integration time of the detector. Time-varying coded apertures can be realized by a spatial light modulator or by a lithographic mask actioned by a piezoelectric device [4], [15], [16]. It was also shown in [15] and [16] that the ensemble of say, $K \ll L$ FPA shots in one-dimensional vectorized form $y = [y_0^T, \dots, y_{K-1}^T]^T$, can be rewritten in the standard form of an underdetermined system of linear equations

$$y = A\theta = H\Psi\theta + \omega, \quad (2)$$

where $A = H\Psi \in \mathbb{R}^{KN^2 \times LN^2}$ is the CASSI sensing matrix, θ is a sparse representation of the data cube in a three-dimensional (3-D) basis Ψ , and ω represents the noise of the system. A

Kronecker basis $\Psi = \Psi_1 \otimes \Psi_2$ is often used, where Ψ_1 is the two-dimensional-wavelet Symmlet-8 basis and Ψ_2 is the cosine basis [17]. The matrix H in (2) accounts for the effects of the coded aperture and the prism. The sensing matrix A thus couples H with the representation basis Ψ . The coded aperture is considered binary, and the dispersive element is considered linear. In practice, it is necessary to take into account the various optical artifacts and nonideal characteristic of the optical system. Furthermore, the underlying principles of CASSI described above are general and are thus, in principle, applicable to imaging with FPAs sensitive to any region of the visible and infrared (IR) radiation of the electromagnetic spectrum.

Signal recovery in CASSI entails solving an underdetermined linear system of equations. Given the set of measurements y , the

inverse CS problem consists on recovering θ such that the $\ell_2 - \ell_1$ cost function is minimized

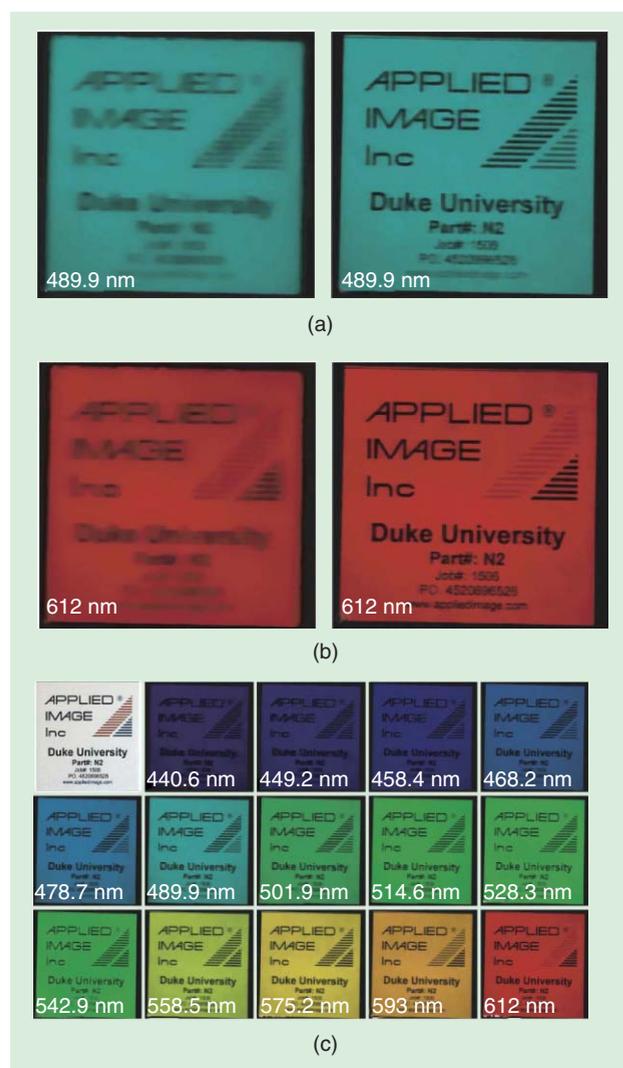
$$\arg \min_{\theta} \|y - A\theta\|^2 + \lambda \|\theta\|_1, \quad (3)$$

where λ is a regularization constant. The inverse problem in (3) will be addressed next. At this point, however, it should be emphasized that the sensing matrix A will play a pivotal role in sensing and thus its design is of key importance. Other cost functions could be used instead of (3). For example, an approach based on the stable recovery of a low-rank and joint-sparse matrix can be used to reconstruct the hyperspectral information. In this case, the optimization problem is regularized to minimize jointly a nuclear norm and a $\ell_2 - \ell_1$ norm [18].

To illustrate the underlying concepts above, a wide spectral bandwidth reconstruction is shown in Figure 3, acquired with a recently developed UV-visible CASSI instrument [19]. The object was illuminated using SoLux daylight emulation bulbs and bandpass filtered by a Baader Planetarium (Germany) UV-IR cut filter with transmission from 420 to 680 nm. A random, binary coded aperture was used in the instrument with a minimum feature size of two pixels and total area on the detector of $1,988 \times 1,988$ pixels. Figure 3 shows the comparison between a snapshot reconstruction and a multishot reconstruction with 24 frames. The prism dispersion is nonlinear, where the shorter wavelengths disperse significantly more than longer wavelengths. This yields higher spectral resolution and lower spatial resolution in the shorter wavelengths and corresponding higher spatial resolution and lower spectral resolution for longer wavelengths, shown in the blue spectral channel in Figure 3(a) and the red spectral channel in (b), respectively.

Code features fewer than two pixels place high demands on the optics, alignment, and calibration of the CASSI instrument, especially for small pixel size detectors. The coded aperture is required to disambiguate the spatial and spectral information that is multiplexed at the detector. For smaller code features, the optical blur, forward model, calibration, detector noise, and pixel sampling of the code have a larger impact on the reconstruction. A code feature size of two pixels (total area of four pixels) guarantees that regardless of where the feature is sampled on the detector, it will cover at least one pixel to fully modulate the signal. Smaller features will generally modulate less, where the worst case scenario involves a one-pixel feature sampled by four pixels, reducing the modulation to 25%. In general, sampling a square wave (in our case, the coded aperture) requires infinite sample points. Even a sine wave would require at least two pixels assuming spatial Nyquist sampling and interpolation under the assumption of lowpass filtering. At the limits of small code features, adequate models [14] are required to compensate for undersampling the coded aperture.

A calibration process is realized in CASSI in which the spectral channels are determined so that the centers of each channel, at the finest scale, are separated by one column of detector pixels. Thus, the position of a channel corresponds to a fixed



[FIG3] A comparison between snapshot and multishot CASSI reconstructions: (a) snapshot, 490 nm and 24-frame reconstruction at 490 nm, (b) snapshot, 612 nm and a 24-frame reconstruction at 612 nm, and (c) selected wavelengths from the 24-frame, reconstructed data cube. An RGB image of the object taken with an SLR digital camera is shown in the upper left part of (c). Notice the relative spatial resolution for the blue and red spectral channels.

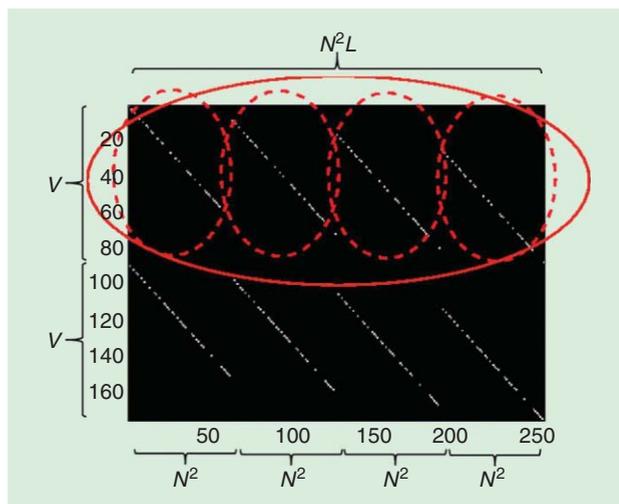
dispersion in terms of detector pixels relative to a fixed spectral channel. Charge-coupled device (CCD) measurements at equally spaced wavelengths are taken after uniformly illuminating the coded aperture with monochromatic wavelength of light within the bandpass of the system. These calibration measurements are used to build a modified system operator that accounts for the optical blur and nonlinear dispersion. The set of measurements is obtained with careful efforts to reduce or remove certain data corruptive processes, including the dark noise on the CCD, the nonuniform spectral intensity of the calibrating light source, and the nonuniform spectral sensitivity of the CCD. An alternative calibration procedure can be realized using specifically designed coded apertures for calibration and by using a high-order model of the CASSI phenomena [14].

CODED APERTURE OPTIMIZATION

The sensing matrix A in (1) plays a crucial role in the mathematics of the inverse CS problem. A large body of literature in CS specifically deals with characterizing the “goodness” of A [5], [6]. In CASSI, A is determined by the matrix H , which can be shown to have the structure shown in Figure 4 [4], [15], [16]. It consists of a set of diagonal patterns (circled) that repeat in the horizontal direction, each time with a unit downward shift, as many times as the number of spectral channels. Each diagonal pattern, circled in Figure 4, is the coded aperture pattern itself. Just below, the next set of diagonal patterns are determined by the coded aperture pattern used in the next FPA shot. The matrix H will thus have as many sets of diagonal patterns as FPA measurements. Initially, commonly used coded apertures in CASSI included Hadamard matrices, S matrices, and Bernoulli random matrices [1].

The use of these coded apertures was principally motivated by the realization that they are well conditioned when used in least square estimation [1]. However, these code designs do not fully exploit the rich theory of CS. Given that the coded apertures determine the nonzero entries of the sensing matrix, the important question can be asked: Can the coded apertures be optimally designed? Remarkably, the answer is yes, where the restricted isometry property (RIP) provides the optimization criteria [6].

The RIP establishes the conditions necessary for A such that the ℓ_2 norm of the underlying 3-D spectral image is approximately preserved under the transformation $A\theta$. More precisely, for each integer $S = 1, 2, \dots$, define the restricted isometry constant δ_s of the matrix A as the smallest constant such that the RIP inequality $(1 - \delta_s) \|\theta\|_2^2 \leq \|A\theta\|_2^2 \leq (1 + \delta_s) \|\theta\|_2^2$ holds for all S -sparse vectors θ [7]. A more intuitive description of the RIP is that it requires that subsets of S columns taken from A to be nearly orthogonal, or equivalently, that all $m \times |\mathcal{T}|$ column submatrices $A_{|\mathcal{T}|}$ of A are well conditioned for all $|\mathcal{T}| \leq S$. This, in turn, implies that all pairwise distances between S -sparse vectors are well preserved in the compressed signal space such that $(1 - \delta_{2s}) \|\theta_1 - \theta_2\|_2^2 \leq \|A\theta_1 - A\theta_2\|_2^2 \leq (1 + \delta_{2s}) \|\theta_1 - \theta_2\|_2^2$ holds for all S -sparse vectors θ_1, θ_2 . If the RIP holds for the sensing matrix A , one can discriminate among S -sparse signals



[FIG4] The sensing matrix H where the (circled) diagonal patterns repeating horizontally correspond to the coded aperture pattern used in the first FPA shot. The second coded aperture pattern determines the next set of diagonals. The figure depicts the sensing of four spectral bands and two FPA shots.

in the compressed domain and consequently it is then possible to accurately reconstruct any θ from its projection $A\theta$.

The RIP inequality can be rewritten as $\| \|A\theta\|_2^2 - \|\theta\|_2^2 \| \leq \delta_s \|\theta\|_2^2$ or equivalently as $|\langle (AA^T - I)\theta, \theta \rangle| \leq \delta_s \|\theta\|_2^2$, where I is an identity matrix. Constraining the vector θ to $\|\theta\|_2^2 = 1$, taking the supremum over all the vectors θ with $\text{supp}(\theta) \subset \mathcal{T}$, $|\mathcal{T}| \leq S$, and taking the maximum with respect to all the subsets \mathcal{T} leads to $\lambda_s = \max_{\mathcal{T} \subset [n], |\mathcal{T}| \leq S} s \lambda_{\max}(A_{|\mathcal{T}|} A_{|\mathcal{T}|}^T - I_{|\mathcal{T}|})$, where $A_{|\mathcal{T}|} A_{|\mathcal{T}|}^T = A_{|\mathcal{T}|}^T A_{|\mathcal{T}|}$, and $n = N^2L$. The probability of satisfying the RIP condition is thus calculated by estimating the statistical distribution of the maximum eigenvalue λ_{\max} of the matrices $A_{|\mathcal{T}|} A_{|\mathcal{T}|}^T - I_{|\mathcal{T}|}$.

The design strategy is then formulated as seeking the set of coded apertures $\{T_{j\ell}^0, \dots, T_{j\ell}^{K-1}\}$, such that

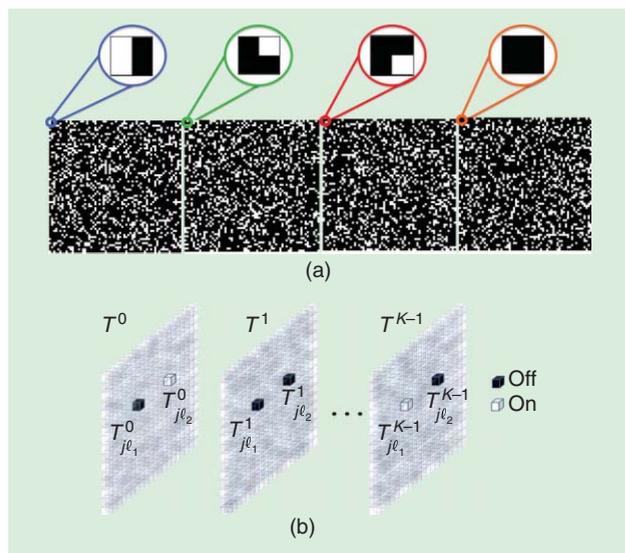
$$T_{j\ell}^0, \dots, T_{j\ell}^{K-1} = \underset{\{T_{j\ell}^0, \dots, T_{j\ell}^{K-1}\}}{\text{argmin}} \max_{\mathcal{T} \subset [n], |\mathcal{T}| \leq S} \lambda_{\max}(A_{|\mathcal{T}|}^T A_{|\mathcal{T}|} - I_{|\mathcal{T}|}), \quad (4)$$

where the entries of $A_{|\mathcal{T}|} A_{|\mathcal{T}|}^T$ are determined by the coded apertures. For a set of K coded apertures $\{T_{j\ell}^0, T_{j\ell}^1, \dots, T_{j\ell}^{K-1}\}$, as depicted in Figure 5, it turns out that the correlation variable [Figure 5(b)]

$$z_{\ell_1 \ell_2}^j = T_{j\ell_1}^0 T_{j\ell_2}^0 + T_{j\ell_1}^1 T_{j\ell_2}^1 + \dots + T_{j\ell_1}^{K-1} T_{j\ell_2}^{K-1} \quad (5)$$

strongly influences (4) and consequently the RIP in CASSI [20]. As depicted in Figure 5, $T_{j\ell_1}^i T_{j\ell_2}^i$ is the product of two elements of the i th coded aperture, both at the j th row, but at the column positions ℓ_1 and ℓ_2 . The first- and second-order statistics of $z_{\ell_1 \ell_2}^j$, specifically the mean $E\{z_{\ell_1 \ell_2}^j\} = m$ and $\text{Var}(z_{\ell_1 \ell_2}^j) = \sigma^2$, have a critical effect on (4). The expectation in this case is over the random selection of the entries $T_{j\ell_1}^i T_{j\ell_2}^i$.

Let the entries of Ψ be $\Psi_{j,k}$, then using the structure of the matrices H in Figure 4, the entries $(A_{|\mathcal{T}|} A_{|\mathcal{T}|}^T)_{k_1 k_2}$ can be obtained as [4]

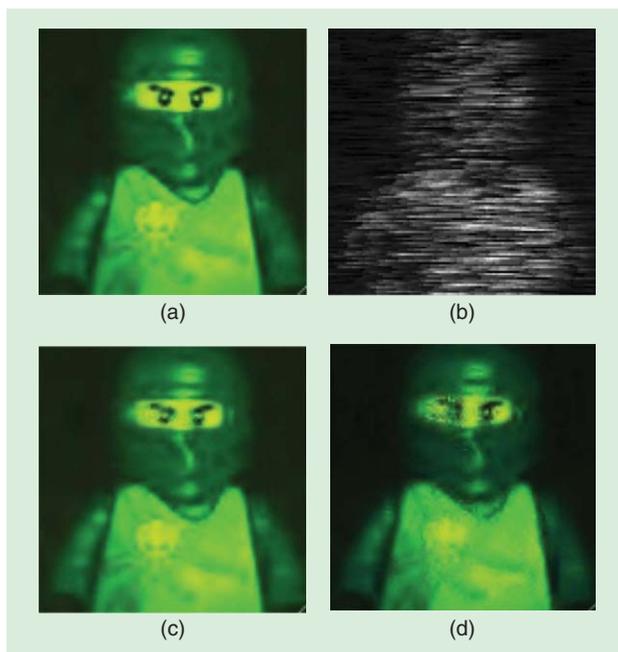


[FIG5] (a) An optimal ensemble of four 64×64 Boolean-coded apertures. White colored squares indicate $T_{jk} = 1$, and black colored squares indicate $T_{jk} = 0$. Zoomed-in areas show that each spatial coordinate in the ensemble contains only one-valued entry. (b) The elements $T_{l_1(l_2-j_1)}^k$ and $T_{l_1(l_2-j_2)}^k$ for $k = 0, \dots, K - 1$ are multiplied and the products are then added to obtain one realization of the random variable z in (5).

$$\sum_{j_1=0}^{N-1} \sum_{j_2=0}^{N-1} \sum_{r_1=0}^{L-1} \sum_{r_2=0}^{L-1} z_{(j_2-r_1)(j_2-r_2)}^{j_1} \Psi_{l_{j_1 j_2 r_1}, \Omega_{k_1}} \Psi_{l_{j_1 j_2 r_2}, \Omega_{k_2}} \quad (6)$$

for $k_1, k_2 = 0, \dots, |\mathcal{T}| - 1$, where Ω_{k_1} and $\Omega_{k_2} \in \mathcal{T}$. Note that the coded aperture entries determine $z_{(j_2-r_1)(j_2-r_2)}^{j_1}$ in (6), which in turn determine the statistical properties of the entries $(A_{|\mathcal{T}||\mathcal{T}|})_{k_1, k_2}$. It was also shown in [20] that the entries in (6) can be roughly approximated as independent and identically distributed Gaussian variables $(A_{|\mathcal{T}||\mathcal{T}|})_{k_1, k_2} \simeq \mathcal{N}(0, C_1(\sigma^2 + m^2))$ such that the concentration of measure for this type of matrices can be applied [21]. A relation between the statistics of (6) and the probability of correct reconstruction is then established using the concentration of measure. The variance of the entries in the Gaussian approximation depends on the variance σ^2 and mean m of the variables z_{l_1, l_2}^i . To guarantee that the diagonal elements of $A_{|\mathcal{T}||\mathcal{T}|}$ satisfy $E((A_{|\mathcal{T}||\mathcal{T}|})_{k, k}) = 1$ for all k , these are normalized by constraining the coded apertures to satisfy $\sum_{i=0}^{K-1} (T_{j_\ell}^i)^2 = C$ for all j and ℓ , where C is a selectable constant. The so-called *Boolean-coded apertures* are defined as those whose entries satisfy $T_{j_\ell}^i \in \{0, 1\}$. In which case, $m = C^2/K^2$ and $\sigma^2 = (K - C)^2 / (K^2(K - 1))$. The term $m^2 + \sigma^2$ can be minimized by setting $C = 1$ implying that in each spatial position, only one coded aperture from the ensemble of K -codes contains a nonzero value. Thus, for each j and ℓ the optimal coded aperture entries under the criterion (4) are obtained by satisfying the constraints $T_{j_\ell}^{i_0} = 1$ and $T_{j_\ell}^i = 0$ for $i \neq i_0$.

Figure 5(a) illustrates an optimal Boolean ensemble for four coded apertures of size 64×64 . As indicated in the zoomed-in regions of the coded apertures, there is only one nonzero element for each j, ℓ position of the ensemble. Figure 6(a)



[FIG6] (a) The original data cube and (b) compressive FPA measurement. Reconstructions from six shots using (c) Boolean (PSNR 40.41 dB) and (d) random (PSNR 27.72 dB). The resulting spectral data cubes are shown as they would be viewed by a Stingray F-033C CCD color camera. (Image courtesy of G.R. Arce and H. Arguello.)

illustrates a portion of the ground truth of a 16-band spectral data cube sensed with a monochromator. Figure 6(b) illustrates a corresponding compressive measurement. Figure 6(c) shows the reconstruction of a data cube of 16 bands using the optimal Boolean-coded apertures. Figure 6(d) depicts the reconstruction using an ensemble with random entries. Notice in Figure 6 that the resulting spectral data cubes are shown as they would be viewed by a Stingray F-033C CCD color camera.

RECONSTRUCTION FROM COMPRESSED MEASUREMENTS

Several numerical algorithms are available to solve the inverse problem in (3). These can be grouped into one of five computational approaches [10]. Algorithms based on greedy pursuit iteratively find an estimate of the solution by selecting atoms of a dictionary and the correspondent weighting factors such that the signal can be represented as a linear combination of these vectors. This approach is implemented by algorithms such as the orthogonal matching pursuit (OMP) and compressive sampling matching pursuit (CoSaMP). The second type of algorithms solve a convex optimization problem. This includes interior-point methods such as ℓ_1 -magic software and gradient-descent methods like the sparse reconstruction via separable approximation (SpaRSA), the two-step iterative shrinkage/thresholding algorithm (TwIST), and the gradient projections for sparse reconstruction algorithm (GPSR) [10], [22], [23]. The third approach uses a Bayesian framework that finds a maximum a posteriori estimator assuming a prior distribution of the unknown coefficients of the signal to recover [8]. Other techniques include

nonconvex optimization [24] and brute force, which attempts to find the solution by trying all possible support sets. A tutorial review of the algorithms in each of these class-types and their associated complexity is found in [10]. Typical computations performed by these algorithms include matrix pseudoinverses, sparse basis transformations, and vector-matrix multiplications. Given that the underlying signals are high dimensional, these calculations require a large number of floating point operations. For instance, in each iteration of the GPSR algorithm, approximately $O(KN^4L)$ operations are computed where K is the number of measurement shots, N is the spatial dimensions, and L is the number of spectral channels of the data cube. Reconstructions of large scenes are indeed overwhelming since they can take hours in desktop architectures [13], [16].

The computational burden can be ameliorated by working with separable sensing operators [25] or fast field-programmable gate array and graphics processing unit implementations of the reconstruction algorithms [26]. Instead of relying on hardware solutions, the computational complexity can be also addressed by exploiting the physical properties of the CASSI optical sensing phenomena. In particular, it is possible to reconstruct the underlying 3-D data cube from a set of reconstructions obtained from nonoverlapping FPA windowed measurements. In this manner, the GPSR reconstruction algorithm performs $O(KB^4L)$ operations per iteration on each $B \times B$ windowed measurement with $B^4 \ll N^4$. After the recovery of the set of images, these are tiled together to assemble the complete data cube reconstruction.

Consider a $B \times B$ measurement window $Y_{m,n}^i$ within the FPA detector at the i th shot, as shown in Figure 7, where the energy in the windowed measurements is traced back through the optical system. After the prism, the energy to be collected by the $B \times B$ FPA window is a coded and dispersed square source cube with L spectral bands. If these voxels are traced back before they impinge on the prism and the coded aperture, the voxels no longer form a cube but instead they form an oblique parallelepiped consisting of L spectral bands, with each one shifted one spatial position in the horizontal axis. Figure 7 illustrates how an oblique parallelepiped F_{mn} of the data cube, which is amplitude

modulated by a coded aperture of size $B \times (B + L - 1)$ and spectrally sheared by the prism, results on a $B \times B$ block of measurements at the detector. In other words, the voxels that are sensed in a $B \times B$ area of the detector emanate from an oblique volume in the source and not from a cube. Furthermore, the oblique parallelepiped volume, once it is sheared by the prism, is transformed into a $B \times B$ cube before it impinges onto the detector.

The energy impinging on an adjacent nonoverlapping window at the FPA can be traced back to the source in a similar manner, such that the entire FPA measurement Y^i can be expressed as an ensemble of $B \times B$ nonoverlapping measurement windows as

$$Y^i = \begin{bmatrix} Y_{0,0}^i & Y_{0,1}^i & \cdots & Y_{0,V'-1}^i \\ \vdots & \vdots & \ddots & \vdots \\ Y_{N'-1,0}^i & Y_{N'-1,1}^i & \cdots & Y_{N'-1,V'-1}^i \end{bmatrix}. \quad (7)$$

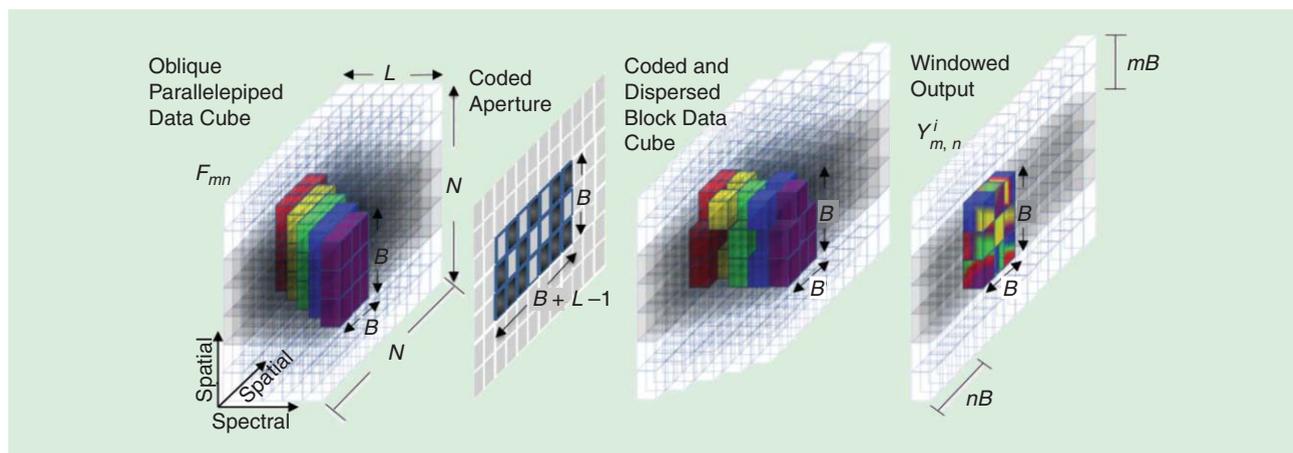
The total number of windows in the set is $N'V'$, with $N' = N/B$ and $V' = \lceil (N + L - 1)/B \rceil$, with B determining the partition block size.

Now consider the reconstruction of the oblique parallelepiped based on its compressive projections $Y_{m,n}^i$. The compressive projection of the parallelepiped block F_{mn} is given by

$$y_{mn}^i = H_{mn}^i f_{mn} + \omega_{mn}^i, \quad (8)$$

where H_{mn}^i is a $B^2 \times B^2L$ submatrix of H^i obtained by choosing the rows and columns that affect each windowed FPA measurement, y_{mn}^i and f_{mn} are the vectorized representations of $Y_{m,n}^i$ and F_{mn} , respectively. Equation (8) is referred to as the *block-model projection*. The set of windowed measurements from sequential FPA shots can then be assembled as in the CASSI model to obtain $y_{mn} = [(y_{mn}^0)^T, \dots, (y_{mn}^{K-1})^T]^T$ and the correspondent matrices H_{mn}^i are assembled as $H_{mn} = [(H_{mn}^0)^T, \dots, (H_{mn}^{K-1})^T]^T$. The matrices H_{mn}^i preserve the structure of H^i in Figure 4 except that the dimensions are now considerably smaller. The multishot block CASSI model can be rewritten as

$$y_{mn} = H_{mn} f_{mn} + \omega_{mn}. \quad (9)$$



[FIG7] Each $B \times B$ window at the detector results from sensing a $B \times B \times L$ oblique parallelepiped block of the data cube. (Reprinted and used with permission from [13].)

Each individual set of windowed measurements y_{mn} in (9) is now used to recover an oblique parallelepiped \hat{F}_{mn} within the data cube. A vectorized representation of \hat{F}_{mn} is then recovered by solving

$$\hat{f}_{mn} = \Psi' \left(\underset{\theta_{mn}}{\operatorname{argmin}} \|y_{mn} - H_{mn} \Psi' \theta'\|_2 + \tau \|\theta_{mn}\|_1 \right), \quad (10)$$

where θ' is a sparse representation of \hat{f}_{mn} in the basis Ψ' . Notice that the inverse problem in (10) is similar to that in (3) of the traditional approach. The difference lies in that the sparsifying basis Ψ' representing the Kronecker product of a wavelet basis and the cosine basis has smaller dimensions than Ψ since F_{mn} is a smaller section of the data cube F .

The full data cube \hat{F} is assembled by tiling all the reconstructed oblique parallelepipeds. Since the number of operations per iteration in CS reconstruction algorithm grows rapidly with the size of the measurement vector, the computational complexity in the block CASSI model decreases rapidly as the number of partitions increase. In particular, the GPCR in CASSI performs approximately $O(KN^4L)$ operations per iteration to recover a $N \times N \times L$ data cube using K FPA measurements. The block model reduces the number of operations to $O(KN^2B^2L)$ operations. If distinct processors

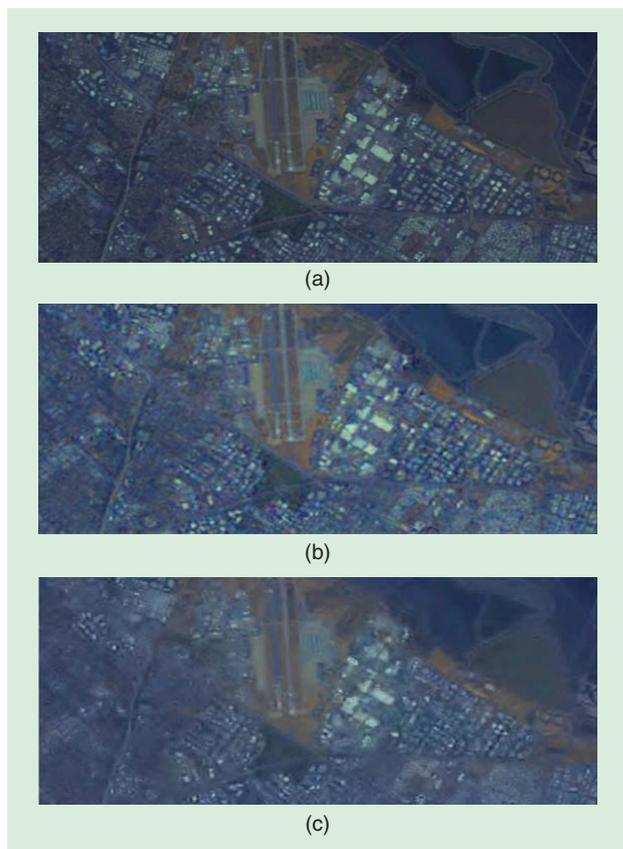
are used to reconstruct separately each of the N'^2 blocks, then the number of operations per core is $O(KB^4L)$.

Figure 8 illustrates the reconstruction quality attained by the block CASSI model. A detailed analysis of the block CS performance in PSNR is presented in [13] where the improvements in image reconstruction are described at length. Figure 8 shows an original $512 \times 512 \times 32$ data cube, the block model reconstructed data cube, and the traditional CASSI model reconstruction. Using an Intel Core i7 3.30 GHz processor PC with 32 GB RAM memory, the block model and the standard CASSI take approximately $0.71K + 3.6$ min and $1.58K + 7.8$ min, respectively, to reconstruct the complete data cube, where K is the number of shots. The reconstruction of one parallelepiped associated with a single block takes $0.045K + 0.28$ min; thus, if multiple processing is available, the block reconstruction can offer significant faster processing.

The block CASSI model is general and can be used with any CS reconstruction algorithm. For instance, the Bayesian reconstruction framework introduced in [27] is noteworthy, with an approximate posterior distribution on model parameters inferred assuming a prior distribution of the unknown coefficients of the signal to recover. In this approach, one seeks to recover not only multiple f_{mn} but to also infer on the underlying dictionary with which the data may be represented. Specifically, we wish to jointly recover all $\{f_{mn}\}$ and Ψ' , where the latter is an over-complete dictionary. It is assumed that $f_{mn} = \Psi' \theta_{mn} + \omega_{mn}$, where $\Psi' \in \mathbb{R}^{B^2L \times U}$ with $U > B^2L$; θ_{mn} is sparse, and ω_{mn} again represents noise. Each measurement is of the form $y_{mn} = H_{mn} \Psi' \theta_{mn} + \omega_{mn}$. The theoretical underpinnings are developed in [28], where a number of illustrative experiments are given. A key distinction of this approach with conventional CS reconstruction algorithms is that Ψ' and θ_{mn} are estimated simultaneously, implying that the measurements are “blind” to the underlying Ψ in which each f_{mn} may be sparsely rendered. This is achievable because $N'V'$ different signals y_{mn} are jointly processed and analyzed. This framework has been coined as *blind* CS [28]. In this framework, a prior is placed on the noise variance, and this is inferred within the analysis. The noise statistics are assumed Gaussian within the prior, but the posterior may differ from the Gaussian assumption. The dictionary learning employed here represents each patch of data in terms of a sparse subset of dictionary elements. One may show that this model has close similarities to Gaussian mixture modeling (GMM) with (near) low-rank covariance matrices [29], with this in turn closely related to recent work on low-rank data modeling [30]. Low-rank models assume the data live in a single linear subspace, with the GMM assume the data live in a union of linear subspaces [29].

COMPUTATIONAL SPECTRAL IMAGING

The coded aperture patterns determine the quality of CASSI measurements. Good codes provide better measurements that, in turn, render more accurate signal reconstructions. Computational spectral imaging goes a step further by jointly optimizing the coded apertures and the computational modules to produce new types of imagery that could benefit vision in



[FIG8] (a) The original RGB and zoomed-in version of the $512 \times 512 \times 32$ data cube. Reconstructions for 10 FPA measurement shots using (b) the block approach with block size $B = 64$, 31.84 dB and (c) the traditional reconstruction approach, 30.99 dB. Block overlapped was used in this example [13].

different applications. This tutorial describes computational algorithms for spectral selectivity and super resolution. Computational tools exist for a number of other applications as detailed in the recent literature [1], [31], [36], [37] and in the companion article [12] in this issue.

SPECTRAL SELECTIVITY

It is often known a priori that objects of interest in a scene do not contain a spectral component over the full range of wavelengths but in a subset of wavelengths. Notably, the coded aperture patterns can be designed so as to maximize the information content on a prespecified subset of spectral bands of particular interest. Spectral selectivity is of interest in many applications, including wide-area airborne surveillance, remote sensing, and tissue spectroscopy in medicine. The optimal spectral bands in airborne surveillance, for instance, depend on atmospheric conditions, time of day, the targets of interest, and the background against which the targets are viewed [4], [15]. Efforts placed on acquiring the entire spectral image cube, to then throw away a large portion of this data is wasteful in many regards.

To this end, coded apertures can be used to attain spectrally selective compressive measurements [4], [15]. A spectrally selective coded aperture, denoted as \mathbf{T}^i , is formed from two concatenated coded apertures, which simplifies the optimization. The first coded aperture \mathbf{w}^i is a structured code optimized to attain the spectral band selectivity. The second coded aperture \mathbf{r}^i is a pseudorandom binary code necessary to attain randomized measurements in CASSI. The coded aperture used in each measurement is obtained by the Hadamard product $\mathbf{T}^i = \mathbf{w}^i \circ \mathbf{r}^i$, and its optimization is divided in two parts. Since row measurements in CASSI are independently coded, the model in (2) is first rewritten as $\mathbf{y} = \sum_{k=0}^{V-1} \mathbf{H}_k \mathbf{C} (\mathbf{\Theta}^T)^{k+1} \mathbf{w}$, where \mathbf{y} is the compressive measurement and \mathbf{H}_k representing a row of the data cube \mathbf{F} and the prism effect, \mathbf{C} is an antidiagonal matrix with all the non-zero entries equal to one, and $\mathbf{\Theta}$ is a permutation matrix [4]. Note that the term $\mathbf{C} (\mathbf{\Theta}^T)^{k+1} \mathbf{w}$ represents the effects of the coded aperture operations on the sheared data cube \mathbf{H}_k . Since the desired spectral bands of interest are known a priori, an approach to design the coded apertures is to create a desired compressive measurement \mathbf{d} where the data cube is first modified such that the spectral bands that are of no interest are zeroed out. A compressive measurement based on the modified data cube would be ideal in the sense that it would only contain information from the spectral bands of interest. The desired compressive measurement is then obtained as $\mathbf{d} = \sum_{k=0}^{V-1} \mathbf{H}_k \boldsymbol{\lambda}$ where the entries of $\boldsymbol{\lambda}$ are $\lambda_j = 1$ if the j th band is desired and zero elsewhere. The desired compressive measurement \mathbf{d} is next equated to the compressive measurement \mathbf{y} attained with the full data cube sensed by the CASSI imager leading to the following structures for the coded apertures

$$\mathbf{w}^j = (\mathbf{\Theta})^{j+1} \mathbf{C}^{-1} \boldsymbol{\lambda} \quad j = 0, \dots, L-1. \quad (11)$$

The solution in (11) requires L vectors, however, the number of vectors can be reduced by exploiting their interdependence. More

specifically, to estimate $K < L$ linearly independent weight vectors, the coded apertures are arranged into the matrix $\mathbf{M}_w = [\mathbf{w}^0, \dots, \mathbf{w}^{L-1}]$. The minimum number of shots K for a given set of bands of interest is the number of independent column K of \mathbf{M}_w determined by $K = \text{rank}(\mathbf{M}_w)$, the rank of the matrix \mathbf{M}_w [4]. The K linear independent columns of \mathbf{M}_w are the linear independent weight vectors $\mathbf{W} = [\mathbf{w}^{\alpha_0}, \dots, \mathbf{w}^{\alpha_{K-1}}]$ selected from \mathbf{M}_w . The remaining $L - K$ columns of \mathbf{M}_w can be estimated using the ensemble of vectors in \mathbf{W} .

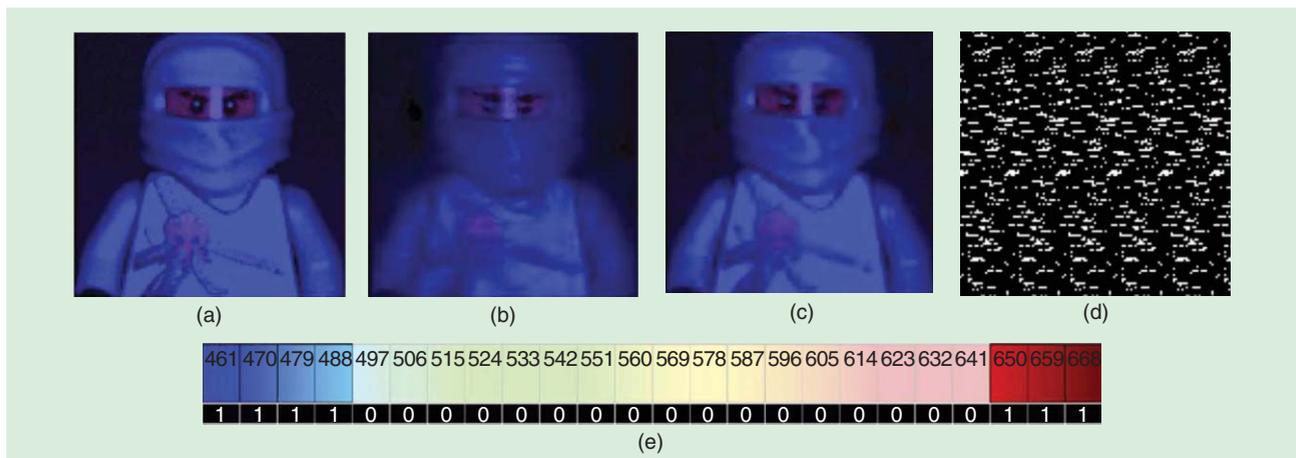
Once the components \mathbf{w}^i of the coded apertures have been optimized, these are fixed and the companion pseudorandom components \mathbf{r}^i are then optimized so as to further minimize the required number of shots. This is achieved by minimizing the rank of the matrix $\mathbf{M}_t = [\mathbf{r}^0 \dots \mathbf{r}^{K-1}] \circ \mathbf{W}$ such that $\text{rank}(\mathbf{M}_t) < \text{rank}(\mathbf{M}_w)$, where \circ is the element by element multiplication operator and where the search of the vectors \mathbf{r}^i is such that the CASSI measurements better satisfy the RIP. The rank minimization of \mathbf{M}_t is an NP hard problem such that a stochastic algorithm can be used to approximately solve this optimization [4].

To illustrate the design of coded apertures with spectral selectivity, consider again the data cube \mathcal{F} with 512×512 pixels of spatial resolution and $L = 24$ spectral bands ranging from 460 nm to 668 nm. The desired spectral bands are set to $\boldsymbol{\lambda} = [461 \text{ nm} - 479 \text{ nm}, 641 \text{ nm} - 668 \text{ nm}]$ as depicted in Figure 9(e). In this case, the initial rank of the matrix \mathbf{M}_t is 24. Using a stochastic based optimization algorithm [4], the rank of the matrix \mathbf{M}_t is minimized to 12. The resulting spectral data cubes are shown as they would be viewed by a Stingray F-033C CCD color camera: Figure 9(a) shows the original bands, Figure 9(b) depicts the reconstruction of the same scene from 12 Bernoulli random coded aperture compressive measurements, Figure 9(c) shows the results when 12 optimized coded apertures are used in the sensing and reconstruction, and Figure 9(d) shows one of the optimal spectrally selective coded apertures. Figure 9 illustrates the gain attained by optimal coded apertures in terms of spectral resolution and higher quality of reconstruction.

SUPERRESOLUTION

While FPAs are available across the IR spectrum, there is continued interest in the development of larger format FPAs for increased resolution [32], [33]. Increasing the size and resolution of FPAs comes with ever-increasing costs. Notably, coded apertures can be designed to yield superresolved reconstruction by leveraging computational imaging [34], [35]. The goal is to translate high-resolution scenes into compressed signals measured by low-resolution or small-format detectors. Superresolution can be attained not only spatially but also spectrally, where the number of spectrally resolved image planes is increased.

Let Δ_c and Δ be the coded aperture pitch and FPA pitch, respectively, and let $r = \Delta/\Delta_c$. A critical requirement to achieve superresolution is that $\Delta_c < \Delta$. If $\Delta = \Delta_c$, the resultant spectral imaging system is equivalent to the standard CASSI architecture. On the other hand, when the pitch resolution of the coded aperture is reduced to the extreme where all the elements



[FIG9] The resulting spectral data cubes are shown as they would be viewed by a Stingray F-033C CCD color camera. (a) The original desired bands. Reconstructed images with 12 shots using (b) random codes (26.92 dB) and (c) optimized codes (31.02 dB). (d) An optimal coded aperture is illustrated. (e) Wavelengths of desired bands. (Reprinted and used with permission from [4].)

of the coded aperture are mapped to just one pixel in the detector ($\Delta = N\Delta_c$), the imaging system becomes the single pixel camera [12]. Superresolution CASSI is realized when the pitch resolution lies somewhere between these two extremes, i.e., when Δ_c lies in the interval $(\Delta/N, \Delta)$. It is important to observe that the minimum value for Δ_c is limited by the diffraction limit of the instrument.

The random projections in superresolution CASSI are given by $y = DH\Psi\theta$, where H is the CASSI measurement matrix shown in Figure 4, and D is a decimation matrix with size $(N(N+L-1))/(r) \times N(N+L-1)$, where r represents the decimation ratio induced by the low-resolution FPA [35]. The decimation matrix D accounts for the integration of light in the detector when pixel mismatch $\Delta_c < \Delta$ is introduced. A set of K low-resolution FPA measurements are first captured, each one having $N' \times M'$ compressed measurements, with $N' = N/\Delta$ and $M' = [(M+L-1)/\Delta]$ being the low-resolution detector height and width, respectively. Superresolution CASSI allows one to exploit subpixel information of a scene to obtain a high-resolution spectral image from low-resolution measurements. As the

number of FPA measurement increases, the superresolved measurements lead to a rapid increase of image reconstruction quality. The standard CASSI measurements, on the other hand, cannot provide improved performance after a few shots. Figure 10 illustrates this concept for $r = 4$, where the CASSI image reconstruction quality after 48 shots (24.95 dB) is compared with the superresolved CASSI, which attains over a 3 dB gain in PSNR by exploiting subpixel information.

CONCLUSIONS

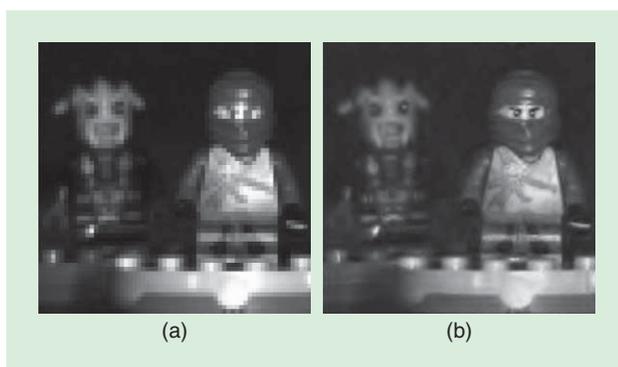
Compressive spectral imaging with coded apertures benefits from the use of simple optical sensing elements to harness compressive projections. CASSI snapshot spectral imagers are both, remarkably simple and surprisingly efficient, making them attractive in a number of applications in remote sensing and surveillance. Their power emerges from the combination of optics with the rich theories of CS and computational imaging. While CASSI spectral imagers naturally embody the congruence of these fields, new spectral imagers as well as more general multidimensional imaging sensors are being discovered by the use of more advanced optical and photonic devices as sensor elements. The potential of coded aperture optimization and optical sensing in multimode and multidimensional imaging holds great promise in the near future, providing fertile ground for signal processing exploration.

ACKNOWLEDGMENTS

This research was supported in part by the Office of Naval Research under contract N00014-10-C-0 199, by the National Science Foundation under grants EECS-0725422 and CCF-0915800, and by the Air Force Office for Scientific Research under grant FA9550-11-1-0194.

AUTHORS

Gonzalo R. Arce (arce@udel.edu) is the Charles Black Evans Professor of electrical and computer engineering at the University of Delaware, Newark. He received the Ph.D. degree in



[FIG10] (a) CASSI reconstruction (PSNR 24.95 dB) versus (b) superresolution CASSI reconstruction (PSNR 29.31 dB) for the sixth spectral band. In both cases 48 shots are used. (Images courtesy G.R. Arce and H. Arguello.)

electrical engineering from Purdue University. He was the 2010 Fulbright–Nokia Distinguished Chair in information and communications technologies. His research interests include statistical signal processing and computational imaging, and he is a coauthor of several books in these areas. He was an associate editor and guest editor of several journals of the IEEE, Optical Society of America (OSA), and International Society for Optics and Photonics (SPIE). He is an IEEE Fellow.

David J. Brady (dbrady@duke.edu) leads the Duke Imaging and Spectroscopy Program (DISP), which builds computational imaging systems. His current DISP projects focus on snapshot gigapixel photography using multiscale optics, X-ray scatter tomography, millimeter wave diffraction tomography, focal tomography, and compressive spectral imaging. He is a Fellow of the IEEE, OSA, and SPIE.

Lawrence Carin (lcarin@ece.duke.edu) earned the Ph.D. degree in electrical engineering at the University of Maryland in 1989. Since 1995, he has been with the Electrical Engineering Department at Duke University, where he is now the William H. Younger Distinguished Professor. He was an associate editor of several journals of the IEEE and the Society for Industrial and Applied Mathematics. He is an IEEE Fellow.

Henry Arguello (henarfu@uis.edu.co) is an associate professor in the Department of Systems Engineering at Universidad Industrial de Santander, Colombia. He received the Ph.D. degree in 2013 from the Department of Electrical and Computer Engineering at the University of Delaware, Newark. His research interests include high-dimensional signal coding and processing, optical imaging, CS, hyperspectral imaging, optical code design, and computational imaging.

David S. Kittle (dsk12@duke.edu) earned the Ph.D. degree in electrical and computer engineering at Duke University in May 2013. He received the B.S. degree in electrical engineering in 2005 from Walla College in College Place, Washington. His research interests lie in computational imaging systems design, testing, and validation.

REFERENCES

- [1] D. J. Brady, *Optical Imaging and Spectroscopy*. Hoboken, NJ: Wiley, 2009.
- [2] M. Eismann, *Hyperspectral Remote Sensing*. Bellingham, WA: SPIE, 2012.
- [3] A. A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, pp. B44–B51, Apr. 2008.
- [4] H. Arguello and G. R. Arce, "Rank minimization code aperture design for spectrally selective compressive imaging," *IEEE Trans. Image Processing*, vol. 22, no. 3, pp. 941–954, 2013.
- [5] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.
- [6] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [7] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [8] J. Shihao, X. Ya, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.
- [9] J. L. Paredes and G. R. Arce, "Compressive sensing signal reconstruction by weighted median regression estimates," *IEEE Trans. Signal Processing*, vol. 59, no. 6, pp. 2585–2601, June 2011.
- [10] J. Tropp and S. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, June 2010.
- [11] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Select. Topics Signal Processing*, vol. 1, no. 4, pp. 586–597, Dec. 2007.
- [12] R. M. Willett, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, "Sparsity and structure in hyperspectral imaging," *IEEE Signal Processing Mag.*, vol. 31, no. 1, pp. 116–126, Jan. 2013.
- [13] H. Arguello, C. V. Correa, and G. R. Arce, "Fast lapped block reconstructions in compressive spectral imaging," *Appl. Opt.*, vol. 52, no. 10, pp. D32–D45, 2013.
- [14] H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, "Higher-order computational model for coded aperture spectral imaging," *Appl. Opt.*, vol. 52, no. 10, pp. D12–D21, 2013.
- [15] H. Arguello and G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 28, no. 11, pp. 2400–2413, Nov. 2011.
- [16] D. Kittle, K. Choi, A. A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Appl. Opt.*, vol. 49, no. 36, pp. 6824–6833, Dec. 2010.
- [17] M. F. Duarte and R. G. Baraniuk, "Kronecker compressive sensing," *IEEE Trans. Image Processing*, vol. 21, no. 2, pp. 494–504, 2012.
- [18] M. Golbabaee and P. Vandergheynst, "Compressed sensing of simultaneous low-rank and joint-sparse matrices," submitted for publication.
- [19] D. S. Kittle, D. L. Marks, and D. J. Brady, "Design and fabrication of an ultraviolet-visible coded aperture snapshot spectral imager," *Opt. Eng.*, vol. 51, no. 7, pp. 071403–1–071403–10, 2012.
- [20] H. Arguello and G. R. Arce, "Restricted isometry property in coded aperture compressive spectral imaging," in *2012 IEEE SSP Workshop*, Aug., pp. 716–719.
- [21] M. Ledoux, *The Concentration of Measure Phenomenon*. Providence, RI: American Mathematical Society, 2001.
- [22] M. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Select. Topics Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.
- [23] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–4.
- [24] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.
- [25] Y. Rivensons and A. Stern, "Compressed imaging with a separable sensing operator," *IEEE Signal Processing Lett.*, vol. 16, no. 6, pp. 449–452, 2009.
- [26] Y. Fang, L. Chen, J. Wu, and B. Huang, "GPU implementation of orthogonal matching pursuit for compressive sensing," *2011 IEEE 17th Int. Conf. Parallel and Distributed Systems (ICPADS)*, pp. 1044–1047.
- [27] A. Rajwade, D. Kittle, T.-H. Tsai, D. Brady, and L. Carin, "Coded hyperspectral imaging and blind compressive sensing," *SIAM J. Imaging Sci.*, vol. 6, no. 2, pp. 782–812, 2013.
- [28] J. Silva, M. Chen, Y. Eldar, G. Sapiro, and L. Carin, "Blind compressed sensing over a structured union of subspaces," *IEEE Trans. Signal Processing*, submitted for publication.
- [29] J. M. Duarte-Carvajalino, G. Yu, L. Carin, and G. Sapiro, "Task-driven adaptive statistical compressive sensing of gaussian mixture models," *IEEE Trans. Signal Processing*, vol. 61, no. 3, pp. 585–600, 2013.
- [30] E. E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
- [31] A. Ramirez, H. Arguello, G. R. Arce, and B. M. Sadler, "Spectral image classification from optimal coded-aperture compressive measurements," *IEEE Trans. Geosci. Remote Sensing*, to be published. DOI: 10.1109/TGRS.2013.2272378
- [32] K. Krapels, R. Driggers, E. Jacobs, S. Burks, and S. Young, "Characteristics of infrared imaging systems that benefit from superresolution reconstruction," *Appl. Opt.*, vol. 46, no. 21, pp. 4594–4603, July 2007.
- [33] A. Rogalski, *Fundamentals of Infrared Detector Technologies*, 2nd ed. Boca Raton, FL: CRC Press, 2011.
- [34] R. Willett, R. Marcia, and J. Nichols, "Compressed sensing for practical optical imaging systems: A tutorial," *Opt. Eng.*, vol. 50, no. 7, pp. 072601–072601-13, July 2011.
- [35] H. Arguello, H. Rueda, and G. R. Arce, "Spatial super-resolution in code aperture spectral imaging," in *SPIE Conf. Defense, Security and Sensing*, Baltimore, MD, Apr. 2012, pp. 83650A–83650A-6.
- [36] A. Ramirez, G. R. Arce, and B. M. Sadler, "Compressive spectral image classification using redundant dictionaries," *IEEE Trans. Geosci. Remote Sensing*, submitted for publication.
- [37] A. Ramirez, G. R. Arce, and B. M. Sadler, "Spectral image unmixing from optical coded-aperture compressive measurements," *IEEE Trans. Geosci. Remote Sensing*, submitted for publication.

[Rebecca M. Willett, Marco F. Duarte, Mark A. Davenport, and Richard G. Baraniuk]

Sparsity and Structure in Hyperspectral Imaging

[Sensing, reconstruction,
and target detection]

Hyperspectral imaging is a powerful technology for remotely inferring the material properties of the objects in a scene of interest. Hyperspectral images consist of spatial maps of light intensity variation across a large number of spectral bands or wavelengths; alternatively, they can be thought of as a measurement of the spectrum of light transmitted or reflected from each spatial location in a scene. Because chemical elements have unique spectral signatures, observing the spectra at a high spatial and spectral resolution provides information about the material

properties of the scene with much more accuracy than is possible with conventional three-color images. As a result, hyperspectral imaging is used in a variety of important applications, including remote sensing, astronomical imaging, and fluorescence microscopy.

While hyperspectral imaging has great potential, acquiring and processing hyperspectral data comes with significant challenges. First, hyperspectral images are extremely high dimensional: in remote sensing applications one routinely encounters images over 1 GB in size. This dimensionality limits our ability to conduct fast and accurate inference (e.g., removing noise or identifying significant spectral signatures). Second, designers of hyperspectral imagers face a myriad of tradeoffs related to photon

Digital Object Identifier 10.1109/MSP.2013.2279507

Date of publication: 5 December 2013

efficiency, acquisition time, dynamic range, and sensor size, weight, power, and cost.

In this article, we review how novel sparse low-dimensional models are enabling sensor designers to tackle many of the above challenges and create new hyperspectral imaging paradigms. We provide an overview the state of the art of hyperspectral image modeling with an emphasis on sparse models that exploit the fact that typical hyperspectral images, while high dimensional, can usually be represented using just a few elements from a basis or dictionary. We also explain how sparse models facilitate the design of novel hyperspectral imaging hardware for remote sensing applications. We pay special attention to cameras based on the compressive sensing (CS) framework that achieve sub-Nyquist measurement rates. We then discuss the imaging design tradeoffs among noise performance, temporal/spatial/spectral resolution, and dynamic range that are afforded by the sensor system, the sparse image model, and noise and quantization errors. Finally, we conclude by describing how the combination of sparse image models and CS architectures can enable fast and accurate target detection.

SPARSE MODELS FOR HYPERSPECTRAL IMAGES

We consider the problem of acquiring a hyperspectral data cube $f \in \mathbb{R}^{d_x \times d_y \times d_\lambda}$, where $f_{i,j,\lambda}$ is the intensity of light in the hyperspectral image at location (i, j) and wavelength λ . For notational simplicity, we also let f denote a vectorized version of the hyperspectral data cube f , which is just a vector in \mathbb{R}^d where $d \triangleq d_x \cdot d_y \cdot d_\lambda$. We model the hyperspectral image acquisition process as $y = Af + w$, where $A \in \mathbb{R}^{n \times d}$ represents the propagation of light through the imaging system, $y \in \mathbb{R}^n$ is a collection of n measurements generated by our imaging system (where n may be less than d), and $w \in \mathbb{R}^n$ is noise.

Due to the significant structure present in hyperspectral data cubes and the linear nature of the aggregation performed by many hyperspectral imagers, low-dimensional signal models for f have received significant attention in the hyperspectral imaging community in a variety of applications, including image compression, denoising, and processing. Most models operate over a partitioning of the hyperspectral data cube into patches along a subset of the dimensions (spatial or spectral) as shown in Figure 1. Spectrum patches collect the intensities for a single spatial location and all wavelengths; band patches collect the intensities for a single wavelength at all spatial locations; and local patches collect the intensities for small intervals of the three dimensions. Denoting the vectorized versions of the patches by the set $\{f_{(1)}, f_{(2)}, \dots, f_{(l)}\}$, the goal of a low-dimensional signal model is to represent each one of these patches using a small number of degrees of freedom: we search for a representation dictionary D that yields patch representations $\theta_{(i)}$ with a small number of nonzeros so that we can write $f_{(i)} = D\theta_{(i)}$, $i = 1, \dots, l$. Below, we discuss two common choices for the dictionary D .

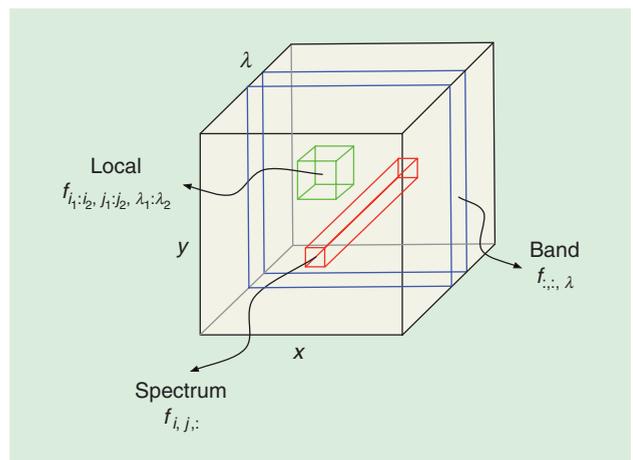
WHILE HYPERSPECTRAL IMAGING HAS GREAT POTENTIAL, ACQUIRING AND PROCESSING HYPERSPECTRAL DATA COMES WITH SIGNIFICANT CHALLENGES.

Principal component analysis (PCA) assumes that the data vectors $f_{(i)}$ lie within or very close to a k -dimensional subspace of \mathbb{R}^p for some $k \ll p$, where p is the patch dimension. In PCA, one computes the empirical cross correlation matrix for the centered data C ; the

top k eigenvalues and corresponding eigenvectors of C are retained so each patch can be accurately represented as a linear combination of these eigenvectors. In practice, the number k is chosen to obtain sufficiently accurate approximations of the patches.

PCA provides an effective and simple way to approximate hyperspectral data. Consider the case in which the image f corresponds to a scene with only a small number $k \ll d_\lambda$ of different types of spectra present across all pixels. In this case, it is clear that the spectral patches $\{f_{:,i,\lambda}\}$ will lie within a k -dimensional subspace of $\mathbb{R}^{d_x d_y}$. PCA has been applied in this manner for hyperspectral image compression [1], classification, segmentation [2], and denoising under Gaussian [3] and Poisson noise models [4], [5]. Furthermore, PCA models can be estimated directly from a sufficiently large number of compressive measurements given enough training data [6].

Sparse signal models are able to capture richer structure than PCA alone. Sparse signal models assume that the data vectors $\{f_{(i)}\}$ lie within (or close to) a union of $\binom{p}{k}$ subspaces of dimension k , where each subspace is spanned by a different choice of k functions from the dictionary D . For instance, these models may rely on a sparsity-inducing orthogonal transform D to obtain coefficient vectors $\theta_{(i)} = D^T f_{(i)}$. In words, the coefficient vector has a small number k of nonzero (or significant) coefficients, and so we can represent the vector $f_{(i)}$ exactly (or approximately) as the linear combination of k components of the transform D . Sparsity models can significantly outperform PCA models in terms of approximation fidelity and are



[FIG1] An illustration of different partitions of a hyperspectral image into patches. The hyperspectral image spans two spatial dimensions (x, y) and one spectral dimension (λ).

SPARSE RECOVERY: METHODS AND GUARANTEES

There are a number of algorithmic approaches to the problem of sparse signal recovery from compressive measurements. We will not provide a complete overview of the possible recovery algorithms here. Instead, we will merely provide a rough outline of what is possible. For further details, we refer the reader to [24] and references therein.

Perhaps the most popular method for sparse recovery is ℓ_1 -norm minimization (also known as basis pursuit or LASSO)

$$\hat{\theta} = \arg \min_{\theta} \|\theta\|_1 \text{ subject to } \|y - A\theta\|_2 \leq \epsilon, \quad (S1)$$

where $\|\theta\|_1 = \sum_i |\theta_i|$ denotes the sum of the magnitudes of the entries of θ and ϵ denotes the tolerable approximation distortion. In addition to (S1), there are also a variety of greedy or iterative strategies, including state-of-the-art methods like compressive sampling matching pursuit (CoSaMP) or iterative hard thresholding (IHT) [26], that treat the vector $A^T y$ as a rough estimate of f and obtain $\hat{\theta}$ by iteratively identifying likely nonzeros. In general, any standard sparse recovery algorithm can be applied to reconstruct a hyperspectral data cube from compressive measurements. However, as detailed in the main body of the article, physical characteristics of real-world compressive hyperspectral compressive imagers should be considered when selecting and implementing such algorithms.

Together with the development of efficient sparse recovery algorithms, there has also been significant recent progress on conditions that ensure that these algorithms obtain provably accurate estimates of the original signal f . One of the more

common assumptions is that the sensing matrix A satisfies the restricted isometry property (RIP), which essentially requires that $\|Af\|_2 \approx \|f\|_2$ for any k -sparse f (i.e., for any f such that we can write $f = D\theta$ where θ has at most k nonzeros). Directly constructing a matrix A that satisfies this property turns out to be rather difficult, but it is possible to show that if we construct A at random, then with high probability it will satisfy the RIP. While a variety of random constructions exist, perhaps the simplest (and most relevant to practical compressive hyperspectral imaging systems) is the so-called “Rademacher ensemble,” where each entry of A is set to be either $1/\sqrt{n}$ or $-1/\sqrt{n}$ with equal probability. Constructing A in this fashion will, with high probability, lead to a matrix satisfying the RIP, provided that $n = O(k \log(d/k))$ [27]. Given such an A and measurements $y = Af + w$, both the approach in (S1) as well as methods like CoSaMP and IHT satisfy a performance guarantee of the form

$$\|\theta - \hat{\theta}\|_2 \leq C_1 \|w\|_2 + C_2 \frac{\|\theta - \theta_k\|_1}{\sqrt{k}}, \quad (S2)$$

where θ_k is the best possible k -sparse approximation to the original θ and C_1, C_2 are absolute constants. From $\hat{\theta}$, we can then obtain the estimate $\hat{f} = D\hat{\theta}$, and when D is an orthonormal basis we can translate this guarantee on $\hat{\theta}$ into one on \hat{f} . Further discussion regarding what can be proven for more specific noise models and in the specific context of compressive hyperspectral imagers is provided in the main body of the article.

predominant in processing and compression of natural images. Examples of sparsity-inducing transforms include the discrete cosine and wavelet transforms. Such transforms can be applied straightforwardly to band patches, as they correspond to intensity images for different light wavelengths.

An additional contribution from the sparsity literature is the application of dictionary-learning algorithms to hyperspectral imaging [7]. These methods use a training data set of image patches to learn a dictionary D , which yields sparse (albeit high-dimensional) representations. However, in contrast to the transformations discussed earlier, the dictionaries learned here do not have orthogonal elements and require the application of custom algorithms for sparse approximation, described in “Sparse Recovery: Methods and Guarantees.” In recent years, sparsity has also been studied in contexts where the types of spectra (called *endmembers*) are known a priori and that each particular pixel is a linear combination of only a few of the endmembers [7], [8]. The sparse representation of the spectrum effectively identifies the component endmembers and their concentrations at each pixel, a process referred to as *hyperspectral unmixing* [8].

Various global sparsifying transforms, to be applied to the entire image rather than its patches, have also been proposed [9]–[11]. Unfortunately, the corresponding increase in dimensionality also increases the computational complexity of the transformation and approximation; furthermore, the improvements in

approximation error are often not found to be significant enough to warrant the additional computational load. Nonetheless, it is possible to formulate global transformations with higher computational efficiency using combinations of patch transformations; a common example is to select a spectrum patch transform D_λ and a band patch transform $D_{x,y}$ and combine them using a Kronecker product $D = D_{x,y} \otimes D_\lambda$ [11]–[13]. PCA models for spectral patches can also be integrated with sparsity models for band patches through the use of Kronecker product matrices [11], [12].

SPARSE MODELS AND HYPERSPECTRAL IMAGERS

CS is the design of signal acquisition strategies that leverage sparse and low-dimensional models such as those described above to ensure accurate signal reconstruction or target detection with relatively few samples. The CS framework has received significant attention in the remote sensing community due to the complexities in hyperspectral imaging hardware designs, the high dimensionality of hyperspectral data sets, and the significant degree of structure and redundancy present in hyperspectral images. In this section, we review baseline designs for hyperspectral imagers and describe several approaches for hyperspectral imaging.

Conventional hyperspectral imagers must address a fundamental design problem: the transformation of a three-dimensional (3-D) signal (in the spatial and spectral domain) into measurements obtained by optical sensing hardware, which is

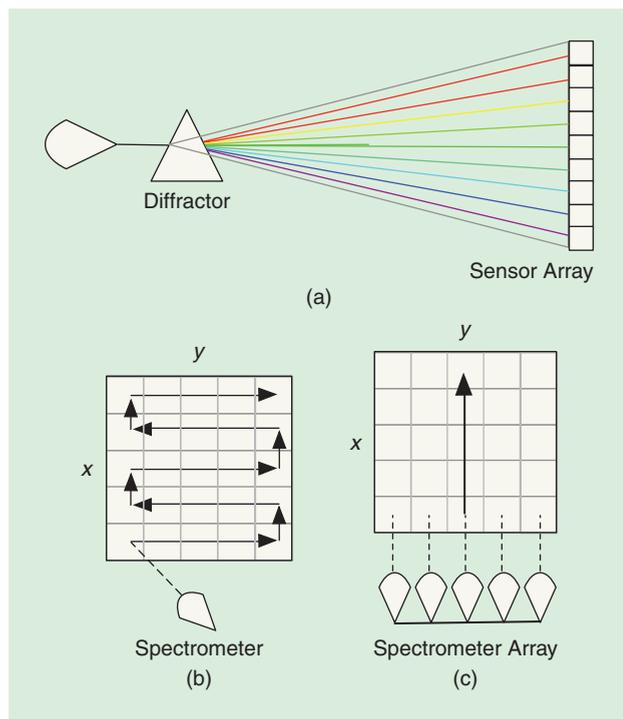
limited to two spatial dimensions. Thus, to design a hyperspectral imager, one must establish a method to record this 3-D data using sensors that do not cover all three dimensions. For example, one can use a one-dimensional (1-D) sensor (i.e., a spectrometer) to obtain a stream of spectrum measurements from the hyperspectral image one pixel at a time. Alternatively, one can use a two-dimensional (2-D) imaging sensor array to capture a single cut or section of the data cube at a time, which could be either a single band or an array of spectra for a single row/column of the spatial dimensions. In this section, we describe the most common designs of hyper spectral imagers as illustrated in Figure 2.

Whiskbroom designs feature optics that focus on a specific spatial location and record either a sequence in time of voxel spectral measurements (using a tunable filter and a single sensor) or an array of samples of the spectra (using a diffraction grating and a linear sensor array). The optical components in whiskbroom designs select a single pixel/spatial location at a time. Whiskbroom designs require a raster scan across the entire field of view and have higher capture latency than other designs; their dwell time on each specific pixel is reduced in comparison with other architectures with matching latency.

Pushbroom designs feature optics that focus along one of the two spatial dimensions (using slit apertures, in comparison with pinhole apertures used by whiskbroom designs) and record a 2-D array of voxels corresponding to a spectral/1-D-spatial cut of the hyperspectral image (using a diffraction grating and a 2-D sensor array). The optical components are usually translated along one spatial dimension to scan the field of view. Although the latency of pushbroom designs is lower than that for whiskbroom designs and their mechanical complexity is comparatively lower, both types of imagers introduce motion in the optics that can result in spatial distortion.

Framing or staring designs feature optics similar to standard imaging cameras that capture 2-D images with additional optics that focus on a single wavelength or band of wavelengths using tunable filters. Their spatial resolution matches that of the sensor array, while spectral resolution is dependent on the tunable filter and latency requirements. The overall design of a staring camera is much simpler than its pushbroom and whiskbroom counterparts. However, the latency due to the tuning of the optical filter is often longer than that of a pushbroom design's scanning system. Furthermore, filtering significantly limits the quantity of light captured at the sensor.

Compressive hyperspectral imagers address a common theme in the design descriptions above: the large number of samples in the spectral data cube results either in high acquisition latency or in significant requirements for the size of the sensor array in the imager. Thus, it can be desirable to reduce the number of measurements necessary for acquisition of the hyperspectral image at a target spatial and spectral resolution. Since one of the central goals in CS is to minimize the required number of measurements (see "Sparse Recovery: Methods and Guarantees" for more details), this has naturally led to its application to hyperspectral imaging. In a compressive hyperspectral



[FIG2] Hyperspectral imager architectures. (a) A spectrometer consists of a diffraction element (grating or prism) and a sensor array that records light intensities at a variety of wavelengths. (b) Whiskbroom designs move the spectrometer spatially throughout the image, scanning one location at a time. (c) Pushbroom designs scan the image along a spatial direction using a spectrometer array.

imager, we continue to model the imaging system as $y = Af + w$ where A is an $n \times d$ matrix, but here we will be specifically interested in the case where n is as small as possible (and hopefully $n \ll d$).

In all of the cases below, the reduction in measurements is achieved through the multiplexing of the voxels of the data cube during acquisition through the optical path. The reduction in measurements can potentially translate to a reduction in acquisition latency and corresponding increase in reconstruction latency introduced by nonlinear sparse recovery algorithms.

The single pixel camera [14]–[17], like whiskbroom designs, relies on a single spectrometer. However, the measurements do not focus on a single spatial location; rather, each measurement aggregates the intensities from a randomly selected subset of pixels of the image. Such selection is performed by programming an optical modulator (such as a digital micromirror device) to reflect light from a subset of the pixels into the spectrometer while masking the light reflected from the rest of the pixels away from the spectrometer. Choosing this configuration for the optical modulator effectively causes the measurement at the single sensor at instance i to correspond to the projection of each spectral band $f_{:, \lambda}$ onto a vector $A_{s, i}$, where $A_{s, i}$ is a binary 0/1 pattern encoding the masking sequence applied by the modulator. By stacking the m vectors as rows of a matrix A_s , the resulting measurement matrix can be expressed as the

Kronecker product $A = I \otimes A_s$, where I is the identity matrix; this measurement operator acts separately on each band.

The compression achieved by the single pixel camera can significantly reduce the acquisition latency compared to whiskbroom designs; however, depending on the number of measurements required for recovery (which is dependent on the complexity of the scene), this design may not outperform pushbroom designs in terms of latency. However, the single-picture architecture can be modified to pushbroom or whiskbroom designs in a straightforward fashion [14]. The spatial resolution of this camera design is given by the resolution of the spatial light modulator, while the spectral resolution of this architecture is given by the characteristics of the single spectrometer.

The coded aperture snapshot spectral imager (CASSI) [18] employs a combination of diffraction prisms, coded apertures, and an optical sensor array to perform multiplexing of the voxels in the hyperspectral image. A dispersive element shears the hyperspectral data cube by enacting a distinct spatial translation for the light field at each wavelength; a coded aperture then masks certain pixels (spatial locations) of the sheared data cube, and a second dispersive element reverses the shearing caused by the spatial translation to result in a modified hyperspectral image with masked voxels. This masked data cube is acquired using an optical sensor that effectively flattens the hyperspectral image into a single snapshot. The imager is a completely static, single-shot design, resulting in a mechanically robust and inexpensive system.

The spatial resolution of this design is governed by the sensor array and the coded aperture (which should have matching resolutions), while the spectral resolution is governed by the degree of dispersion and feature size of the coded aperture. A simplified version of CASSI requires only a single dispersive element and captures the sheared data cube but requires the sensor array size to be $d_x \times (d_y + d_\lambda)$ [19]. This linear acquisition system can be effectively represented by a highly structured $d_x(d_y + d_\lambda) \times d_x d_y d_\lambda$ matrix with binary entries. CASSI is discussed in additional detail in a companion article in this issue [20], including coded aperture design and additional hyperspectral image modeling.

Complementary metal-oxide-semiconductor (CMOS)-based CS approaches have recently emerged for optical imaging [11], [21], [22]. In addition to the aforementioned optics-based designs, it is possible to combine these CMOS-based approaches with standard pushbroom or framing designs to reduce the number of measurements taken with respect to the number of voxels. However, the resulting schemes still require each pixel of the image to be acquired by the CMOS device, and so there is no improvement in properties such as latency, resolution, etc. over those of the CMOS device. Existing implementations of compressive optical sensor arrays perform the computation of the required projections using metal-oxide-semiconductor electronics and are based on random convolution [22], separable transformations [21], block-based transforms [22], structured incoherent transforms like noiselets [11], and randomized integration via Sigma-Delta

analog-to-digital converters (ADCs) [23]. The resulting measurement matrices are expressed in terms of a Kronecker product $I \otimes A_{\text{CMOS}}$, where A_{CMOS} denotes the measurement operator implemented by the CMOS design and the Kronecker product represents the replication of the measurement process among the snapshots required by the particular camera design (e.g., across spectral bands for a staring camera or across shifts in a spatial dimension for a pushbroom camera).

PERFORMANCE LIMITS AND TRADEOFFS FOR RECONSTRUCTING HYPERSPECTRAL IMAGES FROM COMPRESSIVE MEASUREMENTS

The compressive hyperspectral imagers described above enable a range of design tradeoffs among noise performance, temporal/spatial/spectral resolution, and dynamic range. These tradeoffs take different forms depending upon what assumptions we can reasonably make about the sensing matrix A , the sparse or low-dimensional structure of the hyperspectral image f , and the distribution of the noise w . We will first consider the classical CS setting with white Gaussian noise, and then discuss effects such as nonnegativity, quantization, and photon-counting noise.

LIMITS OF CS RECOVERY IN GAUSSIAN NOISE

We begin with the simple observation model $y = Af + w$ but where the noise w , instead of being arbitrary, is independent and identically distributed (i.i.d.) Gaussian with mean zero and variance σ^2 . This leads to slightly different results than those described in “Sparse Recovery: Methods and Guarantees.” Specifically, since the noise w is now random, we consider the expected recovery error. While we could directly apply (S2) and replace $\|w\|_2$ with $\mathbb{E}\|w\|_2 = \sqrt{n}\sigma$, it is possible to get a somewhat tighter result (that does not increase if we take more measurements). In particular, under the assumption that $\|Af\|_2 \approx \beta\|f\|_2$, one can show that most standard sparse recovery algorithms yield an estimate satisfying a guarantee of the form

$$\mathbb{E}\|\theta - \hat{\theta}\|_2 \leq C_1 \sqrt{\frac{k \log d}{\beta}} \sigma + C_2 \frac{\|\theta - \theta_k\|_1}{\sqrt{k}}, \quad (1)$$

where C_1 and C_2 are absolute constants. Note that we have replaced the standard RIP assumption (that $\|Af\|_2 \approx \|f\|_2$) with the more relaxed assumption that $\|Af\|_2 \approx \beta\|f\|_2$ for some constant β , which is equivalent to saying that A/β satisfies the RIP. This can be quite useful since the RIP induces a particular scaling of the matrix A (unit-norm columns), while other scalings of A may be more natural in practice. Naturally, either an increase in β or a decrease in σ (which are essentially equivalent) leads to improved estimation of θ .

One might wonder whether the first term in (1), which represents the impact of the noise w on the recovery error, can be substantially improved. It turns out that this dependence is essentially optimal. In fact, one can show that given the freedom to pick any matrix A (not necessarily satisfying the RIP, but with the same energy as above, i.e., $\|A\|_F^2 = \beta d$) and use any recovery procedure, there is no method that can improve on (1)

by more than a constant factor [28]. In other words, when it comes to sensing a sparse signal in the presence of Gaussian noise, standard CS algorithms are operating at the limit of what any system could achieve given a fixed set of nonadaptive, linear measurements (subject to some energy/signal-to-noise ratio (SNR) constraint on the sensing system A). Moreover, at least if we wish to have an error bound that holds for arbitrary sparse f , we cannot substantially improve this situation even if we pick the rows of this sensing matrix A in a sequential or adaptive fashion [29], [30].

While the bulk of the CS literature has focused on the cases of bounded noise, as in (S2), or white Gaussian noise, as in (1), these may not necessarily be the most natural model in the context of hyperspectral imaging. In particular, Gaussian noise is not a particularly realistic model for photon noise, which arises often as limited available light (constrained by the aperture and latency requirements) is spread across a large number of pixels and spectral bands. We will address this more realistic noise model next. But first, we discuss an important difference between the standard CS framework and the problem of compressive hyperspectral imaging that arises due to the fact that our measurements are constrained to be nonnegative.

EFFECTS OF NONNEGATIVE MATRICES AND OBSERVATIONS

Consider the mechanism described in “Sparse Recovery: Methods and Guarantees” for constructing the sensing matrix A , where we set each element of A to be $\pm 1/\sqrt{n}$ with equal probability. Unfortunately, in the context of linear optical imaging, such a sensing matrix cannot be implemented. In particular, we can think of A as describing how light is propagated through a linear optical system, so that $A_{i,j}$ denotes the fraction of the total amount of light from the j th voxel in the hyperspectral image that contributes to the i th measurement. Clearly, the fractions cannot have negative values, so $A_{i,j} \geq 0$. Furthermore, the total amount of light sensed cannot be greater than the amount of light incident upon the system (i.e., photon flux must be preserved); mathematically, this has several consequences. Most generally, this means that if a_j denotes the j th column of A , then we must have $\|a_j\|_1 \leq 1$, since the entries in a_j correspond to how the light from voxel f_j is distributed across the detector array. This constraint ensures that the total photon flux is preserved, i.e., $\|Af\|_1 \leq \|f\|_1$ for all f (where f , denoting the intensity of light at different locations and wavelengths, also consists solely of nonnegative elements). In some imaging systems, there are additional constraints on the entries $A_{i,j}$. For instance, in the single pixel camera architecture, if we assume that each measurement is allocated an equal amount of time, then the maximum possible value for $A_{i,j}$ is $1/n$ (since

WHILE THE BULK OF THE COMPRESSIVE SENSING LITERATURE HAS FOCUSED ON THE CASES OF BOUNDED OR GAUSSIAN NOISE, THESE MAY NOT NECESSARILY BE THE MOST NATURAL MODELS IN THE CONTEXT OF HYPERSPECTRAL IMAGING.

only $1/n$ of the total amount of light is available during each measurement period).

These restrictions lead to a small gap between the hyperspectral imaging setting and the standard theoretical treatment of CS. While it is possible to develop a specially tailored theory for certain classes of matrices with nonnegative entries, and ultimately obtain

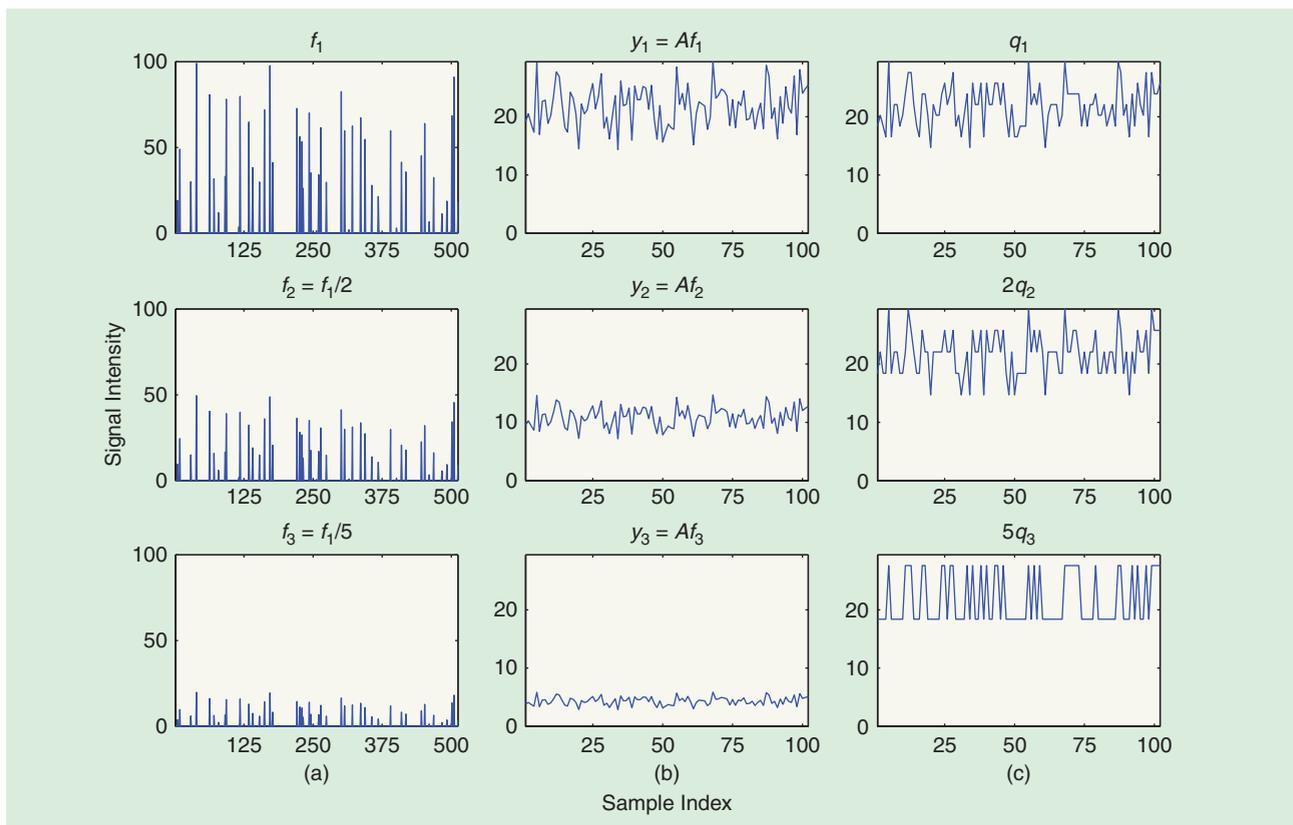
bounds similar to (S2) or (1), it is perhaps more instructive to consider how to relate the desired RIP matrix A with $\pm 1/\sqrt{n}$ entries to a physically realizable matrix \tilde{A} with entries of zero or $1/n$ (with equal probability). Specifically, one can imagine constructing \tilde{A} by adding $1/\sqrt{n}$ to each element of A to make each element either zero or $2/\sqrt{n}$, and then rescaling by $1/(2\sqrt{n})$ to obtain a matrix with entries of zero or $1/n$. In the i.i.d. Gaussian measurement noise model from above, the impact of this shifting and renormalization is that we can write our measurements as

$$y = \tilde{A}f + w = \frac{Af}{2\sqrt{n}} + \frac{\|f\|_1}{2n} + w; \quad (2)$$

that is, we observe a scaled version of what we would ideally like to measure (Af) plus a constant offset proportional to the total amount of light in the scene. The constant offset introduces some unique and nontrivial challenges. As we describe below, it has a significant impact on the noise variance in photon-limited settings. However, even in photon-rich settings, where we may accurately adopt a Gaussian noise assumption, the constant offset may cause challenges.

First, consider recovering f from y using the standard sparse recovery methods described in “Sparse Recovery: Methods and Guarantees.” The nonnegativity of A can lead to some important algorithmic challenges when the recovery algorithm has been specifically designed under the assumption that A satisfies the RIP. In particular, one of the consequences of the RIP is that $A^T A$ acts like an isometry when applied to sparse vectors. This fact is explicitly exploited by greedy algorithms that make decisions based on $A^T y$, and sometimes implicitly exploited by some ℓ_1 -minimization solvers to speed convergence. Unfortunately, this is no longer the case when the entries of A are nonnegative, since in this case all the columns of A are correlated with each other. For the algorithms that rely on this fact, simply plugging y and A into the algorithm without any modifications will yield inaccurate reconstructions and/or slow convergence.

Fortunately, in many cases it is possible to sidestep this issue. For example, in the context of (2), if we can use the data y to accurately estimate $\|f\|_1$ (or can directly obtain an estimate of this value in advance), then we can set $y' = y - \|f\|_1/(2n)$ and then feed y' and $A/(2\sqrt{n})$ into standard sparse recovery methods. This fix can significantly improve the speed and accuracy of reconstruction (although this



[FIG3] An illustration of dynamic range and quantization challenges in compressive hyperspectral imaging. In all plots, the horizontal axis is the sample index and the vertical axis is the signal intensity. (a) The depiction of the same sparse signal at three different intensity levels (brightnesses). (b) The depiction of unquantized compressive measurements of the signals on the left using the sensing matrix construction in (2). (c) The quantized measurements, rescaled for easy visual comparison. We apply the same 4-bit uniform quantizer, designed to quantize values between zero and 30, to each set of measurements. Clearly designing a quantizer capable of quantizing measurements from bright sources limits the accuracy of quantized measurements at lower intensities.

approach can have significant noise implications in the low-light regime; see [31] for details).

Alternatively, it is also often relatively straightforward to modify the algorithm to rely less heavily on the RIP assumption. For example, greedy algorithms can be modified by replacing A^T with the pseudoinverse $A^\dagger = A^T(AA^T)^{-1}$. More generally, this can be viewed as a special case of preconditioning the data y , which is shown to significantly improve reconstruction accuracy [32], [33]. Note that it is also possible to modify standard sparse recovery methods to enforce nonnegativity in f as well [34].

EFFECTS OF QUANTIZATION AND DYNAMIC RANGE

A more significant challenge posed by nonnegativity arises due to the fact that physical systems must ultimately also quantize the measurements y . Typically, a quantizer will have a fixed number of quantization levels arranged to cover the entire range of different values that elements of y may take. When this range is precisely known in advance, each quantization level corresponds to a small interval of different values, yielding accurate measurements. In the context of the model in (2), however, note that we are actually trying to quantize small fluctuations (determined by Af) around a constant offset

(determined by $\|f\|_1$) that will, in general, be unknown a priori. This poses a challenge when using a traditional quantizer since, if the range of the quantizer is set to be too small, the elements of y may fall outside the range of the quantizer, but if the range is too large, the small fluctuations determined by Af will fail to use the full quantization range and the system will lose precision. This is especially problematic when using a quantizer with low bit depth. Thus, in the context of compressive hyperspectral imaging, quantization noise can be a significant source of error. A toy example illustrating this effect is presented in Figure 3, which demonstrates the challenge associated with designing a single mechanism for uniformly quantizing a signal with an unknown intensity or brightness, and hence an unknown constant offset.

We would like to be able to address this challenge in an automatic fashion, without resorting to manual tuning of the quantizer range for each scene of interest. One approach is to simply use very high bit-depth quantizers, but this can be costly and ultimately fails to fully address the challenge for broad ranges of brightnesses. A more robust approach is to compensate for the offset in hardware before quantizing [23], [35]. While this requires specialized sensor circuitry and can be somewhat costly, when designing a system that will be used

to image scenes of widely varying brightness the improvement in performance may be worth this increased cost.

Yet another approach to this problem relies on some of the rather unique properties of randomized measurements. In particular, the randomized measurements typically used in CS are democratic, generally meaning that they each contain roughly the same amount of information, and hence by taking additional measurements we can be robust to having large errors (or even erasures) on a subset of the measurements [36]. This has a number of consequences in the context of quantization. First, while classical systems typically try to set the quantizer range to ensure that saturation occurs with extremely low probability, it has been shown empirically that in CS systems one can obtain improved performance by allowing a nontrivial number of saturation events (e.g., on the order of 5–10%) [36]. Second, it allows for a particularly elegant method for automatically adjusting a quantizer to mitigate the problem described above. In particular, if the measurements are obtained sequentially in time (as in the single pixel camera architecture) then one can perform automatic gain control to dynamically adjust the prequantization gain to ensure that some desired fraction of the measurements saturate the quantizer (on both ends of the quantization range). This approach ensures that the full range of the quantizer is exploited without the need to manually measure the offset in (2), but it has the drawback of requiring a certain amount of “burn-in time” before stabilizing.

Finally, it is worth noting that as long as we can compensate for the unknown constant offset in (2), CS actually has the potential to result in significant gains over noncompressive systems in terms of quantization error and dynamic range. In particular, in a noncompressive system, we typically would quantize each voxel using the same fixed quantization range, but voxel intensity can vary dramatically both spatially and across spectra. This causes saturation and loss of detail in bright and dark regions of the data cube. In contrast, by combining random combinations of voxels into a single measurement, compressive systems dramatically reduce the dynamic range over which the measurements that we must quantize can fluctuate. This has been studied in the context of ADCs in [37] and can be seen by comparing Figure 3(a) and (b). For a given bit depth, this reduced range can allow for reduced quantization error in the compressive case. Exploiting this, along with the fact that by taking fewer measurements in a given time window we can use a lower-rate quantizer with a higher bit depth, there is potential for compressive systems to be more effective at mitigating quantization error than traditional systems.

EFFECT OF PHOTON-COUNTING NOISE

Up to this point, we have considered the impact of noise, non-negativity, and quantization but only when the noise vector w is signal independent. However, in many hyperspectral imaging

ONE OF THE ADVANTAGES OF COMPRESSIVE METHODS FOR HYPERSPECTRAL IMAGING IS THAT THEY ENABLE NEW DESIGN TRADEOFFS.

contexts we are in fact photon limited, so that the total number of photons detected by our system is small relative to the desired resolution. In photon-limited settings, we may model the observations as obeying a Poisson distribution, which has a mean equal to its vari-

ance. This effect introduces serious limitations. In particular, in (2) we saw that the signal of interest was added to a constant offset. Since the mean and variance of Poisson noise are equal, this offset plays a critical role in controlling the noise variance.

Some of the major theoretical challenges associated with the application of CS to linear optical systems in the presence of Poisson noise have been addressed in the recent literature [38], [39]. These works considered two novel sensing paradigms, based on either pseudorandom dense sensing matrices (akin to the shifted and scaled dense sensing matrix described above) or expander graph constructions, both of which satisfy the non-negativity and flux preservation constraints. In these settings, for a fixed signal intensity (i.e., fixed $\|f\|_1$), the error bound actually grows linearly with the number of measurements or sensors, n , since a limited amount of light is spread across an increasing number of detectors, each with a decreasing SNR. In other words, keeping n as small as possible (a central goal in CS) helps maximize SNR and reconstruction accuracy in a way not reflected in conventional CS bounds. Thus, incorporating real-world constraints into the measurement model has a significant impact on the expected performance of a compressive hyperspectral imager, and these constraints should be considered carefully throughout any design process.

EFFECT OF IMPERFECT SYSTEM MODELS

A major challenge in the design of compressive hyperspectral imagers is accurate knowledge of the projection operator A . While we might design a system to have a particular sensing matrix A , calibration errors and optical effects will always introduce inaccuracies. Even if we had the ability to estimate A precisely, there are settings where using an approximation of A has advantages; for instance, when we can approximately compute Af using fast Fourier transforms, conducting sparse recovery is much faster than with a dense matrix representation of A .

When we run a sparse recovery algorithm with an inaccurate sensing matrix A , it corresponds to the observation model $y_i = Af + Ef + w$, where Ef represents the difference between the *true* projections collected by hyperspectral imager and the *assumed* projections in A . The term Ef can be thought of as signal-dependent noise. Analysis of the theoretical ramifications of these kinds of errors allow the designers of spectral imagers to accurately assess tradeoffs between accurate calibration of A and computational efficiency [40].

ADDITIONAL TRADEOFFS

One of the advantages of compressive methods for hyperspectral imaging is that they also enable a range of new design tradeoffs.

For example, the single pixel camera architecture allows us to achieve high spectral resolution while trading off between spatial resolution and latency by adjusting the resolution of the patterns used by the optical modulator (a higher-resolution pattern will also require a larger total number of measurements, increasing spatial resolution at a cost of higher latency). Alternatively, the CASSI system allows for low latency while trading off between spatial and spectral resolution. For all architectures, however, we have a fundamental tradeoff between resolution and the SNR. If we fix the latency (i.e., the total acquisition time, and hence the total amount of light incident upon a hyperspectral imager), then increasing either spatial or spectral resolution means decreasing the amount of light measured for each voxel in the hyperspectral image. As resolution increases, measurements become more photon limited and, therefore, noisy.

HYPERSPECTRAL TARGET DETECTION FROM COMPRESSIVE MEASUREMENTS

In addition to enabling the design of new hyperspectral imaging hardware and acquisition methods, sparsity and other low-dimensional structures provide new ways to efficiently process the data produced by these new sensors, in some cases without ever explicitly estimating the high-dimensional hyperspectral image [41], [42].

In this section, we address the question of whether compressive measurements of hyperspectral images of the form $y = Af + w$ can be used to accurately and efficiently infer whether f belongs to some target class without estimating f directly. As a motivating example, consider the CASSI system discussed earlier: it collects one coded projection of each spectrum in the scene. One projection per spectrum is sufficient for reconstructing spatially homogeneous spectral images, since projections of neighboring locations can be combined to infer each spectrum. Significantly more projections are required for detecting targets of unknown strengths without the benefit of spatial homogeneity. One might ask how several such systems can be used in parallel to reliably detect spectral targets and anomalies from different coded projections.

Hyperspectral imaging introduces several unique target detection challenges. For instance, in remote sensing applications each measured spectrum reflects the mixing of multiple spectra across a relatively large physical area—so that the spectrum of interest may be mixed with other spectra in unknown proportions. A mixed pixel model accounts for such interferences by modeling every spatial location as either a target material corrupted by background, or just background [43]. This background may be modeled using a multivariate Gaussian distribution: $b \sim \mathcal{N}(0, \Sigma_b)$, so that we have mixed observations according to

$$y_m = A(f + b) + w = y + Ab. \quad (3)$$

INCORPORATING REAL-WORLD CONSTRAINTS INTO THE MEASUREMENT MODEL HAS A SIGNIFICANT IMPACT ON THE EXPECTED PERFORMANCE OF A COMPRESSIVE HYPERSPECTRAL IMAGER.

Thus, in the mixed pixel setting our ideal compressive observations are contaminated by Ab , which suggests that the statistics of b must be considered when choosing A .

One approach to this challenge is to apply a prewhitening filter $P \in \mathbb{R}^{n \times n}$ to the mixed observations y_m , with the goal of mitigating the effects of the background

b . The prewhitened observations can be expressed as $z = Py_m = \tilde{A}f + \tilde{w}$, where \tilde{w} is white Gaussian noise with variance one and $\tilde{A} = PA$. This suggests choosing the hyperspectral camera optical design, described by A , in a way that depends on the background covariance Σ_b , so that the product PA facilitates accurate compressive signal classification and detection (e.g., a random $n \times d$ matrix with i.i.d. $\mathcal{N}(0, 1)$ entries, commonly considered in the CS literature) [44]. This approach naturally provides fundamental insight into the robustness of compressive target detection to background contamination.

TARGET DICTIONARIES

The goal of hyperspectral target detection is, in the context of mixed observations, to determine whether $f = 0$ (i.e., no target and only background is present) or which f in a dictionary of target spectral signatures \mathcal{D} corresponds to the observations.

Theoretical performance bounds provide key insight into how error rates scale with the number of measurements collected, the spectral resolution of targets, the amount of background signal present, the SNR, and properties of \mathcal{D} . In particular, let ρ denote the minimum Euclidean distance between any two target spectra in the target class \mathcal{D} , and let $|\mathcal{D}|$ denote the size of the dictionary. Performance can be characterized in terms of a method's positive false discovery rate (pFDR), which measures the fraction of declared targets that are false alarms and is a useful metric in multiple testing scenarios such as this.

A target detection method based on a nearest-neighbor approach applied to prewhitened measurements z yields the bound

$$\text{pFDR} = O\left[\frac{1}{|\mathcal{D}|} \left(\left[1 + \frac{\rho^2}{4n} \right]^{n/2} - \frac{1}{|\mathcal{D}|} \right)^{-1}\right], \quad (4)$$

which decays with the number of measurements n and the size of the target dictionary, but increases with ρ . Thus introducing new candidate targets which are very similar to existing candidate targets can significantly deteriorate performance, regardless of the spectral resolution d . Experimental results show that using these theoretically supported designs of A , which account for background contamination and target dictionary properties, yields significantly better target detection accuracy than simply measuring low-resolution hyperspectral images [44].

TARGET MANIFOLDS

The fixed-dictionary hyperspectral target detection problem formulation above is accurate if the signals in the dictionary are faithful representations of the target signals that we observe. In reality, however, the target signals will differ from those in the dictionary due to the differences in the experimental conditions under which they are collected. For instance, in remote sensing applications, the observed spectrum of a material will not match the reference spectrum observed in a laboratory due to differences in atmospheric and illumination conditions. In this case, one could reasonably model the target signals observed under different experimental conditions as lying in a low-dimensional submanifold of the high-dimensional ambient signal space; this has been shown to be an accurate model for hyperspectral images in [45].

Thus, in many practical settings, rather than differentiate among a finite collection of candidate spectra, we must differentiate among a collection of candidate target manifolds. Target detection in this setting has two key components: 1) a search for the closest point in each candidate target manifold to the observation, followed by 2) a minimum distance-based detection step controlled by the desired false alarm probability. This approach has been dubbed *smashed filtering* in [41]. CS theory and methods yield insights into both these steps. Tradeoffs and performance bounds for the second step are described in the previous section. Furthermore, it is now known that the randomized projections common in CS also preserve the structure of the manifold; this can be shown by adapting the earlier Johnson–Lindenstrauss lemma argument to a sufficiently dense sampling of the manifold [46]. This theoretical result implies that the first step of the smashed filter can be computed directly in the compressive domain.

ANOMALY DETECTION

While in many settings target dictionaries can be formed in a laboratory or using “ground truth” data (usually collected at considerable expense and time), at times target dictionaries are simply unavailable. In such settings, one might be interested in detecting objects not in the dictionary. Here, the target signals of interest are anomalous and are not known a priori to the user. The target detection methods discussed above can be extended to anomaly detection by exploiting the distance preservation property of the sensing matrix A to detect anomalous targets from projection measurements, as detailed in [44], [47], and [48].

CONCLUSIONS AND FUTURE DIRECTIONS

Due to the enormous size of hyperspectral images with high spatial and spectral resolution, approaches that enable efficient data collection, signal reconstruction, and target detection tasks have enormous practical potential. The good news is that typical hyperspectral images have significant structure that can be exploited within the context of sparse models and CS. Armed with such models, we can engineer novel compressive sensors and reconstruction algorithms.

On the surface, the application of the CS theory and algorithms to hyperspectral imaging appears very promising. However, one of the central themes of this article is that these theories and methods cannot be applied blindly to this application arena. For a compressive hyperspectral imaging design to be truly effective, it must account for the physical constraints of the measurements system, use appropriate quantization methods, accommodate realistic noise models (including photon noise, background signal effects, and calibration errors), and use reconstruction algorithms that specifically account for all of these effects. None of these aspects can be considered in isolation, and any system design that ignores these issues has limited potential.

Despite these caveats, researchers are pushing the boundaries of our collective knowledge of how to exploit signal structure for improved sensing and inference. For example, while sequentially selecting the rows of A in an adaptive fashion is of limited benefit in some of the hardest possible sparse recovery problems [29], [30], in high SNR regimes or settings where we have structured or group sparsity (common in hyperspectral imaging), adaptivity can potentially yield significant gains. Exploring the applications of these ideas to practical imaging systems is an important area of ongoing research.

ACKNOWLEDGMENTS

Rebecca Willett was supported by grants NGA HM1582-10-1-0002, NSF CCF-06-43947 and DMS-08-11062, Smithsonian GO1-12098B, and DARPA HR0011-07-1-003. Richard Baraniuk was supported by grants NSF CCF-1117939, DARPA/ONR N66001-11-C-4092 and N66001-11-1-4090, ONR N00014-10-1-0989 and N00014-11-1-0714, and ARO MURI W911NF-09-1-0383.

AUTHORS

Rebecca M. Willett (rmwillett@wisc.edu) is an associate professor in the Electrical and Computer Engineering Department at the University of Wisconsin-Madison. She was an assistant professor and then associate professor of electrical and computer engineering at Duke University between 2005 and 2013. She received the NSF CAREER Award in 2007 and an Air Force Office of Scientific Research Young Investigator Award in 2010.

Marco F. Duarte (mduarte@ecs.umass.edu) is an assistant professor in the Department of Electrical and Computer Engineering at the University of Massachusetts Amherst. His research interests include machine learning, CS, sensor networks, and optical image coding. He received the Presidential Fellowship and the Texas Instruments Distinguished Fellowship in 2004 and the Hershel M. Rich Invention Award in 2007, all from Rice University. He coauthored (with C. Hegde and V. Cevher) the Best Student Paper at the 2009 International Workshop on Signal Processing with Adaptive Sparse Structured Representations.

Mark A. Davenport (mdav@gatech.edu) is an assistant professor with the School of Electrical and Computer Engineering, the Georgia Institute of Technology. His research interests include CS, low-rank matrix recovery, nonlinear approximation,

and the application of low-dimensional signal models in signal processing and machine learning. He received the 2007 Hershel M. Rich Invention Award and the 2010 Ralph Budd Thesis Award, both from Rice University.

Richard G. Baraniuk (richb@rice.edu) is the Victor E. Cameron Professor of Electrical and Computer Engineering at Rice University. His research interests lie in the areas of signal processing, sensors, and machine learning. He was elected a Fellow of the IEEE in 2001 and of the American Association for the Advancement of Science in 2009.

REFERENCES

- [1] S. Lim, K. H. Sohn, and C. Lee, "Principal component analysis for compression of hyperspectral images," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, Sydney, Australia, July 2001.
- [2] C. Rodarmel and J. Shan, "Principal component analysis for hyperspectral image classification," *Surv. Land Inf. Sci.*, vol. 62, no. 2, pp. 115–122, 2002.
- [3] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "A nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Processing*, vol. 22, no. 1, pp. 119–133, 2013.
- [4] J. Salmon, Z. Harmany, C. Deledalle, and R. Willett, "Poisson noise reduction with non-local PCA," *J. Math. Image Vision*, to be published.
- [5] M. Makitalo and A. Foi, "Optimal inversion of the Anscombe transformation in low-count Poisson image denoising," *IEEE Trans. Image Processing*, vol. 20, no. 1, pp. 99–109, 2011.
- [6] J. Fowler, "Compressive-projection principal component analysis," *IEEE Trans. Image Processing*, vol. 18, no. 10, pp. 2230–2242, Oct. 2009.
- [7] A. Charles, B. Olshausen, and C. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE J. Select. Topics Signal Process.*, vol. 5, no. 5, pp. 963–978, 2011.
- [8] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [9] P. Dragotti, G. Poggi, and A. Ragozini, "Compression of multispectral images by three-dimensional SPIHT algorithm," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, no. 1, pp. 416–428, Jan. 2000.
- [10] J. Fowler and J. Rucker, "3-D wavelet-based compression of hyperspectral imagery," in *Hyperspectral Data Exploitation: Theory and Applications*. Hoboken, NJ: Wiley, 2007, pp. 379–407.
- [11] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.
- [12] W. Li, S. Prasad, and J. Fowler, "Integration of spectral-spatial information for hyperspectral image reconstruction from compressive random projections," *IEEE Geosci. Remote Sensing Lett.*, to be published.
- [13] M. Golbabaee, S. Arberet, and P. Vandergheynst, "Compressive source separation: Theory and methods for hyperspectral imaging," *IEEE Trans. Image Processing*, to be published.
- [14] J. Ma, "A single-pixel imaging system for remote sensing by two-step iterative curvelet thresholding," *IEEE Remote Sensing Lett.*, vol. 6, no. 4, pp. 676–680, Oct. 2009.
- [15] T. Sun, D. Takhar, J. Laska, M. Duarte, V. Bansal, R. Baraniuk, and K. Kelly, "Realization of confocal and hyperspectral microscopy via compressive sensing," *Bull. Amer. Phys. Soc.*, vol. 53, no. 2, 2008.
- [16] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single pixel imaging via compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 83–91, 2008.
- [17] Y. August, C. Vachman, Y. Rivenson, and A. Stern, "Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains," *Appl. Opt.*, vol. 52, no. 10, pp. D46–D54, Apr. 2013.
- [18] M. Gehm, R. John, D. Brady, R. Willett, and T. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Opt. Exp.*, vol. 15, no. 21, pp. 14013–14027, 2007.
- [19] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–B51, 2008.
- [20] G. Arce, D. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging," *IEEE Signal Processing Mag.*, vol. 31, no. 1, pp. 105–115, 2014.
- [21] L.-L. Xiao, K. Liu, and D.-P. Han, "CMOS low data rate imaging method based on compressed sensing," *Opt. Laser Technol.*, vol. 44, no. 5, pp. 1338–1345, 2012.
- [22] M. Dadkhah, M. Deen, and S. Shirani, "Block-based compressive sensing in a CMOS image sensor," *IEEE Sensors J.*, to be published.
- [23] Y. Oike and A. El Gamal, "CMOS image sensor with per-column $\Sigma\Delta$ ADC and programmable compressed sensing," *IEEE J. Solid-State Circuits*, vol. 48, no. 1, pp. 318–328, Jan. 2013.
- [24] J. Tropp and S. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, 2010.
- [25] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.
- [26] T. Blumensath and M. Davies, "Iterative hard thresholding for compressive sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.
- [27] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Const. Approx.*, vol. 28, no. 3, pp. 253–263, 2008.
- [28] E. Candès and M. Davenport, "How well can we estimate a sparse vector?" *Appl. Comput. Harmon. Anal.*, vol. 34, no. 2, pp. 317–323, 2013.
- [29] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [30] E. Arias-Castro, E. Candès, and M. Davenport, "On the fundamental limits of adaptive sensing," *IEEE Trans. Inform. Theory*, vol. 59, no. 1, pp. 472–481, 2013.
- [31] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–120, 124, 2007.
- [32] S. Xiong, B. Dai, and P. Z. G. Qian, "Orthogonalizing penalized regression," arXiv:1108.0185, 2011.
- [33] J. Jia and K. Rohe, "Preconditioning to comply with the irrepresentable condition," arXiv:1208.5584, 2012.
- [34] Z. Harmany, D. Thompson, R. Willett, and R. Marcia, "Gradient projection for linearly constrained convex optimization in sparse signal recovery," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Hong Kong, Sept. 2010.
- [35] V. Majidzadeh, L. Jacques, A. Schmid, P. Vandergheynst, and Y. Leblebici, "A (256 × 256) pixel 76.7 mW CMOS imager/compressor based on real-time in-pixel compressive sensing," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Paris, France, May 2010.
- [36] J. Laska, P. Boufounos, M. Davenport, and R. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *Appl. Comput. Harmon. Anal.*, vol. 31, no. 3, pp. 429–443, 2011.
- [37] M. Davenport, J. Laska, J. Treichler, and R. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: Noise folding vs. dynamic range," *IEEE Trans. Signal Processing*, vol. 60, no. 9, pp. 4628–4642, 2012.
- [38] M. Raginsky, R. Willett, Z. Harmany, and R. Marcia, "Compressed sensing performance bounds under Poisson noise," *IEEE Trans. Signal Processing*, vol. 58, no. 8, pp. 3990–4002, 2010.
- [39] M. Raginsky, S. Jafarpour, Z. Harmany, R. Marcia, R. Willett, and R. Calderbank, "Performance bounds for expander-based compressed sensing in Poisson noise," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4139–4153, 2011.
- [40] P.-L. Loh and M. Wainwright, "High-dimensional regression with noisy and missing data: Provable guarantees with nonconvexity," *Ann. Stat.*, vol. 40, no. 3, pp. 1637–1664, 2012.
- [41] M. Davenport, M. Duarte, M. Wakin, J. Laska, D. Takhar, K. Kelly, and R. Baraniuk, "The smashed filter for compressive classification and target recognition," in *Proc. IS&T/SPIE Symp. Electronic Imaging: Computer Imaging*, San Jose, CA, Jan. 2007.
- [42] J. Haupt, R. Castro, R. Nowak, G. Fudge, and A. Yeh, "Compressive sampling for signal classification," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2006.
- [43] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Mag.*, vol. 19, no. 1, pp. 29–43, 2002.
- [44] K. Krishnamurthy, R. Willett, and M. Raginsky, "Target detection performance bounds in compressive imaging," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, 2012.
- [45] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 6, pp. 2706–2717, 1999.
- [46] M. Wakin and R. Baraniuk, "Random projections of smooth manifolds," *Found. Comput. Math.*, vol. 9, no. 1, pp. 51–77, 2009.
- [47] J. Fowler and Q. Du, "Anomaly detection and reconstruction from random projections," *IEEE Trans. Image Processing*, vol. 21, no. 1, pp. 184–195, Jan. 2012.
- [48] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Effects of linear projections on the performance of target detection and classification in hyperspectral imagery," *J. Appl. Remote Sensing*, vol. 5, no. 1, Nov. 2011.

[Christine Guillemot and Olivier Le Meur]

Image Inpainting

[Overview and recent advances]

Image inpainting refers to the process of restoring missing or damaged areas in an image. This field of research has been very active over recent years, boosted by numerous applications: restoring images from scratches or text overlays, loss concealment in a context of impaired image transmission, object removal in a context of editing, or disocclusion in image-based rendering (IBR) of viewpoints different from those captured by the cameras. Although earlier work dealing with disocclusion has been published in [1], the term *inpainting* first appeared in [2] by analogy with a process used in art restoration.

Image inpainting is an ill-posed inverse problem that has no well-defined unique solution.

To solve the problem, it is therefore necessary to introduce image priors. All methods are guided by the assumption that pixels in the known and unknown parts of the image share the same statistical properties or geometrical structures. This assumption translates into different local or global priors, with the goal of having an inpainted image as physically plausible and as visually pleasing as possible.

The first category of methods, known as *diffusion-based inpainting*, introduces smoothness priors via parametric models or partial differential equations (PDEs) to propagate (or diffuse) local structures from the

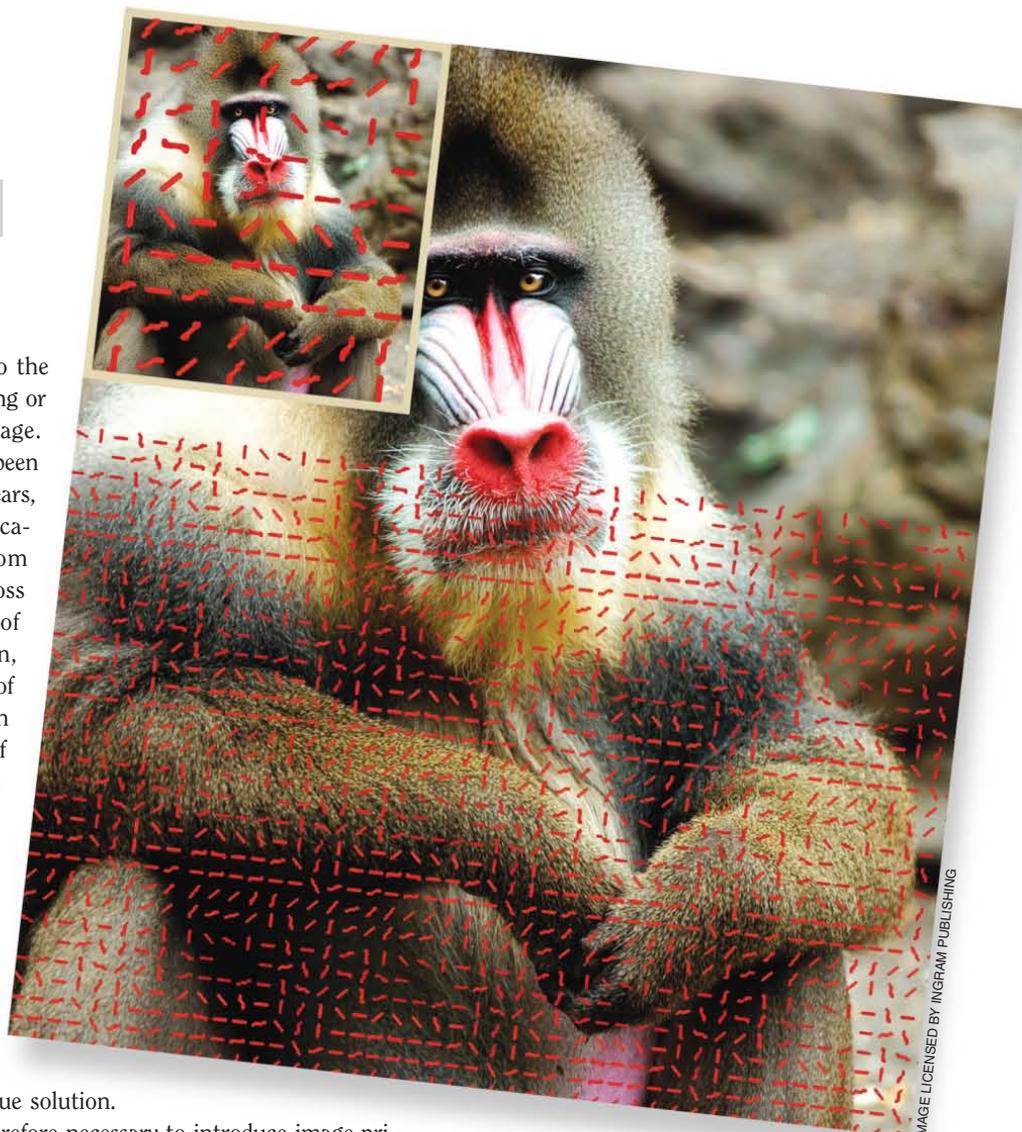


IMAGE LICENSED BY INGRAM PUBLISHING

Digital Object Identifier 10.1109/MSP.2013.2273004

Date of publication: 5 December 2013

exterior to the interior of the hole (as shown in Figure 1, where U denotes the unknown part or the hole to be filled in, and S the source or known part of the image). Many variants exist using different models (linear, nonlinear, isotropic, or anisotropic) to favor the propagation in particular directions or to take into account the curvature of the structure present in a local neighborhood. These methods are naturally well suited for completing straight lines, curves, and for inpainting small regions. They, in general, avoid having unconnected edges that are perceptually annoying. However, they are not well suited for recovering the texture of large areas, which they tend to blur.

The second category of methods is based on the seminal work of Efros and Leung [3] and exploits image statistical and self-similarity priors. The statistics of image textures are assumed to be stationary (in the case of random textures) or homogeneous (in the case of regular patterns). The texture to be synthesized is learned from similar regions in a texture sample or from the known part of the image. Learning is done by sampling, and by copying or stitching together patches (called *exemplar*) taken from the known part of the image. The corresponding methods are known as exemplar-based techniques.

With the advent of sparse representations and compressed sensing, sparse priors have also been considered for solving the inpainting problem. The image (or the patch) is in this case assumed to be sparse in a given basis [e.g., discrete cosine transform (DCT), or wavelets]. Known and unknown parts of the image are assumed to share the same sparse representation. Exemplar-based and sparse-based methods are better suited than diffusion-based techniques for filling large texture areas. Hybrid solutions have then naturally emerged, which combine methods dedicated to structural (geometrical) and textural components.

This article surveys the theoretical foundations, the different categories of methods, and illustrates the main applications.

THE INPAINTING PROBLEM

An image I can be mathematically defined as

$$I: \begin{cases} \Omega \subset \mathcal{R}^n \rightarrow \mathcal{R}^m \\ \mathbf{x} \rightarrow I(\mathbf{x}), \end{cases} \quad (1)$$

where \mathbf{x} represents a vector indicating spatial coordinates of a pixel p_x , which in the case of a two-dimensional (2-D) image ($n = 2$), is defined as $\mathbf{x} = (x, y)$. In the case of a color image, each pixel carries three color components ($m = 3$) defined in the (R, G, B) color space. Each c th image color channel of I is denoted $I^c: \Omega \rightarrow \mathcal{R}$. In the inpainting problem, the input image I (i.e., each color channel of the image) is assumed to have gone through a degradation operator, denoted M , which has removed samples from the image. As a result, the generic definition domain Ω of the input image I can be seen as composed of two parts: $\Omega = S \cup U$, S being the known part of I (source region) and U the unknown part of I , which we search to estimate. The observed degraded version F of the image can also be expressed as $F = MI$.

The goal of inpainting is to estimate the color components of the pixels p_x located at each position \mathbf{x} in the unknown

region U , from the pixels located in S the known region, to finally construct the inpainted image. The objective in terms of quality is that the recovered region looks natural to the human eye, and is as physically plausible as possible. Typical inpainting artifacts are unconnected edges, blur, or inconsistent pieces of texture (also called texture garbage).

DIFFUSION-BASED INPAINTING USING SMOOTHNESS PRIORS

The term *diffusion* comes from the idea of propagating local information with smoothness constraints, by analogy with physical phenomena like heat propagation in physical structures. These phenomena can be formalized with PDEs, and diffusion is therefore performed using PDE-based regularization. Inpainting using diffusion smoothly propagates local image structures from the exterior to the interior of the hole, “imitating” the gesture of professional painting restorators. The considered data are assumed to satisfy smoothness constraints and are iteratively regularized, producing a continuous sequence of smoother images.

Image regularization can be defined locally as the diffusion of pixel values using PDEs, or formulated as the minimization of a functional measuring a global image variation. Moreover, to preserve edges, the regularization (or smoothing) must follow directions given by the local image structure. If the pixel is located on an image contour, the smoothing must be performed along the contour direction and not across boundaries. If the pixel is located in an homogeneous region, the smoothing can be performed in all directions. The first step is therefore to retrieve the local image geometry and then use PDEs or variational methods to describe continuous evolutions of the image and of its structures.

RETRIEVING IMAGE LOCAL GEOMETRY

Local image geometry is in general retrieved by computing gray level lines (also called *isophotes*) or structure tensors.

ISOPHOTES

Isophotes are lines of constant intensity within an image as shown by the red lines in Figure 1. Their directions at a given pixel p_x (located on the front line) are given by the normal to the discretized gradient vector computed at this point, which will be mathematically expressed as ∇I^\perp (instead of $\nabla I_{p_x}^\perp$ for simplifying the notations). In the case of 2-D color images I with three color channels, the derivatives are first computed separately for each color channel and then added to produce the final color gradient in the pixel location p_x . The discretized gradient vector gives the direction of largest spatial changes, while its 90° rotation is the direction of smallest spatial changes, hence, of the isophotes.

STRUCTURE TENSORS

Another method used to retrieve local geometry is based on the computation of the spectral elements of the structure tensor, also called the *Di Zenzo matrix* [4]. The structure tensor for a scalar image (with only one color channel) is computed at each point p_x as $G = \nabla I \nabla I^T$, where the term ∇I denotes

the image spatial gradient, and the term $\nabla \mathbf{I}^T$ denotes the transpose of the image gradient. In the case of color images with three color channels, the structure tensor is computed for each color channel and the results are then added as $\mathbf{G} = \sum_{c=1}^3 \nabla \mathbf{I}^c \nabla \mathbf{I}^{cT}$ to produce the structure tensor of the image \mathbf{I} . By using the spectral decomposition, the structure tensor \mathbf{G} can be expressed as $\mathbf{G} = \mathbf{R} \mathbf{D} \mathbf{R}^T$, where the columns of \mathbf{R} are the eigenvectors (\mathbf{v}_1 and \mathbf{v}_2) of \mathbf{G} and \mathbf{D} is the diagonal matrix whose entries are the corresponding eigenvalues λ_1 and λ_2 . The orientation of the eigenvector \mathbf{v}_2 corresponding to the smallest eigenvalue λ_2 is the orientation with the lowest fluctuations (i.e., of the isophote; see Figure 1). The orientation of the eigenvector \mathbf{v}_1 corresponding to the highest eigenvalue λ_1 gives the gradient direction (i.e., of the normal to the level curve at this point). The retrieved local image geometry can be used to control directions along which pixel values are propagated in the inpainting process, as well as the regularization strength.

In the following paragraphs, we first recall the basics of PDE-based regularization and then present how PDE-based regularization is used for inpainting. We finally give the main ideas of image regularization and inpainting with variational methods.

BASICS OF PDE-BASED REGULARIZATION OR DIFFUSION

The simplest formulation concerns isotropic regularization (or diffusion), which stems from the linear heat (or heat flow) equation

$$\begin{cases} \mathbf{I}(t=0) = \mathbf{F} \\ \frac{\partial \mathbf{I}}{\partial t} = \Delta \mathbf{I}, \end{cases} \quad (2)$$

where \mathbf{F} is a degraded version of the original picture \mathbf{I} . The term $\Delta \mathbf{I}$ denotes the image Laplacian. The PDE evolution is parameterized with a time variable t , which describes the continuous progression of the function \mathbf{I} . Isotropic diffusion, which minimizes these variations in all directions, acts as a low-pass linear filtering suppressing high frequencies in the image. For this reason, the method suffers from blur close to edges and contours. More general formalisms (using nonlinear PDEs), first used for describing physical and fluid dynamics phenomena, have been introduced to better preserve edges and sharpness.

NONLINEAR ISOTROPIC DIFFUSION

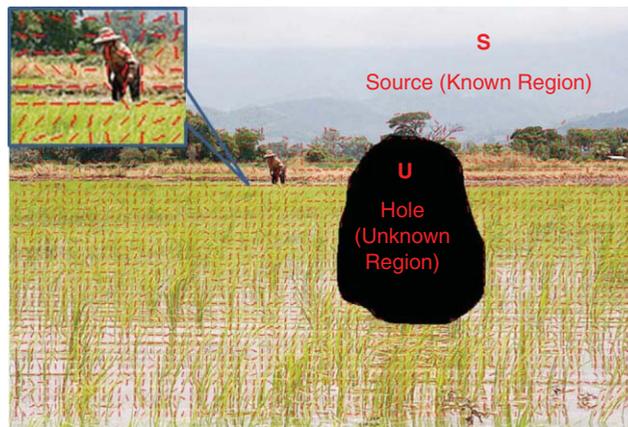
The heat equation can be rewritten in a divergence form as

$$\frac{\partial \mathbf{I}}{\partial t} = \Delta \mathbf{I} = \text{div}(\nabla \mathbf{I}), \quad (3)$$

where the notation $\text{div}(\cdot)$ stands for the divergence operator which measures how fast a vector field (here image intensities) is changing in x and y directions. This led Perona and Malik in [5] to introduce a nonlinear extension of the heat equation, as

$$\frac{\partial \mathbf{I}}{\partial t} = \text{div}(g(\|\nabla \mathbf{I}\|^2) \nabla \mathbf{I}) \quad (4)$$

by introducing a diffusion coefficient (also called conductivity coefficient), which is a scalar in $[0, 1]$ returned by a decreasing



[FIG1] Isophotes or lines (in red) of constant intensity represented for one out of ten pixels with a zoom on one region of the image. (Original image courtesy of [39].)

function $g(\cdot)$ of the image gradient. The goal of the diffusion coefficient is to limit diffusion around edges and avoid smoothing across region boundaries. Two functions have been proposed in [5] for computing the diffusion coefficient

$$g(\|\nabla \mathbf{I}\|) = e^{-\|\nabla \mathbf{I}\|/\alpha} \quad (5)$$

and

$$g(\|\nabla \mathbf{I}\|) = \frac{1}{1 + (\|\nabla \mathbf{I}\|/\alpha)^2}, \quad (6)$$

where α is a constant chosen experimentally to control the sensitivity to edges. The function $g(\cdot)$ vanishes in the neighborhood of steep edges (high gradients) and returns a value close to 1 on flat regions (low gradients).

The method is often referred to as *anisotropic diffusion* even though it is rather a nonlinear diffusion. Indeed, the filter applied locally is isotropic, but its response is adapted locally and varies in space so that the diffusion process is lessened near edges and is strong in homogenous areas. A good review of this model can be found in [6].

ANISOTROPIC DIFFUSION USING TENSOR FIELDS

One step further has been to introduce locally adapted filters for performing truly anisotropic diffusion close to image structures such as edges [7]. Anisotropic regularization (or diffusion) refers to a smoothing in privileged spatial directions with different weights. Weickert in [7] used a field of diffusion tensors to find privileged directions of diffusion.

The diffusion tensor \mathbf{D} (symmetric and positive-definite matrix) is derived in each point $\mathbf{x} = (x, y)$ of the image from the spectral elements of the structure tensor $\mathbf{G} = \nabla \mathbf{I} \mathbf{I}^T$ as [7]

$$\mathbf{D} = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T, \quad (7)$$

where \mathbf{v}_1 and \mathbf{v}_2 are the eigenvectors (λ_1 and λ_2 are the corresponding eigenvalues) of the structure tensor \mathbf{G} . Pixel values are therefore anisotropically smoothed along local directions given by the eigenvectors \mathbf{v}_1 and \mathbf{v}_2 , with weights (or smoothing strength) given by the corresponding eigenvalues. In other words, the eigenvalues determine the diffusivities

in the directions of the eigenvectors. The diffusion is governed by the PDE

$$\frac{\partial I}{\partial t} = \text{div}(D\nabla I). \quad (8)$$

Note that, if D equals the identity matrix, we obtain the heat equation, i.e., isotropic diffusion. Note also that the tensor can be regularized with Gaussian kernels, which leads to a smoothing of the image with small elongated kernels along edges and with Gaussian-like kernels in homogeneous regions. An edge-enhancing diffusion (EED) is further proposed in [7], where the diffusivity decreases with the increased contrast in the direction perpendicular to edges. The effect of the varying diffusivity of EED can be seen in Figure 2. A coherence-enhancing diffusion (CED) is also proposed in [8], which increases the diffusivity along the coherent direction given by v_2 when the coherence measured by $(\lambda_1 - \lambda_2)^2$ increases.

ANISOTROPIC DIFFUSION USING ORIENTED LAPLACIANS

Image smoothness (and its variation) can also be measured by a discretization of the 2-D Laplacian of the image, instead of using gradient or divergence operators. Anisotropic diffusion models are in this case expressed as

$$\frac{\partial I}{\partial t} = \lambda_1 I_{v_1 v_1} + \lambda_2 I_{v_2 v_2}, \quad (9)$$

where the terms $I_{v_1 v_1}$ and $I_{v_2 v_2}$ are the image Laplacians, i.e., the second derivatives of I in the directions given by the vectors v_1 and v_2 that, as above, are derived from the local image structures (gradients, isophotes, or tensor fields). Here again, the above equations can be applied on each color channel of the image separately or on vectors formed by the three components (R,G,B) of color images to smooth multivalued images.

The diffusion is controlled by the knowledge of the smoothing directions v_1 and v_2 , and the corresponding weights λ_1 and λ_2 . This can be seen as heat flows oriented along orthonormal directions given by v_1 and v_2 .

INPAINTING USING DIFFUSION PDES

The use of regularization or diffusion for image inpainting was pioneered by Bertalmio et al. [2] in 2000. All PDE-based image regularization methods (isotropic or anisotropic) described in the previous sections can be used for inpainting. The authors in [2] use an anisotropic model that propagates image Laplacians [as formalized by (9)] from the surrounding neighborhood into the interior of the hole. The directions of the propagation are given by the directions of the isophotes estimated by the perpendicular to the image gradient in each point. The algorithm numerically solves the equation

$$\frac{\partial I}{\partial t} = \nabla(\Delta I) \nabla I^\perp \quad (10)$$

for the image intensity I inside the hole, until a steady state solution ($\nabla(\Delta I) \nabla I^\perp = 0$), which means that the image Laplacians (ΔI) remain constant (no variations captured by the

gradient ∇) in the directions ∇I^\perp of the isophote. The term $\nabla(\Delta I) \nabla I^\perp$ is the derivative of ΔI in the direction ∇I^\perp , leading to a smooth continuation of available information (the Laplacians) inside the region to be inpainted. In other words, the image information is propagated inside the missing region in a way that aims at preserving the isophote directions. The analogy between this propagation of image intensity along smooth level curves and the transport of vorticity in fluid dynamics formalized with the Navier–Stokes equations has been established in [9]. The image intensity is seen as a stream in fluid dynamics and isophote lines are seen as flow streamlines.

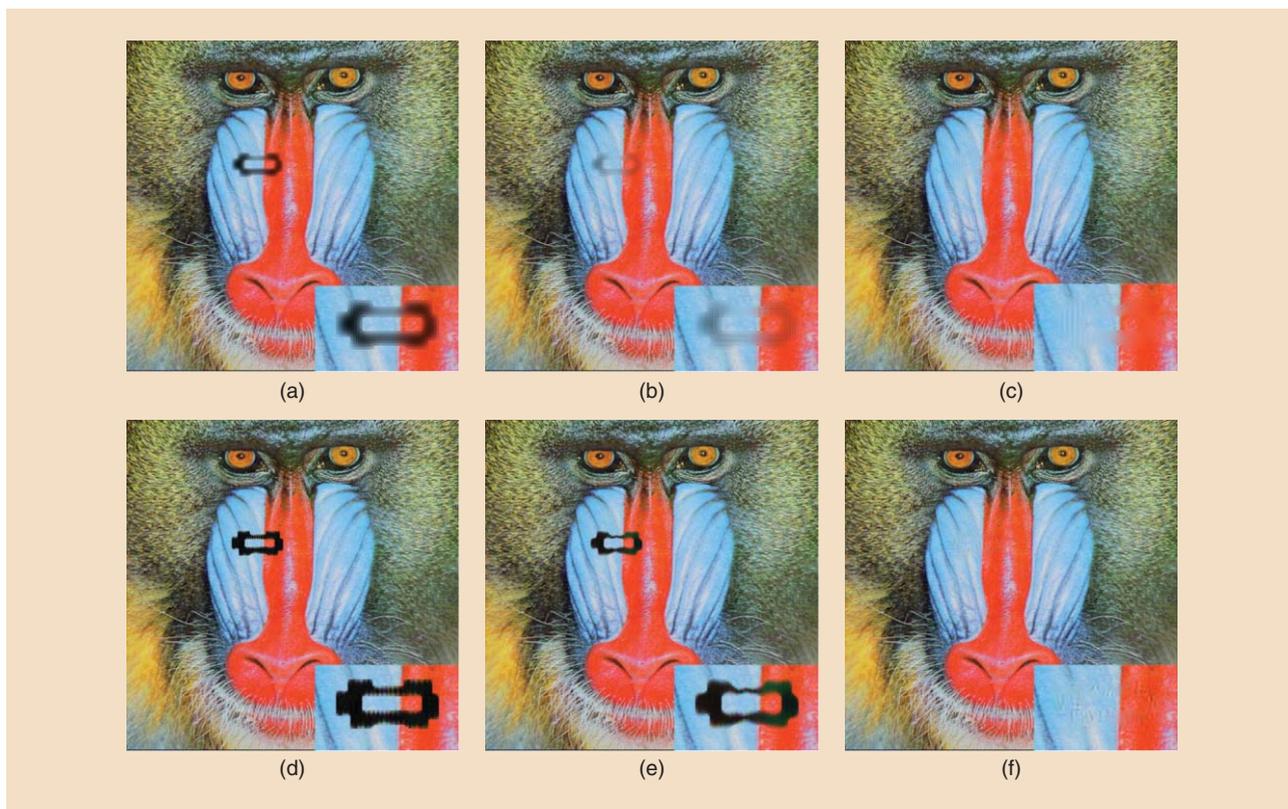
Variants have then been proposed in the use of anisotropic diffusion for image inpainting, for either reducing complexity or better preserving structures. PDE-based methods require implementing iterative numerical methods that are generally quite slow. A fast marching technique is described in [10], which estimates the unknown pixels in one pass using weighted means of already calculated pixels. A trace-based PDE model is proposed in [11] to regularize multivalued images (with multiple color channels). It was observed in [11] that Gaussian behavior inherent to the use of tensors for defining orientation and strength of the diffusion degrades the reconstruction of curved image structures such as corners. This observation led the author in [11] to use heat flows constrained on integral curves to better preserve curvatures in image structures.

Diffusion methods tend to prolong structures (e.g., isophotes) arriving at the border of the region to be filled in. These methods are hence successful for piecewise smooth images, for propagating strong structures, or for filling small gaps. However, they are not well suited for textured images, especially if the region to be completed is large. Although intended to preserve edges, after a few diffusion iterations, the inpainted region appears smooth with a lot of blur when the missing region is large. Figure 2 shows inpainting results after a few iterations when using two different diffusion equations: isotropic diffusion with heat equation, and edge-enhancing diffusion (EED) [7]. Isotropic diffusion introduces a blur over the entire region to be filled in. On the contrary, EED (and similarly for CED) introduces less smoothing across edges. Figure 3 illustrates typical artifacts of diffusion-based inpainting methods when the hole to be filled in is large.

VARIATIONAL INPAINTING

Image regularization, hence image inpainting, may also be formulated as a variational problem where the image is seen as a function of bounded variation (BV). The total variation (TV) image model, first proposed for denoising and deblurring applications in [12], has been applied to image inpainting in [13]–[15]. The energy function is based on the total variational norm, hence the name of the TV image model.

Keeping the goal of a smooth propagation of image intensity, the idea consists in searching for a function $\hat{I}(x)$ of BV on the set Ω ($BV(\Omega)$) on which the input image I is defined, which minimizes a TV energy (defined as the integral of the gradient magnitude) of the image inside the hole, under some



[FIG2] The diffusion process with isotropic diffusion after (a) ten, (b) 20, and (c) 600 iterations. Edge-enhancing diffusion [7] after (d) ten, (e) 50, and (f) 600 iterations.

constraint of fidelity to image observations. The minimization problem is written as

$$J_{TV}(\hat{I}) = \int_{\Omega} |\nabla \hat{I}(x)| dx + \lambda \int_{\Omega \cup U} (I(x) - \hat{I}(x))^2 dx. \quad (11)$$

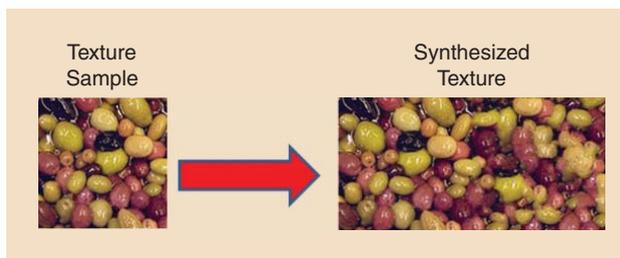
The first integral term represents the regularization term, whereas the second term is a data term measuring the fidelity of the reconstructed image to the input image for the known samples. The difference with the denoising problem in [12] resides in the fact that the second integral is computed over $\Omega \setminus U$ instead

of the entire domain Ω . The regularization strength can be seen as spatially varying, where $\lambda(x) = 0$ in U . It can also vary locally according to the image geometry or to given image models.

Note that a variational formulation has been first applied to the disocclusion problem in [1] where the image is reconstructed from a set of level lines defined as BV functions. The level lines within the hole are geodesic paths joining isophotes arriving at the boundary of the hole. This method is the first variational approach where the detection of the isophotes is followed by their variational continuation across the hole.



[FIG3] Typical blurring effects of (a) and (b) diffusion-based methods (example with the method in [10]) when the hole to be filled in is large. [(a) courtesy of [39]. (b) generated by Christine Guillemot and Olivier Le Meur.]



[FIG4] The texture synthesis problem. The goal is to produce a texture larger than the input sample with a similar visual appearance. (Figure used with permission from the Massachusetts Institute of Technology.)

TV regularization is an effective inpainting technique that is capable of recovering sharp edges with, however, some problems of connectivity. To further improve the edge connectivity, the TV inpainting model has been extended by introducing in the regularization term energy functionals taking into account curve structures, via curvature-driven diffusion models [16], or the Euler elastica functional [17].

EXAMPLAR-BASED METHODS

A second family of inpainting methods has appeared in the last decade based on seminal work on texture synthesis [3], [18] with the aim of better recovering the texture of the missing area. The texture synthesis problem is slightly different from image inpainting. The goal of texture synthesis is to create a texture from a given sample as shown in Figure 4, in such a way that the produced texture is larger than the source sample with a similar visual appearance. This problem is also referred to as *sample-based texture synthesis*. There is a vast body of work on texture synthesis, using either local region growing or a global optimization.

Exemplar-based inpainting has been, for a large part, inspired by local region-growing methods that grow the texture one pixel or one patch at a time, while maintaining coherence with nearby pixels.

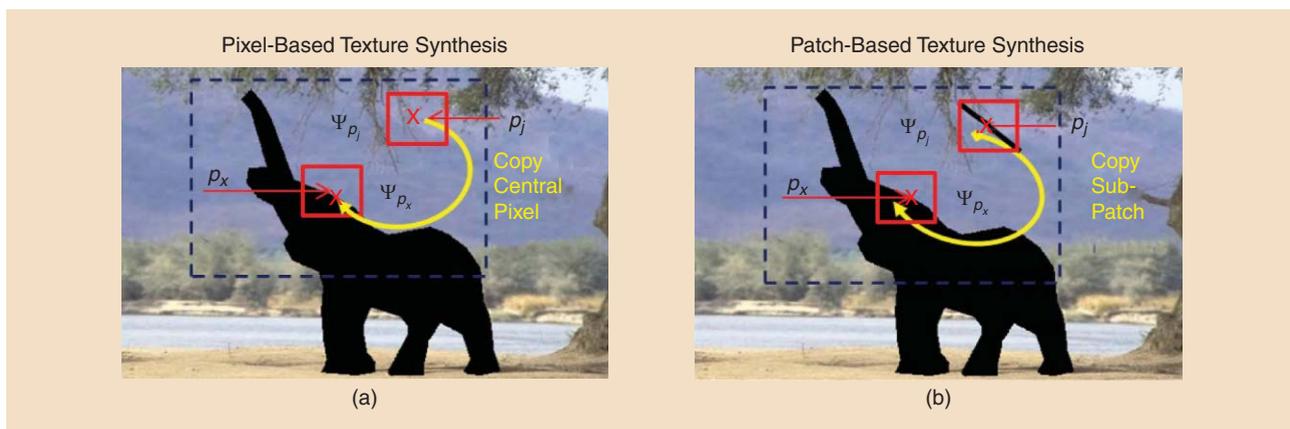
Most local pixel-based synthesis techniques rely on Markov random fields (MRFs) modeling of textures. Instead of running

a complex probabilistic inference on the graphical model of the MRF for learning the missing pixels from the input sample, simpler, yet efficient, approximate solutions have been proposed in [3]. Exploiting both locality (the color of a pixel being assumed to depend on its local neighborhood only) and stationarity (the dependency is independent of the pixel location), the missing pixels are learned by sampling and copying the central pixel of a patch from the sample texture that best matches the known neighborhood of the input pixel to be synthesized, according to a certain distance. Similarly, in [18], the output image is generated pixel-per-pixel in a raster scan order choosing at each step a pixel from the sample image which neighborhood is most similar to the currently available neighborhood in the texture being synthesized. Texture synthesis methods have evolved from pixel-based to patch-based techniques, with recent enhancements relying on elaborated blending and quilting strategies, using, e.g., graph-cut techniques to minimize energy terms along a seam. We come back to this issue in the section “Patch Stitching with Blending and Quilting.”

Texture synthesis methods directly apply to the inpainting problem where the known part of the image can be seen as the input texture sample from which the missing pixels can be learned [as illustrated in Figure 5(a)].

The simple pixel-based texture synthesis technique in [3] proceeds as follows. Let p_x be a pixel located at position x in the image I , and Ψ_{p_x} be the patch centered on the pixel p_x . This patch has a known part ($\Psi_{p_x}^S$) and an unknown part ($\Psi_{p_x}^U$). The idea is to search for the patch Ψ_{p_j} (centered on p_j) the most similar to the input patch Ψ_{p_x} . The central pixel p_j having a neighborhood most similar to the known neighborhood of p_x is then copied to recover p_x . Image information is therefore pixel-per-pixel propagated from the known part to the unknown part of the image.

This pixel-per-pixel recovery algorithm suffers from a high computational cost, even if its complexity can be reduced by constraining the search for best matching patches among the candidates of the neighboring pixels that have been already inpainted [19]. Another limitation is the



[FIG5] The principle of exemplar-based methods: search for the patch the most similar to the known part of the input patch to be completed and copy the central pixel for (a) pixel-based approaches or (b) a set of pixels for patch-based approaches. (Figure used with permission from [24].)

difficulty for this type of approach to synthesize textures that are not frontal (with some perspective transformations) and to fill in large and dispersed holes. Moreover, although performing better than diffusion methods on textured areas, the pixel-based synthesis techniques often suffer from synthesis errors propagation and from repetitive patterns, which look unnatural especially in the case of stochastic textures. They also run into difficulties when synthesizing texture formed by an arrangement of small objects.

Approaches synthesizing entire patches rather than only one pixel at a time have then emerged to cope with the drawbacks just mentioned. Instead of synthesizing the missing region pixel per pixel, the idea of patch-based solutions is to recover entire patches in one step by sampling and copying texture patterns (entire patches) from the source [20]. The first step for estimating the pixels in the unknown part $\Psi_{p_x}^U$ of the patch again consists in searching for the patch Ψ_{p_j} (centered on p_j), which is the most similar to the patch $\Psi_{p_x}^S$. But this time, all the pixels from Ψ_{p_j} which are located at the same position as $\Psi_{p_x}^U$ are copied to estimate the unknown pixels of the input patch, as shown on the right side of Figure 5.

The terminology *exemplar-based inpainting* now mostly refers to these methods that synthesize entire patches by learning from patches in the known part of the image. Since they synthesize entire patches at once, these methods are faster than pixel-based approaches. Many variants have then been introduced to optimize patch-based methods. These variants concern:

- distance metrics for finding best matching patches
- fast search of best matching patches
- patch processing order
- global spatial coherence via constrained search or global optimization
- patch stitching with blending and quilting
- methods to learn unknown pixels from best matching patches
- multiscale refinement.

These variants are discussed in the next section.

ISSUES AND VARIANTS OF EXAMPLAR-BASED INPAINTING

DISTANCE METRICS

Several metrics exist for measuring similarity between images or between image patches. The most widely used metrics can be classified in the following categories: pixel-based metrics measuring the similarity in terms of difference or cross-correlation between pixel color values and statistics-based metrics measuring the similarity between probability distributions of pixel color values in patches. The sum of squared differences (SSD), the L_p norm, as well as the normalized cross-correlation belong to the first category. Statistics-based metrics include the Bhattacharyya distance, the normalized mutual information (NMI), and the Kullback–Leibler divergence.

The most widely used metric to search for similar patches is the SSD. However, as observed in [21], the SSD introduces

some bias towards uniform regions. In other words the SSD favors the copy of pixels from uniform regions.

A weighted Bhattacharyya distance has been proposed in [21] ($d_{(SSD,BC)}$) to cope with this limitation. This metric is computed as

$$d_{(SSD,BC)}(\Psi_{p_x}, \Psi_{p_j}) = d_{SSD}(\Psi_{p_x}, \Psi_{p_j}) \times d_{BC}(\Psi_{p_x}, \Psi_{p_j}), \quad (12)$$

where the term $d_{SSD}(\Psi_{p_x}, \Psi_{p_j})$ is the SSD between the two patches Ψ_{p_x} and Ψ_{p_j} . The term $d_{BC}(\Psi_{p_x}, \Psi_{p_j})$ is a modified Bhattacharyya distance given by $d_{BC}(\Psi_{p_x}, \Psi_{p_j}) = \sqrt{1 - \sum_k \sqrt{h_1(k)h_2(k)}}$, with h_1 and h_2 representing the histograms of patches Ψ_{p_x} and Ψ_{p_j} , respectively.

However, when two patches have the same distribution, their distance d_{BC} is zero, hence the weighted Bhattacharyya distance is also zero, even if one patch is a geometrically transformed version of the other one. A variant of the weighted Bhattacharyya distance addressing this problem is proposed in [22] where the SSD is multiplied by $(1+d_{BC})$. When two patches have the same distribution, the distance value is equal to the SSD between the two patches.

Note that other distance metrics also exist that account for geometrical transformations [23]. Transformed (rotated, scaled, and mirror) versions of existing patches are, in this case, included as possible match candidates [24].

FAST SEARCH OF BEST MATCHING PATCHES

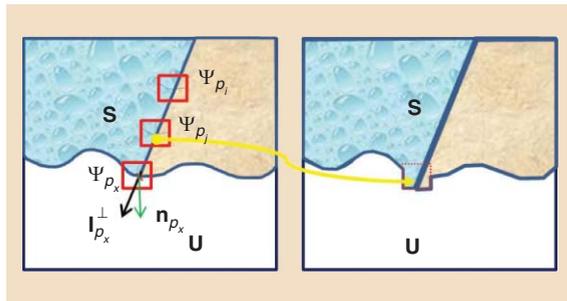
Exemplar-based inpainting methods first search for K -nearest neighbors (K -NNs) within the known part of the image. A naive solution to the NN search problem is to compute the distance from the query patch to all possible candidate patches, treating each patch as a point in a high-dimensional space. Faster and approximate NN search methods exist that organize the candidates in specific space-partitioning data structures, such as the k -dimensional trees (kd-trees) [25] or the vantage point trees (vp-trees) [26], according to their distribution in the search space. The kd-trees are a special case of binary space partitioning trees that divide the space along different coordinates. The set of data points is bisected at the median of all points in a selected dimension to build a binary tree. The vp-tree, instead of splitting along coordinate values, splits the set of data points according to their distance to a specific point called the vantage point. The NN search can then be efficiently done by using the tree properties to quickly eliminate large portions of the search space and check only a small portion of candidates. The kd-tree-based matching is one of the most widely used algorithms for finding the nearest patch. However, its number of searched nodes increases exponentially with the space dimension. When the dimension is large (e.g., higher than 15), its search speed becomes very slow. Several NN search algorithms are assessed in [27] for finding similar patches in images.

The tree-based approximate NN (ANN) search methods treat the queries separately. A randomized patch search algorithm called PatchMatch was introduced in [28], which exploits dependency among the queries to perform some collaborative search. The idea relies on the assumption that images are coherent. Once a pair of similar patches has been

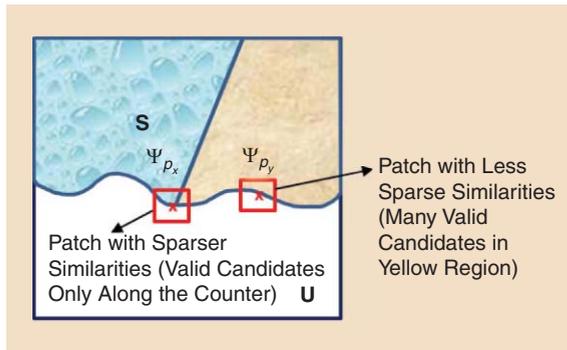
[TABLE 1] PATCH FILLING ORDER.

THE DATA TERM $D(p_x)$ USED FOR FAVORING PATCHES WITH STRONGER STRUCTURES CAN TAKE DIFFERENT FORMS.

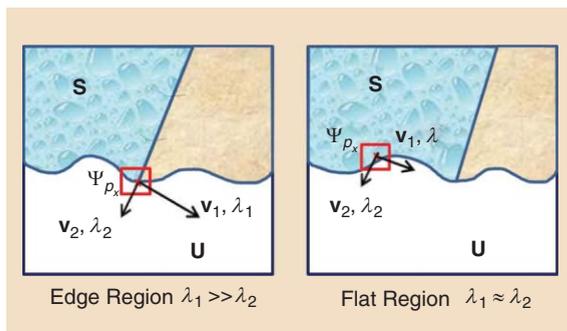
1) GRADIENT-BASED DATA TERM [31]: $D(p_x) = (|\nabla I_{p_x}^\perp \cdot \mathbf{n}_{p_x}|) / \alpha$, WHERE \mathbf{n}_{p_x} IS THE NORMAL TO THE FRONT LINE AT THE POSITION OF THE PIXEL p_x , AND THE PERPENDICULAR TO THE GRADIENT $\nabla I_{p_x}^\perp$ REPRESENTS THE ISOPHOTE.



2) SPARSITY-BASED STRUCTURE CONFIDENCE TERM [32]: $D(p_x) = \|\mathbf{w}_{p_x}\|_2 \times \sqrt{N_s/N}$, WHERE N_s AND N ARE THE NUMBERS OF VALID AND CANDIDATE PATCHES IN THE SEARCH WINDOW.



3) TENSOR-BASED PRIORITY [8], [33]: $D(p_x) = \alpha + (1 - \alpha) \exp(-\eta / (\lambda_1 - \lambda_2)^2)$, WHERE $\eta \geq 0$ ($\eta = 8$) AND $\alpha \in [0, 1]$ ($\alpha = 0.01$). WHEN $\lambda_1 \approx \lambda_2$, FLAT REGION IMPLIES LOW PRIORITY ($D(p_x)$ SMALL). WHEN $D(p_x) \lambda_1 \gg \lambda_2$, PRESENCE OF STRUCTURE IMPLIES HIGH PRIORITY (TENDS TO BE "1").



found in two images, then their neighbors in the image plane (patches shifted by a few pixels) are likely to be also similar. Therefore, the matching result of a query patch can be propagated to the nearby queries, providing a good initial guess that is then updated by some randomly sampled candidates.

PatchMatch is a fast algorithm for computing dense approximate NN correspondences between patches of two image regions. The set of approximate correspondences is called an NN field

(NNF). The method searches for the NNF by proceeding as follows. The NNF is first initialized either by random values or prior information. Initializing the NNF with a random guess is likely to provide few good guesses. The NNF is then iteratively refined by interleaving two operations called propagation and random search at the patch level. The propagation step updates a patch offset with the known offsets of its causal neighborhood, exploiting the image coherency. At even iterations, offsets are propagated from up and left patches, whereas at odd iterations, offsets are propagated from right and bottom patches. The second operation performs a local random search to seed the patch matches, and these matches are then propagated by iterating a small number of times.

The algorithm was shown to be much faster than kd-trees with, however, less accuracy. It can be trapped in a local optimum due to the short-distance propagation. To improve the NN search accuracy, the random search step of PatchMatch is replaced with a hashing scheme in [29]. Good matches are propagated to nearby patches as well as to similar patches that were hashed to the same value (i.e., which are similar in appearance). The algorithm therefore runs faster and turns out to be more accurate. The PatchMatch algorithm has been generalized in [30] to find K -NNs instead of one, and to extend the search space with transformations (rotations, scaling).

PATCH PROCESSING ORDER

The missing regions in an image may be, in general, composed of textures and structures. It has been observed [31] that it was important to separate these two components and start by first recovering the structures. This led to proposing patch processing orders that are defined in such a way that patches on structures are filled in first. In general, the processing order is given by a patch priority measure defined as the product of two terms ($P(p_x) = C(p_x)D(p_x)$). The first term accounts for the amount of known pixels versus unknowns in the input patch [this is a so-called confidence term $C(p_x)$] and the second term $D(p_x)$, called data term, reflects the presence of some structure in the patch. This data term can take several forms (see Table 1).

A gradient-based data term is proposed in [31], which favors patches in which the isophote is perpendicular to the front line at pixel p_x . The data term is defined as the absolute value of the inner product between the isophote direction (perpendicular to the gradient $\nabla I_{p_x}^\perp$) and the normal \mathbf{n}_{p_x} to the front line as

$$D(p_x) = \frac{|\nabla I_{p_x}^\perp \cdot \mathbf{n}_{p_x}|}{\alpha} \quad (13)$$

Thanks to this priority term, unknown pixels at the edge of an object have higher priority than pixels located on flat image areas (see Table 1).

A sparsity-based data term has been proposed by Xu et al. [32] to measure the structure confidence of patches. The structure confidence is based on the sparseness of nonzero patch similarities. Structural patches are assumed to have sparser

nonzero similarities with its neighboring patches compared to textural patches. This assumption stems from the observation that structures (edges, corners) are, in general, sparsely distributed in the image. A similarity weight w_{p_x, p_j} (i.e., proportional to the similarity between the two patches centered on p_x and p_j) is computed for each pair of patches. A vector w_{p_x} is formed with the similarity weights between all the pairs (p_x, p_j) where $j = (i_1, j_1), \dots, (i_N, j_N)$, with N being the number of all candidate patches within a search window. The sparsity-based data term is defined as

$$D(p_x) = \|w_{p_x}\|_2 \times \sqrt{\frac{N_s}{N}}, \quad (14)$$

where N_s and N are the numbers of valid (all their pixels are known) and candidate patches in the search window centered on the pixel p_x .

A large value of the structure sparsity term $D(p_x)$ means sparse similarity with neighboring patches—a good confidence that the input patch is on some structure—whereas a small value indicates that the current input patch is highly predictable by many candidates, hence is likely to be a textural patch. This is illustrated in Table 1, where the patch Ψ_{p_x} has sparser similarity than the patch Ψ_{p_j} for which a larger number of matching patches can be found in the uniform yellow region. This term, compared to the one defined on isophotes, better distinguishes structural from textural patches.

A tensor-based data term is proposed in [33] based on Di Zeno's structure tensor [4] computed on the color components (R,G,B) as $\mathbf{J} = \sum_{c=1}^m \nabla I^c \nabla I^{cT}$. As explained above, the eigenvectors and eigenvalues of the structure tensor give an indication of the local geometry. Based on the discrepancy of the eigenvalues, the degree of anisotropy of a local region can be evaluated. The tensor-based data term is therefore written as

$$D(p_x) = \alpha + (1 - \alpha) \exp\left(-\frac{\eta}{(\lambda_1 - \lambda_2)^2}\right), \quad (15)$$

where $\eta \geq 0$ (e.g., $\eta = 8$) and $\alpha \in [0, 1]$ (e.g., $\alpha = 0.01$). On flat regions ($\lambda_1 \approx \lambda_2$), no direction is favored for the propagation (isotropic filling order). When $\lambda_1 \gg \lambda_2$, which indicates the presence of a structure, the data term $D(p_x)$ tends to be "1."

The patch processing order has a strong impact on the quality of the inpainted image: giving a higher priority to structural patches leads to a better recovery of object boundaries, avoiding, e.g., unconnected boundaries in the inpainted image, as shown in Figure 6. Studies of vision psychology show that human observers are quite annoyed by unconnected edges [34].

GLOBAL IMAGE COHERENCE VIA CONSTRAINED SEARCH OR GLOBAL OPTIMIZATION

The approaches in [3], [24], and [31] progress patch per patch in a greedy fashion; for this reason, they do not ensure a global image coherence. The visual quality of the inpainted image can be improved by maximizing similarity between the synthesized patch and original patches in the known part of the image. This can be achieved in pixel-based synthesis approaches by searching for

candidate pixels in the neighborhood of those already used to inpaint the neighbors of the input pixel to be estimated [19].

The authors in [35] constrain the search for candidate patches by using a global coherence measure $\text{Coh}()$ defined as

$$\text{Coh}(\Psi_{p_x}^U) = \min_{p_j \in S} \{d_{\text{SSD}}(\Psi_{p_x}^U, \Psi_{p_j}^U)\}. \quad (16)$$

This coherence measure is used to favor patches giving a synthesized part $\Psi_{p_x}^U$ most similar to the pixels $\Psi_{p_j}^U$ of the known part S of the image. Among the best candidate patches, the algorithm keeps the one that minimizes the above coherence measure. This prevents from pasting in the unknown region a texture that would be too different from original textures. This measure therefore limits texture garbage. Examples of texture garbage are visible in Figure 6.

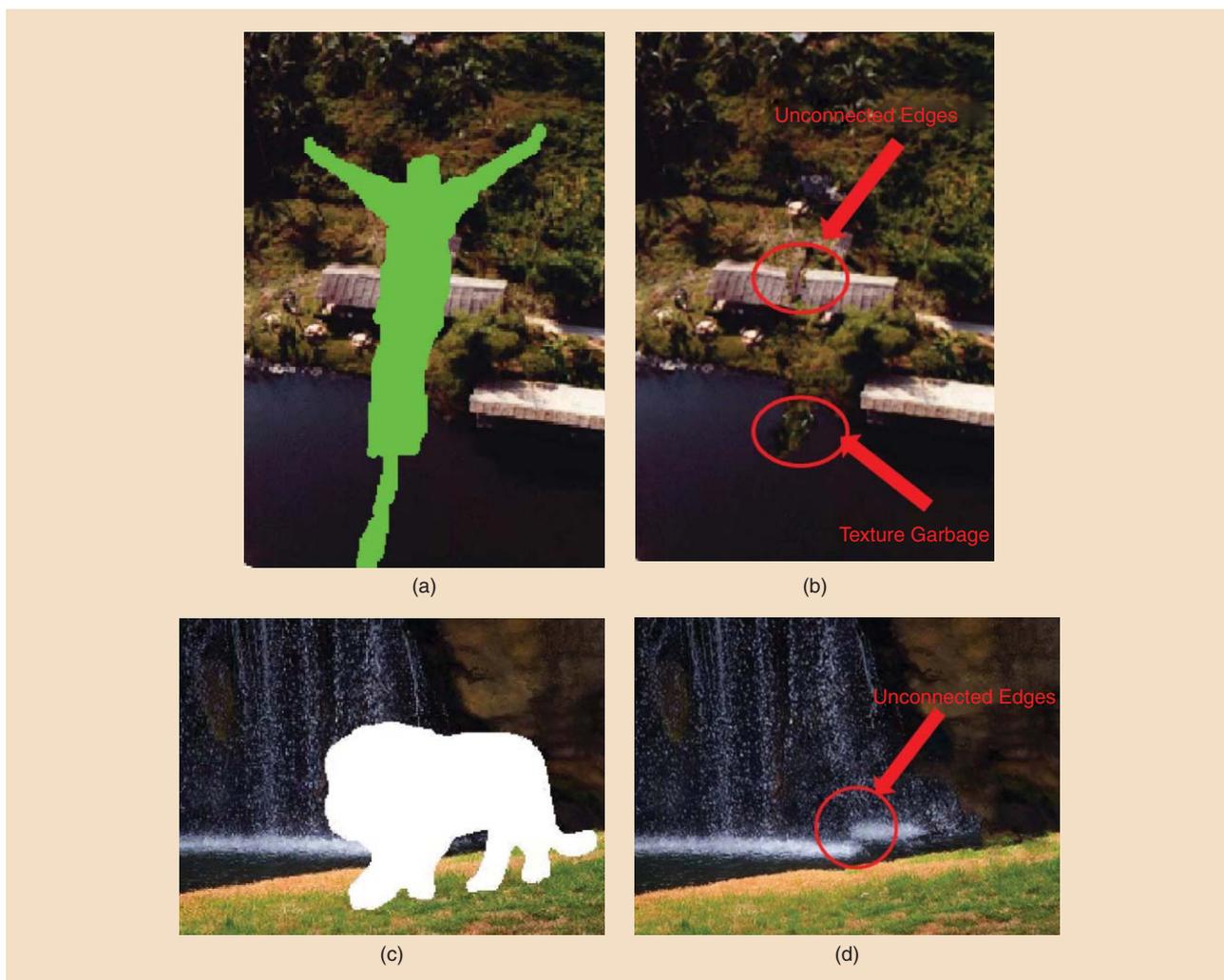
Spatial coherence can also be naturally ensured via a global optimization of MRF energy functions over the entire image. Patch, pixel locations, or offsets are optimized in the MRF by using belief propagation [36] or graph cuts [37]. The shift-map method [37] indeed treats inpainting as a global optimization on the entire image. The method, inspired from the patch transform described in [38], computes a vector field (called shift-map), which maps each pixel in the hole to be filled in to a pixel in the known part of the image. The optimal shift-map is computed as a graph labeling optimization minimizing a global cost function. A node in the graph corresponds to a pixel in the output image labeled by a shift (translational displacement).

The authors in [39] match similar patches in the known part of the image and compute patch offsets. They observed that a majority of patches have similar offsets, and the peaks in the offset distribution correspond to dominant offsets. A stack of images is formed with shifted versions of the input image according to these dominant offsets. The shifted images are then combined by optimizing a global MRF energy function. These methods produce visually pleasing and coherent inpainted images, in particular when the hole to be filled in has homogeneous texture and few structures.

PATCH STITCHING WITH BLENDING AND QUILTING

Filling the unknown part of the input patch may lead to stitching together pieces of texture that are not consistent in terms of color or contrast, or which may contain structures that do not align. Color bleeding or boundary artifacts may therefore be visible in a boundary or transition region. To reduce the boundary effect, the pieces of texture must be stitched together along an optimal seam that will be the best boundary between the two regions.

A quilting method is introduced in [40], which consists of finding an optimal path (called a *seam*) cutting the overlapped regions \mathcal{O} , as illustrated in Figure 7. The seam, which is an eight-connected path of pixels in the overlap region, goes from top to bottom for a vertical seam and from left to right for an horizontal seam. The path determines which patch contributes to pixels at different positions in the overlap region. A vertical (respectively, horizontal) seam contains only one pixel per row



[FIG6] (a)–(d) Typical artifacts of exemplar-based methods: unconnected boundaries and texture garbage. The bungee jumper has been inpainted with the method in [31] and the lion image with [22]. [(a) Original bungee jumper image courtesy of M. Bertalmio. (b) and (d) Courtesy of Christine Guillemot and Olivier Le Meur. (c) Courtesy of [39].]

(respectively, column) in the overlap region. An energy function $e(i, j)$ is defined to evaluate the contrast of the current pixel with respect to its neighbors. A gradient-based energy is commonly used. The energy of a seam s is given by $E(s) = \sum_{(i, j) \in O} e(i, j)$, where (i, j) are the coordinates of a pixel of the seam s inside

the overlap region O . The optimal seam s^* having the minimum energy

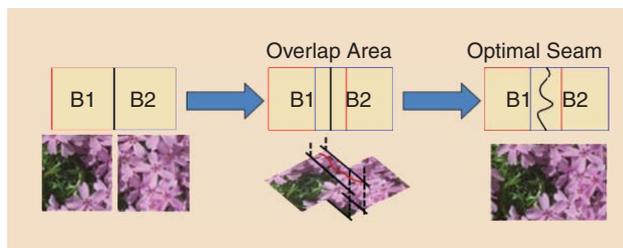
$$s^* = \arg \min_s E(s) \tag{17}$$

is searched by using Dijkstra's shortest path algorithm, dynamic programming [40], or by using graph cuts [23], [41].

Dynamic programming allows solving this problem efficiently by computing the cumulative minimum energy M for all possible connected seams as

$$M(i, j) = e(i, j) + \min(M(i - 1, j - 1), M(i - 1, j), M(i - 1, j + 1)). \tag{18}$$

For a vertical seam, we look for the minimum value on the last row and backtrack from this minimum to find the optimal vertical path. The quilting approach is limited by the number of possible orientations at each pixel. For instance, for the case of a vertical seam, only three directions (bottom-right, bottom, and bottom-left) can be used to build the seam path. In



[FIG7] The quilting texture. Two texture chunks are patched together. An optimal seam is computed to cut the overlap region. On the left side of the seam, texture from block B1 is copied into the final texture. On the right side, pixels belonging to texture B2 are copied. (Figure adapted from [40] and used with permission from the Massachusetts Institute of Technology.)

addition, the quilting method is greedy since the seams are defined either in left-to-right or top-to-bottom order.

To cope with these limitations, a graph-cut method can be used to find the optimal seam. A graph cuts consist of first constructing a specialized graph for the energy function to be minimized such that the minimum cut on the graph, computed by max-flow algorithms, minimizes the energy [42], [43]. Graph cuts have been used for a wide variety of vision problems, particularly in [44] to seamlessly combine and stitch different textures together.

Blending (feathering, pyramid blending, alpha blending, or Poisson blending [45]) and image melding [46] can also be used to seamlessly merge the new patch. Blending aims at handling color inconsistency by smoothly interpolating the error in the transition region. Image melding [46] synthesizes the texture in the transition region by transforming one patch (or image) to be stitched into the other.

EMBEDDING NEAREST NEIGHBORS

The first exemplar-based methods were estimating the unknown part of the input patch by copying pixels of the single best match among possible candidates. A better estimate of the unknown pixels can be obtained by computing a linear combination of several candidate patches, e.g., of the K -NN, as illustrated in Figure 8. This places the inpainting problem in a neighbor embedding framework. Different approaches can be used to compute the weights of the linear combination of the K -NN. The authors in [47] proposed a nonlocal means approach. The unknown pixels of the patch to be completed are inferred by the nonlocal means of several candidate patches instead of taking the single best match. The weights of the linear combination are computed using a similarity kernel as

$$w_k = \exp\left(-\frac{\|\Psi_{p_x}^S - \Psi_{p_j}^S\|_2^2}{h}\right), \quad j = (i_k, j_k), k = 1 \dots N, \quad (19)$$

where N is the number of all candidate patches in the search window. The parameter h acts as a filtering parameter. This kernel gives higher weights to patches that are more similar to the known samples of the input patch. This approach is inspired from the nonlocal means (NLM) algorithm used for denoising in [48] and for texture synthesis in [3].

The weights can also be computed using least square approximations of the known pixels of the input patch under various constraints. A constraint that the weights sum to one leads to placing the problem in the locally linear embedding (LLE) framework [49]. The search for the weights of the linear combination of the K best matching patches, centered on pixels p_j located at positions $j = (i_k, j_k), k = 1 \dots K$, is expressed as

$$\begin{aligned} \operatorname{argmin}_{w_k, k=1 \dots K} & \left\| \Psi_{p_x}^S - \sum_k w_k \Psi_{p_j}^S \right\|_2^2 \\ \text{s.t.} & \sum_k w_k = 1, j = (i_k, j_k), k = 1 \dots K. \end{aligned} \quad (20)$$

The nonnegativity constraint places the problem in the non-negative matrix factorization [50] framework, which leads to

formulating the weight computation as the constrained minimization

$$\begin{aligned} \operatorname{argmin}_{w_k, k=1 \dots N} & \left\| \Psi_{p_x}^S - \sum_k w_k \Psi_{p_j}^S \right\|_2^2 \\ \text{s.t.} & w_k \geq 0, j = (i_k, j_k), k = 1 \dots K. \end{aligned} \quad (21)$$

Once the weights w_k are computed, they are used to linearly combine the pixels of the candidates patches $\Psi_{p_j}^U$, which are located at the same position as the unknown pixels of the input patch (i.e., $\Psi_{p_x}^U = \sum_k w_k \Psi_{p_j}^U$).

MULTISCALE REFINEMENT

The search for similar patches can be improved by introducing a priori a rough estimate of the inpainted values using a multiscale approach. The missing regions are iteratively approximated using some guidance from coarse to fine levels [24]. The use of a multiscale approach offers a number of key advantages. The inpainting of a coarse version of the input picture is much easier than performing the inpainting at the full resolution. It is easier to retrieve the main and dominant structures since the local singularities in terms of orientation as well as luminance are less numerous. The results are moreover less sensitive to noise. Drori et al. [24] used, for instance, a rough estimate of the inpainted values to improve the K -NN search method. The drawback of multiscale refinement, as mentioned in [36], is that if an error occurs at a coarse scale, inpainting errors can spread across the finer scale. The authors in [22] perform the inpainting on a coarse resolution of the input picture, and then use a single-image superresolution method to retrieve the high-frequency details of the inpainted areas.

INTRODUCING SPARSITY PRIORS

The inpainting problem can also be solved assuming image sparsity priors. In this case, the image \mathbf{I} is assumed to be a sparse-land signal [51], meaning that the image \mathbf{I} is sparse in a given basis. The basis can be formed by predefined elementary waveforms (also called atoms) which are stored in a dictionary matrix \mathbf{A} . The dictionary matrix \mathbf{A} can also be learned using dictionary-learning methods. An image \mathbf{I} is said to be k -sparse in a given basis stored in the dictionary matrix \mathbf{A} , if it can be represented by a vector \mathbf{v} having only k nonzero elements (i.e., its l_0 norm is $\|\mathbf{v}\|_0 = k$) verifying $\mathbf{I} = \mathbf{A}\mathbf{v}$. The sparsity of a signal depends on the considered basis, i.e., of the matrix \mathbf{A} .

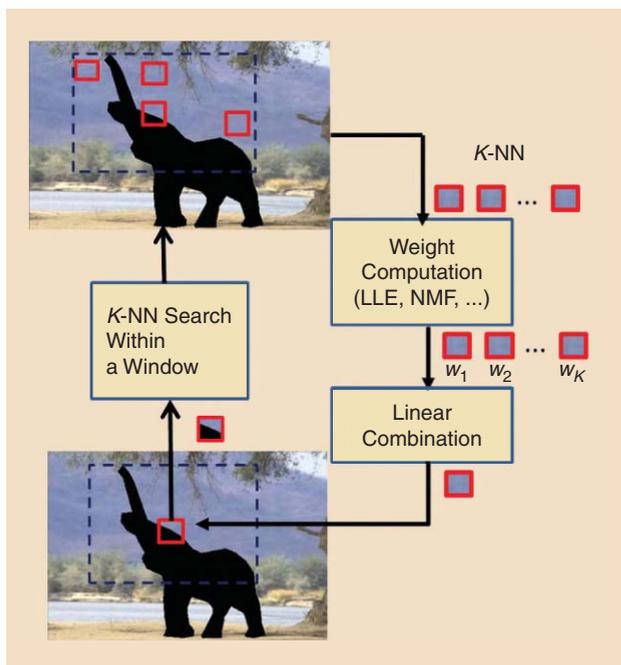
Considering the degraded image $\mathbf{F} = \mathbf{M}\mathbf{I}$, the inpainting problem is therefore formulated as searching for the sparse representation vector \mathbf{v} of the image \mathbf{F} , by solving

$$\min \|\mathbf{v}\|_0$$

such that

$$\mathbf{F} = \mathbf{M}\mathbf{A}\mathbf{v}. \quad (22)$$

Many solutions exist for searching for the sparse vector \mathbf{v} , with the most popular ones belonging to the family of greedy matching pursuit algorithms. A good overview of these matching pursuit algorithms can be found in [51, Ch. 3].



[FIG8] The estimation of unknown pixels with neighbor embedding. The first step consists in searching for the K -NN of the input patch. The algorithm then searches for the weights of the linear combination of these K -NN patches that best approximates the known pixels of the input patch to be completed. The weights computation can be done using different methods (e.g., NLM, LLE, and NMF).

The above inverse problem is usually solved patch-per-patch rather than directly on the entire image F . For each patch Ψ_{p_x} of the image F formed by a known part $\Psi_{p_x}^S$ and an unknown part $\Psi_{p_x}^U$, one searches for the sparse vector v_{p_x} , which best approximates the known part of the input patch $\Psi_{p_x}^S$ as $\Psi_{p_x}^S = A^S v_{p_x}$ where A^S is a matrix obtained by masking the rows of the matrix A corresponding to the positions of the unknown pixels $\Psi_{p_x}^U$ in Ψ_{p_x} , as shown in Figure 9.

The same sparse linear combination of atoms is then used to approximate the unknown pixels $\Psi_{p_x}^U$, as $\Psi_{p_x}^U = A^U v_{p_x}$, taking this time the masked samples of the atoms (these samples correspond to the positions of the unknown pixels). This general formulation is also used for other image processing problems like denoising and superresolution. The sparse representation area has recently evolved into the more general compressive sampling framework that also naturally applies to image recovery problems [52].

Variants have been introduced exploiting sparsity priors. Assuming that images are composed of locally uniform regions separated by edges, the author in [53] uses adaptive sparse representations. The algorithm performs a nonlinear approximation with adaptively determined sparsity constraints. Hybrid sparse representations enforcing both local and nonlocal sparsity are considered in [54]. The nonlocal sparsity is defined as the sparsity of a three-dimensional (3-D) data array formed by the input patch and its K -NN in the known part of the image, whereas the local sparsity is defined as the sparsity of the 2-D

patch, the sparsity constraints being enforced by hard-thresholding in a predefined waveform basis (DCT, fast Fourier transform). Local and nonlocal sparse representations are then combined via Bayesian model averaging [54] to satisfy both constraints of local smoothness and nonlocal similarity, with a constraint of fidelity to the known samples.

Patch-based methods show that texture patches can be relevant dictionary elements. Therefore, instead of using predefined waveforms, a linear combination of candidate patches regularized by a sparseness prior on the weighting coefficients can be used for inferring the unknown pixels [32]. Sparsity is also used in [32] for determining structural patches to be processed first, as explained in the section "Patch Processing Order." The patch sparse representation is moreover constrained by local patch consistency.

HYBRID METHODS SEPARATING STRUCTURE FROM TEXTURE

Diffusion methods work well for small and sparsely distributed gaps. They are also appropriate for piecewise smooth images and for propagating strong structures. But they are unable to restore texture. On the contrary, exemplar-based methods work amazingly well in textured regions with homogeneous or regular patterns. Nevertheless, they are not so well suited for preserving edges or structures, or for images with many small distributed holes.

Yet, natural images contain composite structures and textures. The structures constitute primal sketches of an image (e.g., edges, corners) and textures are regions with homogeneous patterns or feature statistics. To handle composite textures and structures, it is therefore natural to combine different types of approaches. Two main strategies have been considered. The first strategy consists in first separating the image components (texture and structure), and inpainting them separately with the most suitable method (e.g., diffusion or exemplar-based). The two inpainted components are then added together as in [55], [56]. A second strategy consists in combining different approaches in one unique energy function using a variational formulation [21], [36].

STRUCTURE/TEXTURE SEPARATION

Structure can be identified in a supervised way, as in [57], where the user specifies curves corresponding to important missing structures (e.g., object boundaries) in the unknown region. A structure propagation is then performed by copying patches located in the direction of these curves in the known region. The remaining unknown pixels are in a third step estimated using a texture synthesis method as in [19].

Texture and structure can also be separated in an automatic manner, using, for example, a variational method as in [55], where the authors decompose the image as a sum of two functions, one being of BV representing the image structure, and a second one capturing the texture. The structure image is a sketchy approximation of the input image, containing only edges separating smooth regions. These piecewise smooth

images are also referred to as cartoon images in [55]. In [55], the texture layer is inpainted using the texture synthesis method of [3] while the geometric layer is inpainted using the diffusion method of [2], and the two inpainted components are added together to produce the final result.

The authors in [56] proposed a method based on sparse representations for decomposing the image into a texture and a geometry component, called layers. Using two dictionaries of different characteristics A_g and A_t , the image is decomposed into structural and textural components as $I = A_g v_g + A_t v_t$. The method is called morphological component analysis (MCA). The two dictionaries are mutually incoherent, i.e., each dictionary gives a sparse representation for one component while yielding a nonsparse representation for the other component. Both dictionaries are grouped into a big dictionary which is then used by a basis pursuit algorithm to find the sparse representation of each layer. The authors in [58] propose an algorithm based on this sparsity seeking image separation into two components. Instead of a separate processing of the two components as in [55], the sparse vectors for the two components are obtained by minimizing

$$\min \|v_g\|_p + \|v_t\|_p$$

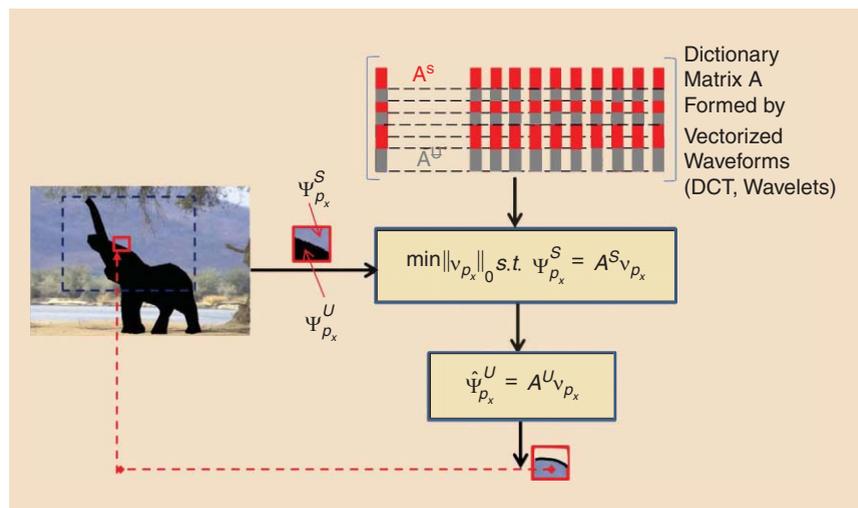
such that

$$F = M(A_g v_g + A_t v_t), \quad (23)$$

where $\|\cdot\|_p$ denotes the L_p norm, with p often equal to zero or one. To solve this minimization problem, the constraint is introduced as a penalization term. In [58], a TV penalization term is added to regularize the sparse approximation of the image geometry. This approach can fill in a region with composite textures and structures. It however introduces blur when the missing region is large.

COMBINING DIFFERENT METHODS VIA ONE ENERGY TERM

Methods best adapted to different components of the image can also be combined by introducing several energy terms in one unique energy function which can then be globally optimized, without prior separation of structure from texture. In [36], an energy term comprising a texture synthesis term and a term measuring how patches to be stitched together agree in the overlap area is minimized using belief propagation. The authors in [21] combine energy terms related to texture synthesis, coherence, and geometry (by minimizing the TV of the structure of the image) as proposed in [59], into one single energy functional. A correspondence map between pixels to be filled in



[FIG9] Estimation of unknown pixels with sparse priors in transform domain. A dictionary matrix is constructed from waveforms (DCT, wavelets). The grey part of the columns of the dictionary represents the masked rows corresponding to the position of the unknown pixels in the input patch.

and pixels in the known part of the image, is then searched to minimize this energy functional which is the sum of three energy measures (self-similarity, coherence, and diffusion). The self-similarity energy term, as in texture synthesis, computes the similarity between the patch centered on the pixel to be filled in and the patch centered on the candidate pixel in the known part of the image. The diffusion term is the energy of the discrete Laplacian of the inpainted part of the image for a given correspondence map. The spatial coherence term measures the similarity between patches corresponding to neighboring pixels.

GLOBAL METHODS

In ill-posed image processing problems such as inpainting, image priors play a very critical role. Statistical and structural priors capture and exploit stationarity and similarity in a local neighborhood or throughout the entire image. In patch-based methods, the missing pixels are patch-wise computed in terms of their neighbors, whereas in diffusion pixel-based methods, they are filled in by propagating neighboring pixels in a way that favors good edge continuation. Rather than searching to capture local relationships or dependencies, one can instead capture the global structure of the input data, using models which reproduce key statistical properties of images or of textures of interest. These models include probabilistic models of coefficients in transformed domains [60], sparsity models (see the section “Introducing Sparsity Priors”) or low-rank models.

When using sparse priors, the input texture or image is assumed to have sparse representations in a certain basis. Forcing sparse priors on the recovered image, the missing region is synthesized as a sparse linear combination of elements from an overcomplete dictionary. Similarly, given the spatial coherence and self-similarity which characterize natural images, the high-dimensional input textures or images lie in a subspace of reduced dimension. This low-rank image model has led to a

class of methods based on low-rank matrix and tensor completion, which aim at representing this subspace characterizing the “global” information of the image.

Let I be the input image to be completed. Let us assume that the image I is of dimension $P \times Q$ in which each pixel in position $\mathbf{x} = (x, y)$ carries three color components (i.e., $I \in \mathcal{R}^{P \times Q \times 3}$). Each image color channel of I forms a 2-D matrix I^c with known and unknown entries. The matrix completion problem is to search for a low-rank approximation of the matrix I^c , given the known subset of its entries. This problem is formulated as

$$\min_{\hat{I}^c} \frac{1}{2} \|\hat{I}^c - I^c\|_S \text{ s.t. } \text{rank}(\hat{I}^c) \leq r, \quad (24)$$

where S is the set of known pixels in the image (known entries in the image), and the rank is the number of nonvanishing singular values of the matrix. This optimization is not convex. The problem can however be reformulated as a convex programming problem by minimizing the sum of the singular values, i.e., the nuclear (or trace) norm instead of the rank [61]. The problem is in this case of finding the matrix with the minimum nuclear norm agreeing with the observed entries.

Given that a tensor is simply a generalization of a matrix to higher dimensions, the low-rank matrix completion problem has been naturally extended to low-rank tensor completion. The order of a tensor is the number of dimensions also known as ways or modes. The three color component input image I can therefore be seen as a three-order tensor to be completed. The problem can be formulated in a similar manner as in (24), and solved using the notion of trace norm for tensors as in [62].

Matrix and tensor completion methods work well for inpainting when the missing areas are not too large or when the rank of the original image is quite low. It is hence suitable for restoring images from scratches, or for removing overlaid text, but not so well suited for applications such as disocclusion, object removal or loss concealment.

APPLICATIONS

The problem of inpainting is encountered in various image processing applications: image restoration, editing (e.g., object removal), disocclusion in image-based rendering, interpolation, loss concealment, texture synthesis or image resizing (e.g., enlargement). Inpainting has also been considered in the context of lossy image compression: blocks within the image that can be recovered by inpainting are not transmitted. The goal

here is not to assess or benchmark in terms of inpainting results the numerous methods which exist, nor to give an exhaustive list of all potential applications. It is instead to illustrate the main applications with some examples of algorithms, showing the limited applicability of some of them for particular use cases. A taxonomy of the methods is given in Table 2. Links to publicly available software code implementing some of the above methods are given in Table 3.

IMAGE RESTORATION

The image restoration problem is concerned with recovering an original image from various forms of degradations. The origin of the degradations to deal with depends on the application: it can be text overlay or scratches in digital photography, in digital cinema, or in pictures taken of ancient paintings [63]. It can also be degradations that result from the capturing process like specular reflections, spots, and cracks in medical images (e.g., in endoscopic images [64]). Another application area is fingerprint restoration in automatic fingerprint identification systems [65].

In the restoration problem, the missing region is generally not too large, hence, local diffusion and patch-based or global methods give satisfactory results. Figure 10 illustrates the problem and gives inpainting results obtained with methods of different types: anisotropic diffusion [11], exemplar-based inpainting, the hybrid method in [21], which performs a global minimization of an energy term, and the tensor completion method [62]. The exemplar-based method used here is a simple approach computing the patch processing order as in [31], but using similarity weights [47] to combine candidate patches rather than inferring the unknown pixels from the single best match. Given the small gaps to be filled in, most methods (diffusion, exemplar, a fortiori combinations of the two, as well as more global techniques) give satisfactory results.

OBJECT REMOVAL

Another natural application of inpainting is image editing in which the user removes objects, hence, uncovering unknown parts of the image foreground. This application is well illustrated by the images of Figure 11 in which one foreground object has been removed, leaving a hole to be filled. Figure 11 shows inpainting results with methods of different categories. Figure 11(b) illustrates the limitations of diffusion methods when the gap to be filled in is large. Diffusion introduces smoothing and blurring artifacts in the synthesized region.

[TABLE 2] A TAXONOMY OF INPAINTING METHODS.

FEATURES	PDE-BASED DIFFUSION	EXAMPLAR-BASED INPAINTING	HYBRID METHODS	GLOBAL
PRIORS	SMOOTHNESS	SELF-SIMILARITY, SPARSITY	SMOOTHNESS + SIMILARITY/SPARSITY	STATISTICAL, LOW RANK
OPTIMIZATION	GREEDY	GREEDY OR GLOBAL	GREEDY OR GLOBAL	GLOBAL
SENSITIVITY TO SETTING	LOW	HIGH	HIGH	HIGH
HOLES	SMALL	MEDIUM TO BIG	MEDIUM TO BIG	SMALL TO MEDIUM
APPLICATIONS	RESTORATION	RESTORATION, EDITING, DISOCCLUSION, CONCEALMENT	RESTORATION, EDITING DISOCCLUSION, CONCEALMENT	RESTORATION

To recover the texture of the hole, exemplar-based methods [66], methods using sparse representations [32], or solutions combining structure diffusion and exemplar-based texture recovery as, e.g., [21] are more appropriate. However, exemplar-based techniques, although conceptually very simple, work strikingly well for these type of applications. The method used in Figure 11(c) uses the same patch processing order as in [31] but performs a linear combination of best matching patches with weights computed using the constrained least squares approximation given in (20). The method in [39], based on statistics of patch offsets, also gives very good results in this case. However, global methods based on low-rank approximation or tensor-based completion [62] are, in general, not suitable for this type of application, except if the hole is small or the input image is low rank.

DISOCCLUSION

Inpainting methods are also needed in 3DTV rendering on stereoscopic or autostereoscopic displays, as well as in the context of free viewpoint rendering of a 3-D scene. For the user to navigate in the 3-D scene, virtual views that may be distant from original ones must be synthesized. The intermediate views (or virtual views) are usually computed with IBR algorithms that use the original images, scene geometry information (depth

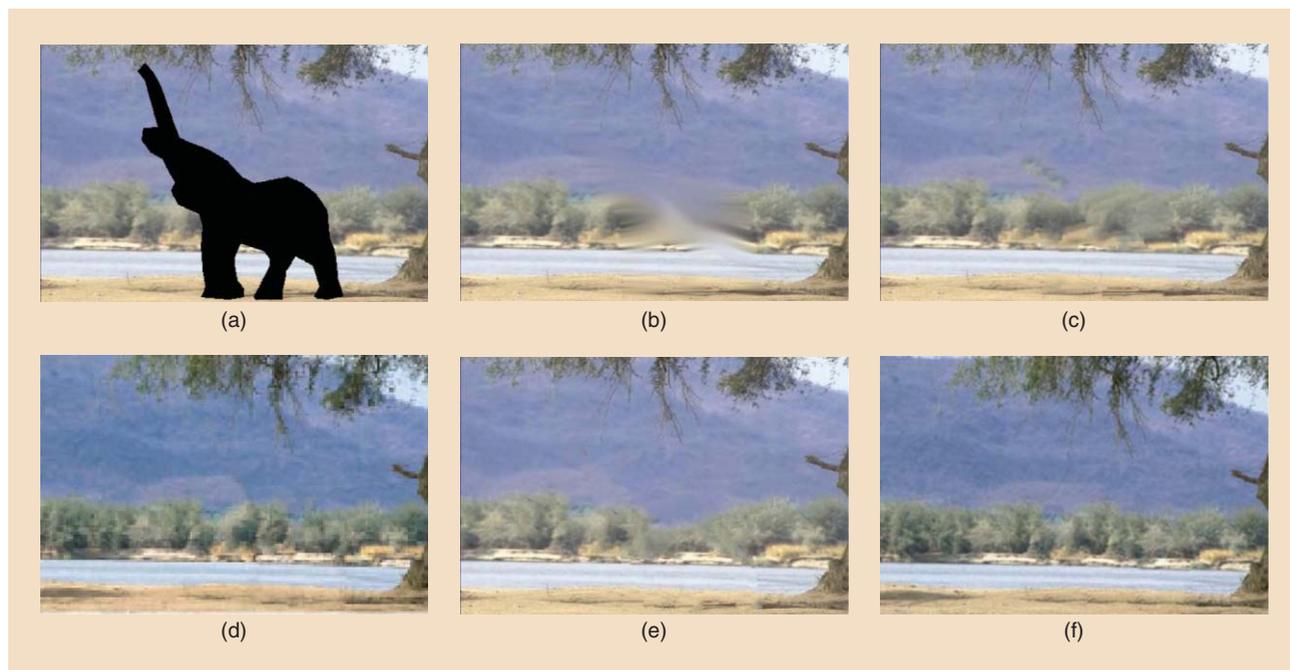
[TABLE 3] LINKS TO PUBLICLY AVAILABLE CODE.

METHODS	URL
PDE-BASED DIFFUSION	FAST MARCHING DIFFUSION [10] (IN OPENCV): http://docs.opencv.org/modules/photo/doc/inpainting.html
EXAMPLAR BASED	THIRD PARTY CODE OF [31]: http://www.csee.wvu.edu/~xinl/source.html http://daviddoria.com/?p=126 [22]: http://people.irisa.fr/Olivier.Le_Meur/publi/2013_TIP/index.html
SPARSE METHODS	[51]: http://www.cs.technion.ac.il/~elad/software/ADAPTIVE [53]: http://eeweb.poly.edu/~onur/source.html#recover_code LOCAL/NONLOCAL [54]: http://www.csee.wvu.edu/~xinl/demo/inpainting.html TEXTURE/STRUCTURE SEPARATION (MCA) [56]: http://jstarc.free.fr/mca.html
GLOBAL/HYBRID METHODS	TENSOR COMPLETION [62]: http://pages.cs.wisc.edu/~ji-liu/

maps), and camera parameters. A 3-D point on the reconstructed geometry is projected onto the image plane of a virtual camera. The color of the 3-D point is first computed by inverse projection of the original image into the 3-D space. During this projection process, some parts of the 3-D scene may be visible in the virtual view but not in the original views: they may be



[FIG10] Image restoration. (a) The original image (courtesy of [62]), (b) mask, (c) inpainting results with anisotropic diffusion, (d) exemplar-based method with similarity weights (NLM), (e) with the hybrid method (courtesy of [21]), and (f) with the tensor completion technique (courtesy of [62]).



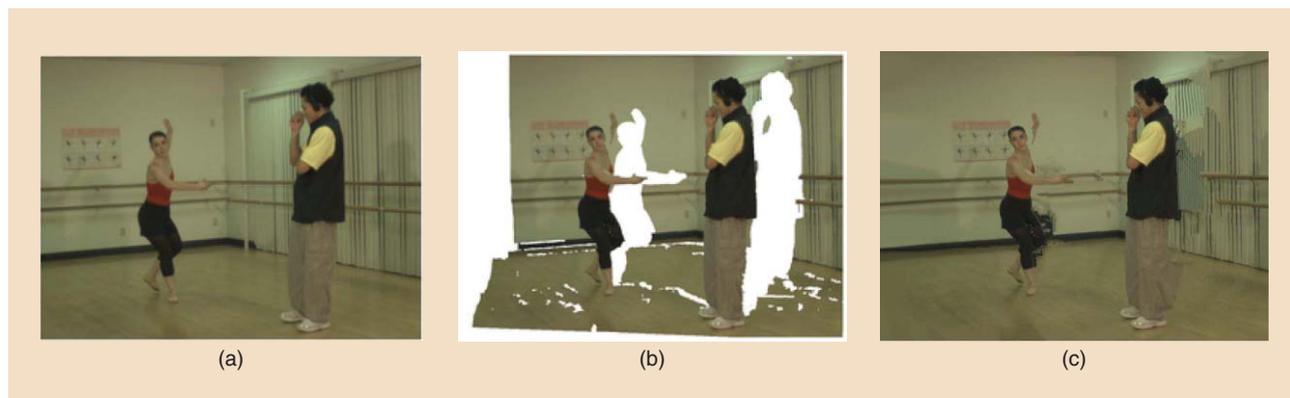
[FIG11] Object removal application (a) mask and inpainting results with methods from different categories, (b) anisotropic diffusion (courtesy of [11]), (c) exemplar-based with LLE (courtesy of [66]), (d) patch sparse representation (courtesy of [32]), (e) hybrid with one global energy minimization (courtesy of [21]), and (f) patch offsets (courtesy of [39]). [(a) courtesy of www.magazinehive.com.]

hidden by foreground objects. When synthesizing a virtual view, these parts become disoccluded, resulting in pixels with unknown color. These pixels need to be estimated using inpainting techniques.

Figure 12 illustrates this application by showing in (a) a reference view (view four of the multiview sequence called ballet (see [67]) and in (c) the result of the projection on a virtual viewpoint corresponding to a displacement to the right of the camera. Disoccluded areas due to the projection appear and have to be filled in. When the camera is moving to the right, missing areas (e.g., disocclusion areas) appear on the right side of foreground objects. Therefore, to prevent the propagation of foreground patches into the background, the filling has to be performed from the right to

the left side. Otherwise for a leftward shift of the camera, the missing areas should be filled in from the left to the right side. Figure 12(c) shows the inpainted result using the exemplar-based method in [66] adapted so that the candidate patches are searched in a window shifted according to the camera displacement (i.e., shifted to the right for a rightward camera shift).

The gaps to be filled in do not result from explicit object removal as in an image editing application. However, the difficulties and characteristics, as far as the inpainting algorithm is concerned, are very similar for both applications. In this application, the inpainting can nevertheless also benefit from taking into account the depth information in addition to the three color components, e.g., for finding best matching patches.



[FIG12] The view synthesis for the sequence ballet. (a) The known reference view projected into virtual viewpoints. (b) Projection of the reference view on a virtual viewpoint with rightward shift of the camera (white areas are disoccluded areas). (c) The inpainted version of the projected image. (Images used courtesy of [67].)

LOSS CONCEALMENT

Image and video transmission over best-effort packet-based networks suffers from packet losses that result in missing areas in the decoded images. The loss concealment is a postprocessing performed after decoding that searches to recover lost parts of an image by exploiting dependencies within the image or between adjacent images in a video sequence. The positions of the missing areas depend on the data packetization scheme. Transmission schemes, such as the flexible macroblock ordering, avoiding placing adjacent blocks of pixels in the same packet, have been specified in the H.264 standard, to ease the concealment process in case of packet losses.

In transmission and decoding systems that are practically used today, the loss concealment is usually performed using a simple copy from previous frames that have been correctly received or using simple spatiotemporal interpolation. However, inpainting techniques are useful tools in this context as well, although complexity issues become quite crucial in this context, where the inpainting must be done in real time. This application is illustrated in Figure 13, where the inpainting is done using the hybrid local/nonlocal sparse representation method. As far as the inpainting methods are concerned, the problem is similar to the object removal application, therefore, in this case as well, patch-based methods using exemplars or sparse priors, or hybrid methods combining diffusion and exemplar-based techniques give very good results.

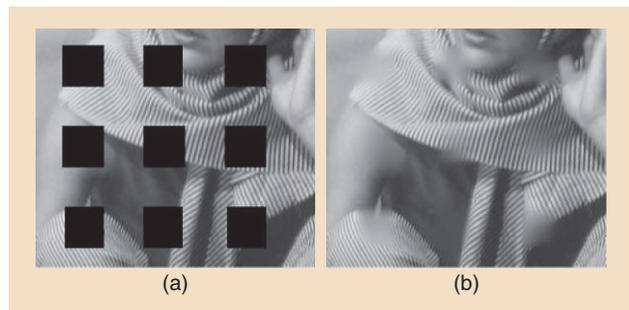
CONCLUSIONS

Image inpainting has received a lot of attention in the past few years. Numerous and different types of approaches have been proposed with varying applicability in restoration, object removal, disocclusion, or in texture synthesis. These algorithms have, however, limited direct applicability for video inpainting, which remains an open problem, despite preliminary solutions making assumptions in terms of moving objects or camera motion. Tracking moving objects in a video (the ones to be removed or others impacted by missing data due to losses or occlusions) in an unsupervised manner remains a difficult problem.

The quality assessment of inpainted images is another open and difficult issue, as no quantitative metrics exists. Fidelity metrics cannot be used given that, for most inpainting applications, the ground truth is in general unknown. One has to rely on a subjective assessment to evaluate whether the inpainted images are visually pleasing and physically plausible.

AUTHORS

Christine Guillemot (Christine.Guillemot@inria.fr) is the director of research at the Institut National de Recherche en Informatique et en Automatique and head of a research team dealing with image and video modeling, processing, coding, and communication. She received her Ph.D. degree from Ecole Nationale Supérieure des Télécommunications, Paris, and a "Habilitation for Research Direction" from the University of Rennes. From 1985 to 1997, she was with France Telecom, where she was involved in



[FIG13] An illustration of the loss concealment problem. The lost areas of the image are recovered with the hybrid local/nonlocal sparse representation [54].

various projects in the area of image and video coding for television, high-definition television, and multimedia. From 1990 to 1991, she was a visiting scientist at Bellcore, New Jersey. She has (co)authored 15 patents, eight book chapters, 50 journal papers, and 140 conference papers. She has been an associate editor of *IEEE Transactions on Image Processing* (2000–2003), *IEEE Transactions on Circuits and Systems for Video Technology* (2004–2006), and *IEEE Transactions on Signal Processing* (2007–2009). She is currently an associate editor of *EURASIP's Journal on Image Communication* and is a member of the editorial board of *IEEE Journal on Selected Topics in Signal Processing* (2013–2015). She is a member of the IEEE Image, Video, and Multidimensional Signal Processing Technical Committees. She is a Fellow of the IEEE.

Olivier Le Meur (olivier.le_meur@irisa.fr) received his Ph.D. degree from the University of Nantes in 2005. From 1999 to 2009, he worked in the media and broadcasting industry. In 2003, he joined the research center of Thomson-Technicolor at Rennes, where he supervised a research project concerning the modeling of the human visual attention. He has been an associate professor of image processing at the University of Rennes 1 since 2009. He is a part of the IRISA/TEMICS team, which includes computational modeling of the visual attention and saliency-based applications (video compression, objective assessment of video quality, and retargeting). His research interests are dealing with the understanding of the human visual attention.

REFERENCES

- [1] S. Masnou and J. Morel, "Level-lines based disocclusion," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Chicago, IL, Oct. 1998, vol. 3, pp. 259–263.
- [2] M. Bertalmio, G. Sapiro, C. Ballester, and V. Caselles, "Image inpainting," in *Proc. ACM SIGGRAPH*, July 2000, pp. 417–424.
- [3] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *Proc. Int. Conf. Computer Vision (ICCV)*, Sept. 1999, pp. 1033–1038.
- [4] S. Di Zenzo, "A note on the gradient of a multi-image," *Comput. Vision, Graphics, Image Process.*, vol. 33, no. 1, pp. 116–125, Jan. 1986.
- [5] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, 1990.
- [6] F. Voci, S. Eiho, N. Sugimoto, and H. Sekiguchi, "Estimating the gradient threshold in the Perona-Malik equation," *IEEE Signal Processing Mag.*, vol. 21, no. 3, pp. 39–46, May 2004.
- [7] J. Weickert, "Theoretical foundations of anisotropic diffusion in image processing," *Comput. Suppl.*, vol. 11, pp. 221–236, 1996.
- [8] J. Weickert, "Coherence-enhancing diffusion filtering," *Int. J. Comput. Vis.*, vol. 32, no. 2–3, pp. 111–127, 1999.

- [9] M. Bertalmio, A. Bertozzi, and G. Sapiro, "Navier-stokes, fluid dynamics, and image and video inpainting," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Dec. 2001, pp. 355–362.
- [10] A. Telea, "An image inpainting technique based on the fast marching method," *J. Graphics Tools*, vol. 9, no. 1, pp. 25–36, 2004.
- [11] D. Tschumperlér, "Fast anisotropic smoothing of multi-valued images using curvature-preserving PDE's," *Int. J. Comput. Vis.*, vol. 68, no. 1, pp. 65–82, 2006.
- [12] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, Nov. 1992.
- [13] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling in by joint interpolation of vector fields and gray levels," *IEEE Trans. Image Processing*, vol. 10, no. 8, pp. 1200–1211, Aug. 2001.
- [14] T. Chan and J. Shen, "Local inpainting models and TV inpainting," *SIAM J. Appl. Math.*, vol. 62, no. 3, pp. 1019–1043, 2001.
- [15] A. Levin, A. Zomet, and Y. Weiss, "Learning how to inpaint from global image statistics," in *Proc. Int. Conf. Computer Vision (ICCV)*, Oct. 2003, pp. II. 305–313.
- [16] T. Chan and J. Shen, "Non texture inpainting by curvature-driven diffusion," *J. Visual Commun. Image Representation*, vol. 12, no. 4, pp. 436–449, 2001.
- [17] T. Chan, S. Kang, and J. Shen, "Euler's elastica and curvature-based inpainting," *SIAM J. Appl. Math.*, vol. 63, no. 2, pp. 564–592, 2003.
- [18] L. Wei and M. Levoy, "Fast texture synthesis using tree-structured vector quantization," in *Proc. ACM SIGGRAPH*, July 2000, pp. 479–488.
- [19] M. Ashikhmin, "Synthesizing natural textures," in *Proc. ACM Symp. Interactive 3D Graphics*, Mar. 2001, pp. 217–226.
- [20] L. Liang, C. Liu, Y. Xu, B. Guo, and H. Shum, "Real-time texture synthesis by patch-based sampling," *ACM Trans. Graph.*, vol. 20, no. 3, pp. 127–150, 2001.
- [21] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro, "A comprehensive framework for image inpainting," *IEEE Trans. Image Processing*, vol. 19, no. 10, pp. 2634–2645, Oct. 2010.
- [22] O. Le Meur and C. Guillemot, "Super-resolution-based inpainting," in *Proc. European Conf. Computer Vision (ECCV)*, Oct. 2012, pp. 554–567.
- [23] Y. Zhang, J. Xiao, and M. Shah, "Region completion in a single image," in *EURO-GRAPHICS*, Sept. 2004.
- [24] I. Drori, D. Cohen-Or, and H. Yeshurun, "Fragment-based image completion," *ACM Trans. Graph.*, vol. 22, no. 2003, pp. 303–312, 2005.
- [25] J. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, Sept. 1975.
- [26] P. Yanilos, "Data structures and algorithms for nearest neighbor search in general metric spaces," in *Proc. 4th Annu. ACM/SIGACT-SIAM Symp. Discrete Algorithms (SODA)*, 1993, pp. 311–321.
- [27] N. Kumar, L. Zhang, and S. K. Nayar, "What is a good nearest neighbors algorithm for finding similar patches in images?" in *Proc. European Conf. Computer Vision (ECCV)*, Oct. 2008, pp. 364–378.
- [28] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "Patchmatch: A randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 24:1–24:11, July 2009.
- [29] S. Korman and S. Avidan, "Coherency sensitive hashing," in *Proc. Int. Conf. Computer Vision (ICCV)*, 2011, pp. 1607–1614.
- [30] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein, "The generalized PatchMatch correspondence algorithm," in *Proc. European Conf. Computer Vision (ECCV)*, Sept. 2010, vol. 6313, pp. 29–43.
- [31] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based inpainting," *IEEE Trans. Image Processing*, vol. 13, no. 9, pp. 1200–1212, Sept. 2004.
- [32] Z. Xu and J. Sun, "Image inpainting by patch propagation using patch sparsity," *IEEE Trans. Image Processing*, vol. 19, no. 5, pp. 1153–1165, May 2010.
- [33] O. Le Meur, J. Gautier, and C. Guillemot, "Exemplar-based inpainting based on local geometry," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Sept. 2011, pp. 3401–3404.
- [34] G. Kanizsa, *Organization in Vision*. New York: Praeger, 1979.
- [35] Y. Wexler, E. Shechtman, and M. Irani, "Space-time video completion," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, June 2004, vol. 1, pp. I-120–I-127.
- [36] N. Komodakis and G. Tziritas, "Image completion using global optimization," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2006, pp. 442–452.
- [37] Y. Pritch, E. Kav-Venaki, and S. Peleg, "Shift-map image editing," in *Proc. Int. Conf. Computer Vision (ICCV)*, Kyoto, Sept. 2009, pp. 151–158.
- [38] T. Cho, M. Butman, S. Avidan, and W. Freeman, "The patch transform and its applications to image editing," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, June 2008, pp. 1–8.
- [39] K. He and J. Sun, "Statistics of patch offsets for image completion," in *Proc. European Conf. Computer Vision (ECCV)*, Oct. 2012, vol. 7773, pp. 16–29.
- [40] A. Efros and W. Freeman, "Image quilting for texture synthesis and transfer," in *Proc. Annu. Conf. Computer Graphics and Interactive Techniques, ACM SIGGRAPH*, July 2001, pp. 341–346.
- [41] S. Avidan and A. Shamir. (2007). Seam carving for content-aware image resizing. *Proc. ACM SIGGRAPH*. [Online]. Available: <http://doi.acm.org/10.1145/1275808.1276390>
- [42] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.
- [43] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)*, vol. 26, no. 2, pp. 147–159, Feb. 2004.
- [44] V. Kwatra, A. Schodl, I. Essa, G. Turk, and A. Bobik, "Graphcut textures: image and video synthesis using graph cuts," in *ACM SIGGRAPH*, July 2003, vol. 22, pp. 277–286.
- [45] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," in *ACM SIGGRAPH*, 2003, pp. 313–318.
- [46] S. Darabi, E. Shechtman, C. Barnes, D. B. Goldman, and P. Sen, "Image melding: Combining inconsistent images using patch-based synthesis," *ACM Trans. Graph. (TOG) (Proc. SIGGRAPH 2012)*, vol. 31, no. 4, pp. 82:1–82:10, 2012.
- [47] A. Wong and J. Orchard, "A nonlocal-means approach to exemplar-based inpainting," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, 2006, pp. 2600–2603.
- [48] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *SIAM J. Multi Scale Modeling Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
- [49] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, Dec. 2000.
- [50] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Adv. Neural Inform. Process. Syst. (NIPS)*, pp. 556–562, Nov. 2000.
- [51] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York: Springer, 2010.
- [52] C. Studer, P. Kuppinger, G. Pope, and H. Bolckei, "Recovery of sparsely corrupted signals," *IEEE Trans. Inform. Theory*, vol. 58, no. 5, pp. 3115–3130, May 2012.
- [53] O. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising—Part II: Adaptive algorithms," *IEEE Trans. Image Processing*, vol. 15, no. 3, pp. 555–571, Mar. 2006.
- [54] X. Li, "Image recovery via hybrid sparse representations: A deterministic annealing approach," *IEEE J. Select. Topics Signal Process.*, vol. 5, no. 5, pp. 953–962, Sept. 2011.
- [55] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Processing*, vol. 12, no. 8, pp. 882–889, Aug. 2003.
- [56] J. Stark, M. Elad, and D. Donoho, "Image decomposition via the combination of sparse representations and variational approach," *IEEE Trans. Image Processing*, vol. 14, pp. 1570–1582, Oct. 2005.
- [57] J. Sun, L. Yuan, J. Jia, and H. Shum, "Image completion with structure propagation," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 861–868, 2005.
- [58] M. Elad, J.-L. Stark, P. Querre, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *J. Appl. Comput. Harmonic Anal.*, vol. 19, pp. 340–358, Nov. 2005.
- [59] J. Aujol, S. Ladjal, and S. Masnou, "Exemplar-based inpainting from a variational point of view," *SIAM J. Math. Anal.*, vol. 42, no. 3, pp. 1246–1285, Mar. 2010.
- [60] J. Portilla, E. Simoncelli, and Y.-J. Zhang, "A parametric texture model based on joint statistics of complex wavelets," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 49–71, 2000.
- [61] E. Candes and B. Recht, "Exact matrix completion via convex optimization," *Commun. ACM*, vol. 55, no. 6, pp. 111–119, 2012.
- [62] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.
- [63] S.-C. Pei, Y.-C. Zeng, and C.-H. Chang, "Virtual restoration of ancient Chinese paintings using color contrast enhancement and lacuna texture synthesis," *IEEE Trans. Image Processing*, vol. 13, no. 3, pp. 416–429, Mar. 2004.
- [64] M. Arnold, A. Ghosh, S. Ameling, and G. Lacey, "Automatic segmentation and inpainting of specular highlights for endoscopic imaging," *EURASIP J. Image Video Process.*, vol. 2010, article 9, ID 814319, pp. 1–12, Dec. 2010.
- [65] F. Bornemann and T. Marz, "Fast image inpainting based on coherence transport," *J. Math. Imaging Vis.*, vol. 28, no. 3, pp. 259–278, July 2007.
- [66] C. Guillemot, M. Turkan, O. L. Meur, and M. Ebdelli, "Image inpainting using LLE-LDNR and linear subspace mappings," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013.
- [67] C. Zitnick, S. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality video view interpolation using a layered representation," in *ACM SIGGRAPH*, 2004, pp. 600–608.

Kiho Choi and Euee S. Jang

Royalty-Free Video Coding Standards in MPEG

On 7 March 2013, the Moving Picture Experts Group Licensing Association (MPEG LA) and Google announced that they have entered into an agreement granting Google a license to techniques, if the patents in MPEG LA might be essential to VP8. Under this agreement, hardware and software companies are free to use the VP8 technology when developing their own products. Considering that it is now common to find patent disputes in headline news, the patent issues related to video coding standards are no exception. In this article, we report on the recent developments in royalty-free codec standardization in MPEG, particularly Internet video coding (IVC), Web video coding (WVC), and video coding for browser, by reviewing the history of royalty-free standards in MPEG and the relationship between standards and patents.

MPEG STANDARDS AND ISO/IEC PATENT POLICY

MPEG has been developing video coding standards under the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) for the past 20 years, with MPEG-2 and MPEG-4 being very successful examples. Like other ISO/IEC standards, MPEG standardization is not directly related to patents and licensing issues. Standardization groups do not have any direct responsibility for patent issues, but they provide guidance with respect to patent policy to deal with intellectual property rights (IPR) issues. All patent issues including standardized technology should be managed by

outside organizations (e.g., the World Intellectual Property Organization). ISO/IEC standards can be categorized into three types with regard to patent licensing, depending on the patent holders' licensing policy [1]; see Figure 1.

If a patent holder decides on Type-3, the technology related to the patent would not be included in a standards specification. So, most patent holders in ISO/IEC standardization choose Type-1 or Type-2. Joint Photographic Experts Group (JPEG) is a typical example of a Type-1 standard, whereas MPEG is a Type-2 standard.

THE MARKET DEMAND FOR ROYALTY-FREE CODECS LED MPEG TO INVESTIGATE THE FEASIBILITY OF ROYALTY-FREE CODEC STANDARDIZATION IN THE MPEG PORTFOLIO.

While Type-1 standards can be used royalty free in principle, they can be vulnerable to an attack by submarine patents. Submarine patents indicate the patents that are generally hidden or not considered during standard development. Patent suits on JPEG are a well-known example of such a case [2].

Type-2 standards are considered a useful compromise between patent holders and licensees. The guaranteed advantage of a Type-2 standard is to provide IPR with fair, reasonable, and non-discriminatory conditions and pricing to market. To maximize the merit of Type-2 standards, managing licenses under a patent pool is a common approach.

LICENSING STANDARD PATENTS

A patent pool is a consortium of at least two companies agreeing to cross-license patents related to a particular technology. The creation of a patent pool can save both owners and licensees time and money. The patent pool provides a service that brings all parties together by enabling one-stop shopping. MPEG LA is a well-known patent pool that manages patents of many MPEG standards, including MPEG-2 and MPEG-4 [3].

The formation of a licensing patent pool for standards is a desirable solution to licensees who want to use the standards. However, some patent holders may want to pursue royalties outside of the patent pool, which will make it more challenging for standard users to license the technologies. AT&T is one example of a nonparticipating patent holder in MPEG-LA, who is believed to have several essential MPEG-4 video compression patents [4]. In such a case, licensees need to enter additional licensing contracts with patent holders who are not part of the pool. This situation is not favorable to licensees since there is a possibility that the licensing cost for the entire standard may increase.

The licensing business of standard patents in MPEG is not expected to be as lucrative as it used to be. While the size of the digital TV market has increased substantially, the number of patents has increased exponentially [5]. According to the analysis of issued U.S. patents including terms "video coding" or "video compression" on United States Patent and Trademark Office (USPTO) by following the same procedure used in [5], the number of related patents issued in 2012 is similar to that issued in the entire 1990s decade (Figure 2). The business model that enables earning a reasonable profit from patented technology should be reconsidered.

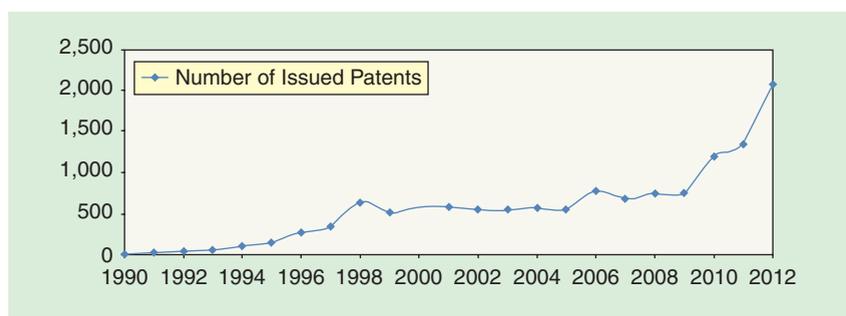
Digital Object Identifier 10.1109/MSP.2013.2282413

Date of publication: 5 December 2013

standards **IN A NUTSHELL** continued

- Type-1: "The patent holder is willing to negotiate licences free of charge with other parties on a nondiscriminatory basis on reasonable terms and conditions. Such negotiations are left to the parties concerned and are performed outside ITU-T/ITU-R/ISO/IEC."
- Type-2: "The patent holder is willing to negotiate licences with other parties on a nondiscriminatory basis on reasonable terms and conditions. Such negotiations are left to the parties concerned and are performed outside ITU-T/ITU-R/ISO/IEC."
- Type-3: "The patent holder is not willing to comply with the provisions of either paragraph type 1 or paragraph type 2; in such case, the Recommendation/Deliverable shall not include provisions depending on the patent."

[FIG1] Three types of licensing declaration for ITU-T, ITU-R, and ISO/IEC.



[FIG2] The number of issued U.S. patents including the terms "video coding" or "video compression" from 1990 to 2012.

Additionally, patent holders within a patent pool are concerned about the emergence of several competing codecs targeting the same market. Today, the competition in the video codec market is fierce: there are the conventional standard codecs in MPEG (e.g., MPEG-1, MPEG-2, and MPEG-4), along with newly introduced standard codecs [e.g., MPEG-H High-Efficiency Video Coding (HEVC)], and codecs outside MPEG (e.g., VC-1, VP8).

During the development of MPEG-2, a patent pool was the right approach because MPEG-2 was the only standard available at the time, with many different patent holders. A significant increase in the number of patents in a patent pool would result in a substantial decrease of return on investment per patent, which can be observed in many video coding standards today. Some patent holders have tried to overcome this situation by setting up an independent patent licensing program without joining a patent pool [6]. The appearance of several competing codecs makes the problem even

worse by lowering the royalty for any given codec. Most patent holders are not nonpracticing entities but manufacturers, which means that the dominance of a standard in the market is more important than the royalty from standard. From this perspective, royalty-free video coding standards are becoming a reasonable alternative to be more competitive than royalty-bearing standards.

INDUSTRY DEMAND FOR ROYALTY-FREE CODECS

In several media market segments, there are requests for royalty-free video coding standards. Standard development organizations (SDOs) have begun to consider developing standards to address these royalty-free requirements. For example, the World Wide Web Consortium (W3C), a well-known royalty-free standardization group, is trying to include a codec specification for the upcoming HTML5 for the first time. Considering that W3C has made royalty-free standards, a royalty-free technology would be the key criteria of the codec specification for HTML5.

At the same time, MPEG-1 and MPEG-2 are becoming royalty-free codecs as the essential patents on the prevalent MPEG-1 and MPEG-2 have been or are expiring. Nevertheless, the market demand for a royalty-free coding standard is not currently fulfilled because the coding performance of these and other available royalty-free codecs is not yet satisfactory. Considering the fact that the compression efficiency of a new coding standard is usually twice better than the previous one, it is natural that the expected coding performance of a new royalty-free codec should be comparable to other current coding standards.

Another option is to find a proper royalty-free codec outside of standards. In 2010, Google released the VP8 codec software under a Berkeley Software Distribution (BSD)-like license and the VP8 bitstream format specification under an irrevocable free patent license [7]. The key application of VP8 are Internet-based technologies [such as HyperText Markup Language (HTML), Hypertext Transfer Protocol (HTTP), Transmission Control Protocol (TCP), and Internet Protocol (IP)], and Google provides VP8 to be open and freely implementable based on open-source policy. VP8 is currently being proposed to several SDOs to spread the codec widely. In fact, Google submitted a document describing the bitstream specification for its VP8 video compression algorithm for publication to the Internet Engineering Task Force (IETF) [8]. Furthermore, Google is currently participating in the development of royalty-free codec standardization in MPEG [9]. Although Google puts in a great deal of efforts to standardize VP8, it is not the only card that Google is playing with. Very recently, Google released VP9 as another royalty-free codec claiming that the coding performance of VP9 is slightly better than that of HEVC [10].

DEVELOPMENT OF ROYALTY-FREE VIDEO CODING STANDARDS IN MPEG

As reviewed in the previous section, the market demand for royalty-free codecs led MPEG to investigate the feasibility of royalty-free codec standardization in the

MPEG portfolio. MPEG has started a new activity for a royalty-free-based coding standard. Currently, there are three tracks with similar goals: IVC, WVC, and video coding for browser.

IVC

In 2009, MPEG started a discussion on royalty-free coding standards, and the IVC group was formed to begin exploring this topic [11]. The IVC test model (ITM) was developed based on the MPEG-2 coding structure. The latest version of ITM is claimed to show a comparable coding performance to AVC/H.264 baseline profile [12].

WVC

Instead of developing a new coding standard for a royalty-free codec, making an existing coding standard royalty-free is also one good solution. The MPEG-4 AVC/H.264 constrained profile has been identified in MPEG as the starting point of WVC.

In its 104th meeting, MPEG requested the patent holders of MPEG-4 AVC/H264 constrained profile to declare their position with regard to their intellectual property rights (IPR) in the WVC project [15]. Considering the aim of

WVC, the declaration from patent holders is crucial to ascertain the completion of WVC as a Type-1 standard. Yet, at this point in time, it is still an open question whether all patent holders will agree on the activity and allow the use

**THE FORMATION OF
A LICENSING PATENT
POOL FOR STANDARDS
IS A DESIRABLE
SOLUTION TO LICENSEES
WHO WANT TO USE
THE STANDARDS.**

of the MPEG-4 AVC/H264 constrained profile free of charge.

Different from the conventional process in MPEG, IVC and WVC should consider the IPR issue during the development of the standard. With regard to the Type-1 standard, a proposed technology for a royalty-free coding standard should be examined carefully not only in terms of its technical merit but also from an IPR perspective. If the IPR issue of the proposed technology is not clear,

the technology should not be adopted in the standard. Currently, the issue of how to resolve the IPR issues during the standard development is under discussion in MPEG. There is a distinction between the analysis of the IPR issues by technical experts from that by patent experts. The objective of the IPR analysis within MPEG is not about the IPR itself, but more of a technical review on how a technology has evolved from one technology to another, whether the technology is published as a technical paper or registered as a patent.

VIDEO CODING FOR BROWSER

Originally, Google joined the IVC activity and proposed the inclusion of VP8. The joining of VP8 in the IVC activity resulted in a competition between VP8 and ITM. Since they have different coding structures, it was very hard to evaluate the tool-level performance of both codecs. However, owing to the competition between these two codecs, the coding performance has been improved significantly during the exploration of IVC. Based on these developments, MPEG issued an additional call for the royalty-free coding standard at the 104th MPEG meeting and established a new royalty coding standard,

[TABLE 1] TOOL-LEVEL COMPARISON OF CODECS.

TOOL	MPEG-2	ITM 6.0	AVC/H.264 (CONSTRAINED BASELINE PROFILE)	VP8	HEVC
BLOCK PARTITION SIZE	16 X 16	16 X 16	16 X 16	16 X 16	8 X 8–64 X 64
INTRAPREDICTION	DC MODE	DC MODE	FOUR MODES FOR 16 X 16, NINE MODES FOR 4 X 4	FOUR MODES FOR 16 X 16, TEN MODES FOR 4 X 4	35 MODES FOR PREDICTION UNIT
MOTION VECTOR PREDICTION	ONE NEIGHBORING BLOCK	FOUR NEIGHBORING BLOCKS	THREE NEIGHBORING BLOCKS	EIGHT NEIGHBORING BLOCKS	SIX NEIGHBORING BLOCKS
SUBPEL ACCURACY	1/2	1/2, 1/4	1/2, 1/4	1/2, 1/4, AND 1/8	1/2, 1/4, AND 1/8
INTERPREDICTION	O	O	O	O	O
P/B-FRAME PARTITION	16 X 16	16 X 16, 16 X 8, 8 X 16, 8 X 8	16 X 16, 16 X 8, 8 X 16, 8 X 8, 4 X 4	16 X 16, 16 X 8, 8 X 16, 8 X 8, 4 X 4	4 X 4–64 X 64
NUMBER OF REFERENCES	UP TO TWO	UP TO FIVE	UP TO 16	UP TO THREE	UP TO 16
TRANSFORM	FLOATING DCT 8 X 8	INTEGER DCT 8 X 8	INTEGER DCT 4 X 4	INTEGER DCT 4 X 4	INTEGER DCT 4 X 4–32 X 32, INTEGER DST 4 X 4
ENTROPY CODING	VLC	CONTEXT-BASED BAC	CONTEXT-BASED VLC	CONTEXT-BASED BAC	CONTEXT-BASED BAC
LOOP FILTER	X	X	DEBLOCKING FILTER	DEBLOCKING FILTER	DEBLOCKING FILTER, SAMPLE ADAPTIVE OFFSET

standards **IN A NUTSHELL** continued

called video coding for browser, based on VP8 [13], [14].

TECHNICAL REVIEW OF IVC, WVC, AND VIDEO CODING FOR BROWSER

Codecs under discussion for royalty-free standards are the following: MPEG-2, AVC/H.264 constrained profile, VP8, and ITM 6.0. The patents of MPEG-2 are expiring, AVC/H.264 constrained profile is being discussed for WVC activity, ITM 6.0 is the candidate for the IVC activity, and VP8 is a basic structure for video coding for browser. Considering that the video coding structure has not changed since MPEG-1, the general coding structure of the above codecs is similar to one another. The difference between codecs can be found at the tool level. MPEG-2 and MPEG-H HEVC represent two extremes here: If a tool is quite similar to MPEG-2, we can consider it close to royalty free. On the other hand, a tool should be carefully analyzed if the tool is closer to HEVC than to MPEG-2. Table 1 shows a tool-level summary of codecs under discussion.

BLOCK PARTITION

The block partition size has evolved from a fixed size to a flexible size. HEVC supports variable partition sizes while the other codecs use only 16×16 as the partition size. From an IPR aspect, the fixed 16×16 size can be considered safe to be used in royalty-free coding standards

INTRAPREDICTION

The development of intraprediction has increased the number of intramodes by supporting various directions. Although ITM 6.0 uses only direct current mode, more modes should be supported for further improving coding efficiency.

MOTION VECTOR PREDICTOR (IN INTERPREDICTION)

The more neighboring information is used, the higher the likelihood that coding efficiency is improved. Motion vector prediction has been developed to increase the number of context used to predict the current motion vector (e.g., the motion vector information

from reconstructed neighboring blocks). Except for MPEG-2, all codecs use more sophisticated motion vector prediction methods.

SUBPEL ACCURACY (IN INTERPREDICTION)

Although HEVC and VP8 support 1/8 pixel accuracy for subpel interpolation, 1/4 accuracy is normally used for the other codecs. The choice of subpel accuracy seems to be more dependent on picture resolution than IPR issues.

AS THE DEVELOPMENT OF THE CODECS HAS PROGRESSED, THE CRITICAL ISSUE FOR ROYALTY-FREE STANDARD DEVELOPMENT IS NOT SO MUCH THE TECHNICAL MERIT BUT THE IPR.

FRAME TYPE (IN INTERPREDICTION)

The types of frames have not been changed since MPEG-2. Paradoxically, this implies that quite a few free technologies (or nonroyalty bearing patents) are still in use in the latest Type-2 standards.

PARTITION (IN INTERPREDICTION)

The development trends of a partition for interprediction is similar to that of the block partition size. It has been developed to support various block sizes and shapes.

TRANSFORM

All the codecs are based on the DCT transform; only HEVC supports the additional DST for the 4×4 block size. The development trend of the partition in transform size is to support integer-based implementations and the various block sizes.

ENTROPY CODING

Until MPEG-2, arithmetic coding was not often chosen even in Type-2 standards due to IPR issues as well as its implementation complexity. Both issues are no longer valid as they have been resolved over time.

LOOP FILTER

The technology related to loop filters was developed to reduce blocking artifacts. Originally, it was not used as a normative part of coding standard but the AVC/H.264 constrained profile adopted a loop filter method for improving coding efficiency as well as visual quality by including a deblocking filter into the coding loop. The use of loop filters is becoming more frequent as HEVC adopts additional filters (i.e., sample adaptive offset). Currently, ITM 6.0 does not support any loop filter.

SUMMARY

In this article, we reviewed the recent development in royalty-free coding standardization by presenting the history of royalty-free standards in MPEG and analyzing the relationship between patents and standards. As the development of the codecs has progressed, the critical issue for royalty-free standard development is not so much the technical merit but the IPR. Therefore, IPR clearance is mandatory for the project, which marks an unprecedented work process within MPEG by reviewing both technical and IPR aspects. MPEG has been very successful in defining Type-2 standards, with a huge impact in the digital media market. Now, it is time to see if MPEG can continue its success story in Type-1 standard development.

RESOURCES**MPEG**

The MPEG document management system Web site, <http://phenix.int-evry.fr/mpeg/>, has links to MPEG contributions. ISO/IEC WG11 members can access working documents under investigation by MPEG.

VP8

The WebM project home page, <http://www.webmproject.org/>, provides information about VP8, including open-source software and specification.

ACKNOWLEDGMENT

This work was supported by the Industrial Strategic Technology Development

(continued on page 155)

Yuhong Liu and Yan (Lindsay) Sun

Securing Digital Reputation in Online Social Media

As computing and communication systems evolve rapidly and ubiquitously, it has become convenient and almost effortless for individual users to generate, share, and exchange information on online social media. Through online social media, a wide range of digital content, which covers blogging, forums, reviews, social networking, question-answer databases, digital video, mobile phone photography, and wikis, is created by users and has dramatically changed the way people work and interact. However, the simplicity of creating such digital content online has also led to an increase of users' concern about the trustworthiness of such information. To address the issue of trustworthiness, a widely recognized approach is to evaluate the quality of the online information based on feedback from large scale, virtual word-of-mouth networks where individuals share their own opinions and experiences. The aggregated result of such feedback is called *digital reputation*. Digital reputation has already been widely adopted by current online social media. For example, viewers on YouTube may "like" or "dislike" a video clip, buyers on Amazon share their purchasing experiences, travelers evaluate hotels or restaurants on Yelp, and readers can either "dig" or "bury" a piece of social news on Reddit. The reputation-based solution is playing an increasingly important role in influencing users' online social interactions. For example, eBay sellers with established reputations can expect about 8% more revenue than new sellers marketing the same goods [1]; the

survey in [2] reveals that the services receiving five-star ratings will attract 20% more revenue than the same services receiving four-star ratings.

Driven by the increasing profits in online social media, manipulations against digital reputation systems are gaining popularity, which in turn motivates security researchers around the world to prevent reputation manipulations. Most current studies focus on modeling various attacks and developing defense schemes and have already achieved some exciting results. Besides these well-studied attack and defense models, there are two under-investigated aspects: 1) how to obtain reliable data for investigating attack and defense in online reputation systems and 2) how to understand the impact of real-world reputation on digital reputation security. We briefly review the current studies of reputation attack and defense models and then discuss in details these two underinvestigated aspects.

ATTACKS AND DEFENSES IN DIGITAL REPUTATION SYSTEMS

REPUTATION ATTACKS

Due to the increasing impact of reputation systems on individual users' online social interactions, the incentive to manipulate digital reputation is growing. For example, some eBay users are artificially boosting their reputation by buying and selling feedback. On IMDB, a movie named *Resident Evil: Afterlife* had kept an overly inflated reputation score of 8.5 out of 10 with more than 1,800 ratings during its first month of release, whereas its reputation fell down to only 5.9 after the promotion period. For just US\$9.99, a company named "Increase

YouTube Views" can provide 30 "I like" ratings or 30 real user comments to video clips on YouTube. Weibo, the microblog in China where users can follow others as "fans," is one of the most popular social Web sites with billions of users. Some companies are making profits by selling millions of "zombie fans," which are automatically generated fake accounts, to boost customers' popularity. Recently, many online businesses that provide diverse "reputation repairing" services are emerging and gaining popularity. So-called professionally trained writers are provided to write positive reviews/articles and spread them all over the social Web sites, blogs, and forums to repair or boost the reputation of individual customers. The business customers can pay to eliminate/suppress negative reviews, such as bad ratings on the Better Business Bureau Web site (www.BBB.com) and ComplaintsBoard.com, and receive positive reviews on different reputation Web sites, such as Yelp, Google Places, CitySearch, Amazon, and TripAdvisor.

Many scientific studies have been conducted to investigate possible attacks against reputation systems. In [3], many existing attack approaches have been summarized according to their evolution trend, from simple to complicated attacks. For example, the simplest attack is the whitewashing attack where an attacker simply discards its disreputable identification (ID) and re-enters the system by registering a new ID with fresh reputation. A slightly more complex attack is the traitor attack, where an attacker restores reputation by performing good behaviors and then behaves badly again. In advanced attacks, such as sybil attacks,

Digital Object Identifier 10.1109/MSP.2013.2282414

Date of publication: 5 December 2013

the attacker registers multiple user IDs to collaboratively provide unfair feedback. One of the most recent attack models, RepTrap [4], can overturn the reputation of a large number of online items (from positive to negative) and undermine the fairness of the entire reputation system. These studies have shown that reputation attacks may greatly distort reputation scores, undermine users' confidence in the fairness of the reputation systems, and lead to unfair business practices.

REPUTATION DEFENSES

Extensive reputation defense studies have been done from online feedback (i.e., online ratings/reviews) anomaly detection and user-behavior modeling. Studies from the former aspect consider user ratings as random variables and assume dishonest ratings to have statistical distributions different from normal ratings. The approach in [5] assumes that the normal ratings follow a Beta distribution and identifies the ratings outside the majority's opinions as dishonest ratings. In [6], dishonest ratings are eliminated through controlled anonymity and cluster filtering. The defense approaches focusing on user behavior modeling include the iteration refinement approach proposed in [7], which computes the "judging power" for each user as the inverse of this user's rating variance. Users with larger judging power have higher weights in reputation calculation. A personalized trust model is proposed in [8] to enable customized trust evaluations for different users. An in-depth survey on defense approaches can be found in [3].

SUMMARY

Reputation attack and defense studies are developing rapidly. The evolution of one side will inspire the development of the other, and there is always an "arms race" between the reputation attack and defense schemes. Although the reputation attack and defense studies have attracted much research attention, there are still two challenging issues not fully investigated: 1) to obtain reliable and real attack data for studying reputation attacks and defenses and 2) to under-

stand how the digital reputation interacts with the real-world reputation.

DATA COLLECTION

The collection of real user attack data is important for both the study of reputation attack strategies and the evaluation of reputation defense schemes. However, it is costly and inefficient to collect attack data by arbitrarily crawling online social media and manually identifying attacks. It is also extremely difficult to obtain the ground truth of such data (i.e., whether a piece of feedback is honest or dishonest). Therefore, many studies rely on simulated data [5], [7], [8]. However, the simulated data often only represents a few types of attacks, which may have already been considered in the design stage of the defense schemes. Such an evaluation may not reflect the defense performance in practical settings, where attackers may develop diverse and even unknown attacks. The lack of realistic attack data is surely a hurdle in reputation security research.

To address this issue, one promising approach is to collect data through crowdsourcing, where we can launch open calls to an unknown group of solvers (i.e., a crowd). Companies and institutions can use crowdsourcing to help their decision making, problem solving, and data collection. There are four advantages to collect attack data through crowdsourcing: 1) the cost is relatively low; 2) it is much easier to discover the ground truth, if we provide normal data (i.e., honest feedback) and ask the crowd to provide attack strategies; 3) the collected attack strategies are generated by real human users and are therefore more realistic, and 4) more diverse attack strategies can be obtained due to the different knowledge background of the crowd. These advantages make crowdsourcing a promising approach to collect data for reputation research.

A COMPETITION CROWDSOURCING: CANT COMPETITION

The Challenge-of-Attack-on-Network-Trust (CANT) was launched in 2008 to collect reputation attack data. In the competition, we built a virtual reputation

system with normal rating data. The crowd (i.e., players) was required to provide attack strategies to downgrade the reputation score of a given product as much as possible, and the winners received cash rewards. The competition lasted for 18 days and attracted more than 630 registered players with 826,980 valid submissions. The collected data set has provided rich information for investigations of the real user attack behaviors and served as testing data set to evaluate the attack-resistance properties of reputation defense schemes. Figure 1 is the user interface of the CANT competition.

In the competition, each player registered one and only one player ID, which was used to track the player's submissions, score, and rank. Each player ID submitted attack strategies as many times as he or she could. In each specific submission, a player P could use u malicious user IDs to insert r unfair ratings, where $0 < u < U, 0 < r < R$. Here, U and R were the largest number of malicious user IDs and unfair ratings, respectively.

All submissions were divided into groups according to their u and r values. Specifically, the group $G_{u,r}$ contained all submissions that used u malicious user IDs and r unfair ratings. Within a group, the submission that yielded the strongest attack (i.e., downgrading the reputation score of product O_1 the most) was marked as the group winning submission. Note that there might be a tie, leading to multiple winning submissions in one group. Let $s_{u,r}$ denote the number of winning submissions in $G_{u,r}$.

In each group, the winning submissions equally split one point. If there was only one winning submission in $G_{u,r}$ (i.e., $s_{u,r} = 1$), the player who submitted the winning submission gained one point. Then, the overall score of a player was the sum of his/her winning submission points.

A CHALLENGING ISSUE: CHEATING BEHAVIORS

Surprisingly, cheating behaviors were found during the CANT competition. In



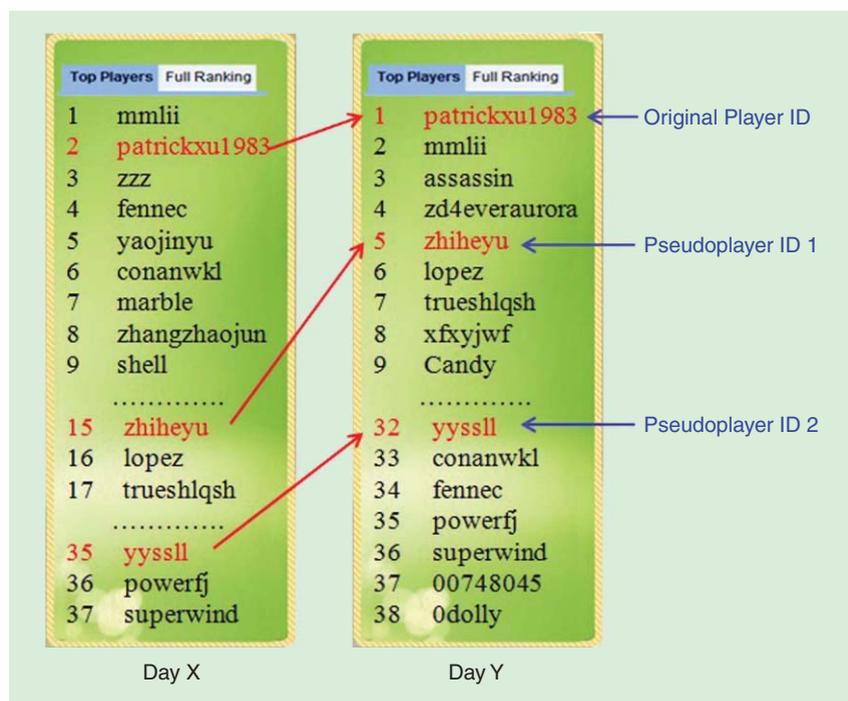
[FIG1] The user interface of the CANT competition.

particular, one player (denoted by cheater C) registered three player IDs. These player IDs had shared the same set of winning submissions and took second, fifth, and 32nd place, respectively, as shown in Figure 2. In the CANT competition, the top 19 players won cash prizes. By using pseudoplayer IDs, the cheater C could increase his rewards, if not detected.

How does such cheating behavior work? Assume that two players, P_1 and P_2 , both have winning submissions in group $G_{u,r}$. Assume P_1 has one winning submission and P_2 has two. Without cheating, P_1 gets $1/3$ points and P_2 gets $2/3$ points, respectively in group $G_{u,r}$. Then, P_1 decides to cheat and registers a new player ID P_d . Through the new player ID P_d , the player P_1 can submit the same winning submission again. By doing so, $s_{u,r}$ is increased from three to four. Then, P_1 gets $1/4$ point and P_2 gets $1/2$ point. Before cheating, the difference between P_1 and P_2 is $1/3$ point. After cheating, the difference between P_1 and P_2 is reduced to $1/4$ point. Through the new player ID P_d , 1) the score gap

between P_1 and P_2 is narrowed down, and the advantage of P_2 over P_1 is reduced; and 2) the player ID P_d grabs

more points and gains a higher rank. This is exactly what happened in the CANT competition.



[FIG2] Rank boosting through multiple pseudo-IDs in the CANT competition.

[TABLE 1] PSEUDO-ID DETECTION RESULTS OF DIFFERENT SCHEMES.

	GROUND TRUTH	THE PROPOSED SCHEME	SCORE-BASED SCHEME
PSEUDO-ID	5, 32	5, 20, 32	5
ORIGINAL ID	2	2, 1	NONE

This type of cheating behavior can exist in other competition crowdsourcing cases, where competition rewards are provided. It is important to detect such cheating behavior since it will not only contribute no meaningful data/solutions to the system, but also ruin the fairness of the competition.

DETECTION OF CHEATING BEHAVIOR

We detect cheating behaviors in the CANT competition by building an implicit social network among players. Most existing social networks inherently describe collaboration among users (e.g., Facebook users who are connected are friends). Can the social network concept be used in a competition environment, in which nodes (i.e., players) have to defeat others to achieve their goals? We define a competition social network to describe such scenarios, in which the nodes' behaviors are dramatically different from these in collaborative social networks.

In the context of the CANT competition, we introduce the following concepts.

- A competition relationship exists and only exists between two player IDs when they have winning submissions in the same group (e.g., $G_{u,r}$).
- The competition value is computed for each pair of players with competition relationship. Assume $t_{u,r}^i$ denote the points obtained by player P_i in group $G_{u,r}$. We define

$$H_{u,r}^{i,j} = \begin{cases} 0 & \text{if } t_{u,r}^i \cdot t_{u,r}^j = 0 \\ 1 & \text{if } t_{u,r}^i \cdot t_{u,r}^j \neq 0 \end{cases}$$

It is easy to see that $H_{u,r}^{i,j} = 1$ if and only if both P_i and P_j win points (or compete) in group $G_{u,r}$. The competition value from P_i to P_j and from P_j to P_i are

$$V_{pi \rightarrow pj} = \sum_{u=1}^U \sum_{r=1}^R t_{u,r}^i \cdot H_{u,r}^{i,j}$$

and

$$V_{pj \rightarrow pi} = \sum_{u=1}^U \sum_{r=1}^R t_{u,r}^j \cdot H_{u,r}^{i,j},$$

respectively.

- In the competition network, each player ID is a node. If two player IDs have competition relationship (i.e., $t_{u,r}^i \cdot t_{u,r}^j \neq 0$), there is a bidirectional link between them and two competition values are computed.

- The competition degree of a node is the number of links connected to this node in the competition network.

Although we focus on the CANT competition here, the concept of a competition network can be extended to other scenarios as long as one can define a quantitative competition value between two players. Ideally, the competition network can be updated whenever a new submission is received. To simplify the computation, we divide the overall time of the competition into 36 equal time frames, where one frame roughly represents a half day. We only update the competition network at the end of each time frame. We refer to the main ID controlled by the cheater as the *original ID* and to the other IDs controlled by the cheater as the *pseudo-IDs*. The goal is to detect the pseudo-IDs and their association with the original ID.

Unlike normal players whose winning submissions are accumulated gradually, pseudo-IDs usually share winning submissions from the original ID within a short time, leading to either a low competition degree (i.e., sharing only unpopular winning submissions) or a sudden increase in the competition degree (i.e., sharing popular winning submissions). The player IDs that fulfill either of these two conditions will be identified as pseudo-IDs. Furthermore, if the competition value between a player ID and an identified pseudoplayer ID is much larger than the average competition value, this

player ID is marked as the associated original ID.

With the competition social network, we detect that Player 5 and Player 32 are the pseudo-IDs of Player 2, and Player 20 is the pseudo-ID of Player 1. We compare the proposed detection scheme with a simple score-based scheme, where a player is considered as pseudoplayer if his/her score suddenly increases. With a simple score-based scheme, we can detect Player ID 5 as a pseudoplayer ID, while the original ID cannot be detected. The results are shown in Table 1. The ground truth is obtained by interviewing players after the competition.

THE ROLE OF REAL-WORLD REPUTATION IN REPUTATION SECURITY RESEARCH

Although digital reputation is an important factor in influencing users' decision making, it is not the only one. Beyond digital reputation, users also make decisions based on the real-world reputation from the words of their friends, neighbors, and coworkers. Although the real-world reputation does exist, it has seldom been considered in the reputation security research, since the digital reputation is believed to be a dominant factor. The digital reputation may dominate in global markets where few users know each other. However, how about in a closely connected social community? Will the real-world reputation play an important role? If so, how will it influence the reputation security research?

COMPARING THE IMPACT OF DIGITAL REPUTATION AND REAL-WORLD REPUTATION

We discuss a study on mobile application (i.e., app) installation, in which both digital reputation and real-world reputation affect users' decision on whether to install an app. Similar to other markets, in the app market, people believe that the digital reputation heavily influences users' shopping decisions. Since app rating and download number are the two most important factors in the calculation of digital reputation, most manipulation is launched against these two factors. A well-known attack is the pay-per-install

model, where app sellers pay for each installation to boost the download number. Some companies, such as App Lifter, provide services for app sellers to directly pay users for installing their apps. Some other companies, such as Tapjoy and Flurry, manage pay-per-install networks composed of plenty of apps. Apps in such networks encourage their users through virtual currency or level upgrading to download other apps in the same network. App ratings/reviews can also be manipulated; for example, Molinker, the app developer with more than 1,000 apps, has been revoked from the app market due to a review scam [9].

On the other hand, with the popularity of tablet computers and smartphones, many users have experiences of installing mobile apps, and they often share such experiences with their local connections (e.g., friends, family members, colleagues). Thus, within a local community (i.e., university campus), apps may also have their real-world reputation, which provides us with an opportunity to evaluate the impact of the real-world reputation on users' decisions.

TESTING DATA

The testing data is a real user data set collected by the Massachusetts Institute of Technology Media Lab [10]. This data set records the installations of 821 apps from 55 participants, are residents living in a graduate student residency of a major U.S. university, from March to July 2010. In this data set, the following information has been collected:

- Users' app installation information (i.e., which user installed which app at what time).
- Call log and Bluetooth hits information. During the data collection period, each participant was given an Android-based cell phone with a built-in sensing software to capture all call logs and Bluetooth hits among the given phones. Call logs were used to indicate participants' interactions through phone calls. Bluetooth hits recorded participants' face-to-face interactions, during which the phones were within each other's vicinity. These two types of

information described participants' daily interactions.

■ Users' friendship, affiliation, and race information was collected through a survey. In the survey, each participant provided his/her affiliation and race and rated his/her friendship relationship to other participants. Such information reflected more about participants' long-term relationship.

This data set perfectly matches our requirements due to two reasons. First, it contains rich information about users' real-world interactions, i.e., call log, Bluetooth hits, friendship, affiliation, and race, which represents the real-world reputation. Second, users' app installation information, which is rarely available in other data sets, makes it possible to analyze the installation decision for each specific user. Beyond this information, we further collect the app rating and download number information to represent the digital reputation.

IMPACT OF DIFFERENT INFORMATION FACTORS

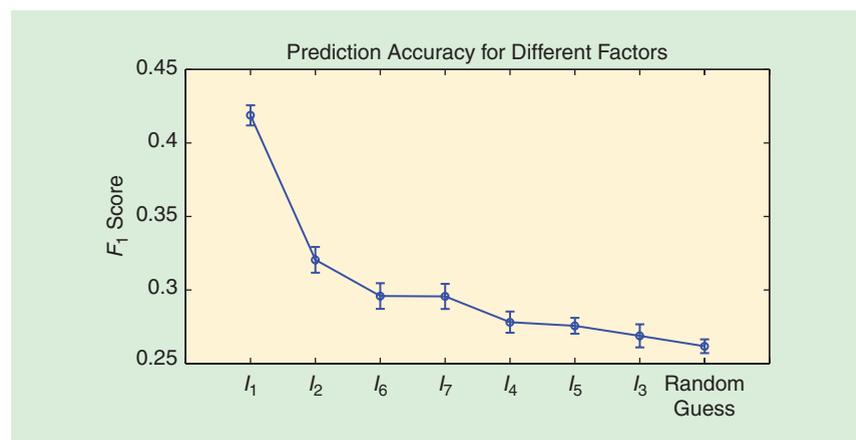
To evaluate the impact of different information on users' app installations, the first step is to accurately estimate app installations. We adopt the model in [10], which predicts app installations by constructing a composite network containing multiple sources of information. To the best of our knowledge, this is currently the most suitable model in

terms of predicting app installations from diverse information sources.

In [10], the goal is to derive the optimized model to combine all the pieces of information so that users' app installations can be predicted with high accuracy. We use this model to understand which information has larger impact on users' app installations. The assumption is that if one factor has larger influence on users' app installations, the optimized prediction based on this factor should yield a higher accuracy. Therefore, each time, we take only one factor as the input to optimize the prediction model and calculate the prediction accuracy. The impact evaluation of different factors is based on the comparison of the prediction accuracy.

Specifically, to calculate the prediction accuracy, we adopt the F_1 -score as the performance measurement, which is computed as $(2 \times \text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$. Here the precision is the number of correct results divided by the number of all returned results, and the recall is the number of correct results divided by the number of results that should have been returned. The optimal F_1 score is obtained by computing F_1 scores for each point on the precision-recall curve and selecting the largest F_1 value.

Figure 3 demonstrates the optimized prediction accuracy for each different factor. The y-axis represents the prediction accuracy (i.e., optimal F_1 score),



[FIG3] The impact of different information factors on the prediction accuracy (I_1 : call log; I_2 : Bluetooth hits; I_3 : friendship; I_4 : affiliation; I_5 : race; I_6 : app download number; and I_7 : app rating).

and the x-axis represents different input factors. From Figure 3, we can make several observations:

- Compared to random guess, where no information is available, all these factors achieve higher F_1 scores thus suggesting that each of these factors will provide some information for the app installations. In other words, all of these factors have impact on users' app installations.
- Prediction with call log information yields the highest F_1 score, indicating that call log has more impact on users' app installation decisions than other factors. If a user has many frequently called friends who have installed a certain app, the call-log number for this user and this particular app is high. The detailed computation of call-log and other factors can be found in [10].
- The daily interaction information, i.e., call log and Bluetooth hits, has a much higher impact compared to the long-term relation information, such as affiliation, race, and friendship. This suggests that users' app installations may easily be influenced by people who contact them every day. Even if two users are friends, if they do not contact each other frequently, their impact on each others' app installations is limited.
- The impact of app rating and download number, which represents the digital reputation, is lying in the middle thus suggesting that to install apps, users in a closely connected social community, such as a university campus community, may first refer to people whom they contact frequently, then check out rating/download information that is publicly available, and at last refer to other people who are in the community but have less frequent contact.

Based on these observations, it seems that in a closely connected social community, the real-world reputation has larger impact than the digital reputation. Furthermore, among different social connections, users tend to be influenced more by their frequent contacts. Note that this study is based on the data from a very special community, a university campus and

may not generalize to a broader consumer base.

THE ROLE OF THE REAL-WORLD REPUTATION ON REPUTATION SECURITY RESEARCH

The experimental results shown above suggest that when a person knows his/her friends' opinion about an app, this person will pay much less attention or even ignore the online reviews and ratings. Although the real-world reputation has rarely been considered in the reputation security research, it is the dominant factor in influencing users' decisions in a closely connected social community. Then how will this influence the reputation security research? We would like to discuss it from both the attack and the defense perspectives.

A GOOD UTILIZATION OF THE REAL-WORLD REPUTATION CAN WELL COMPLEMENT THE DIGITAL REPUTATION AND HELP IMPROVING ITS ATTACK-RESISTANCE PROPERTIES.

From the attack perspective, to influence users' decisions in a closely connected social community, manipulating the real-world reputation may be more effective than manipulating the global digital reputation. Therefore, the pay-per-install model may not be a good attack strategy for a closely connected social community. On the other hand, the app developers are suggested to advertise their apps on online forums, blogs, and social networks (e.g., Facebook, Twitter) to boost their real-world reputation and cultivate potential customers. Although such promotions are usually more costly and take longer, they may be more effective than manipulating the digital reputation only.

From the defense perspective, the designer takes advantage that real reputation overwrites digital reputation in closely connected social communities. For example, if the designer can identify the users who belong to the same closely

connected social community (e.g., university campus) and capture the reputation of an app in this community (e.g., ratings from this community), the designer can then investigate whether this local reputation agrees with the global digital reputation. Any significant difference may lead to further investigation. Another interesting direction is to make the real-world reputation more accessible. A user may benefit from a more personalized reputation system that considers this user's social community. Given these, we believe that a good utilization of the real-world reputation can well complement the digital reputation and help improving its attack-resistance properties. This will be an interesting direction for future research.

SUMMARY

In this column, we have discussed security issues of digital reputation in online social media. In particular, digital reputation has already been proven to be an effective approach to ensure information quality in the rapidly developing online social media. Driven by the low costs and large potential profits of manipulating digital reputation, diverse attacks are emerging, which attracts much research attention. However, due to the lack of realistic attack data, the evaluation of the reputation defense schemes has been a challenging task. A promising approach is to collect the real user attack data through crowdsourcing. Unexpectedly, cheating behaviors may also exist in the crowdsourcing process. To address this issue, we described a "competition social network" to effectively model the crowd's behavioral patterns and to detect anomaly. Finally, we have compared the digital reputation and the real-world reputation, and believe that the integration of the two types of reputation will be an interesting direction for research aiming to provide secure and trustworthy reputation in online social media.

RESOURCES

- Everything you need to know about eBay feedback. [Online]. Available: <http://www.newlifeauctions.com/feedback.html>

- R. Kalla. IMDB, Whats going on with your scores? [Online]. Available: <http://www.thebuzzmedia.com/imdb-whats-going-on-with-your-scores/>
- Increase views and ratings on YouTube videos. [Online]. Available: <http://www.earnersonline.com/increase-youtube-video-views/>
- S. Millward. Of Sina Weibo's 500 million registered users, Are 90% actually zombies? [Online]. Available: <http://www.techinasia.com/sina-weibo-90-percent-users-zombies/>
- Reputation repair. [Online]. Available: <http://www.reputationchanger.com/>
- App Lifter. [Online]. Available: <http://applifter.com/>
- Tapjoy. [Online]. Available: <http://developers.tapjoy.com/how-it-works/>
- Flurry. [Online]. Available: <http://www.flurry.com/appCircle-a.html>.

ACKNOWLEDGMENT

This research was sponsored by the National Science Foundation under award 0643532.

AUTHORS

Yuhong Liu (yuhong@ele.uri.edu) is an assistant professor at Penn State, Altoona.

Yan (Lindsay) Sun (yansun@ele.uri.edu) is an associate professor at the University of Rhode Island.

REFERENCES

- [1] P. Resnick and R. Zeckhauser, "Trust among strangers in internet transactions: Empirical analysis of Ebay's reputation system," in *The Economics of the Internet and E-Commerce* (Advances in Applied Microeconomics, vol. 11). Amsterdam, The Netherlands: Elsevier, 2002, pp. 127–157.
- [2] comScore Inc. and T. K. Group. (2007, Nov.). Press release: Online consumer-generated reviews have significant impact on offline purchase behavior. [Online]. Available: <http://www.comscore.com/press/release.asp?press=1928>
- [3] Y. L. Sun and Y. Liu, "Security of online reputation systems: Evolution of attacks and defenses," *IEEE Signal Processing Mag.*, vol. 29, no. 2, pp. 87–97, Mar. 2012.

- [4] Y. Yang, Q. Feng, Y. Sun, and Y. Dai, "Reputation trap: An powerful attack on reputation system of file sharing p2p environment," in *Proc. 4th Int. Conf. Security and Privacy in Communication Networks (SecureComm'08)*, Istanbul, Turkey, Sept. 2008.
- [5] A. Josang and R. Ismail, "The beta reputation system," in *Proc. 15th Bled Electronic Commerce Conf.*, Bled, Slovenia, June 2002, pp. 324–337.
- [6] C. Dellarocas, "Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior," in *Proc. 2nd ACM Conf. Electronic Commerce*, Minneapolis, MN, Oct. 2000.
- [7] P. Laureti, L. Moret, Y.-C. Zhang, and Y.-K. Yu, "Information filtering via iterative refinement," in *Europhys. Lett.*, vol. 75, pp. 1006–1012, 2006.
- [8] J. Zhang and R. Cohen, "A personalized approach to address unfair ratings in multiagent reputation systems," in *Proc. 5th Int. Joint Conf. Autonomous Agents and Multiagent Systems (AAMAS) Workshop on Trust in Agent Societies*, Hakodate, Japan, 2006, pp. 89–98.
- [9] C. Sorrel. Apple expels 1,000 apps from store after developer scam. [Online]. Available: <http://www.wired.com/gadgetlab/2009/12/apple-expels-1000-apps-from-store-after-developer-scam/>
- [10] W. Pan, N. Aharony, and A. Pentland, "Composite social network for predicting mobile apps installation," in *Proc. 25th Conf. Artificial Intelligence (AAAI-II)*, San Francisco, CA, Aug. 2011.

[SP]

standards **IN A NUTSHELL** (continued from page 148)

Program (10047438, Development and International Standardization for MPEG Type-1 Standard Technology) funded by the Ministry of Trade, Industry, and Energy of Korea.

AUTHORS

Kiho Choi (aikiho@gmail.com) is a post-doctoral researcher at Hanyang University in Seoul, South Korea. He is the chair of the IVC Ad Hoc Group in MPEG.

Euee S. Jang (esjang@hanyang.ac.kr) is a professor at Hanyang University. He has been participating in MPEG in various leading roles since 1996.

REFERENCES

- [1] ISO standards and patents [Online]. Available: http://www.iso.org/iso/standards_development/patents

- [2] M. Hicks. (2007). JPEG hits new patent-infringement snag. [Online]. Available: <http://www.eweek.com/c/a/Enterprise-Applications/JPEG-Hits-New-PatentInfringement-Snag/>

- [3] MPEG-LA official Web site. (2004). [Online]. Available: <http://www.mpegla.com>

- [4] AT&T patent licensing Web site. [Online]. Available: <http://www.att.com/gen/sites/ipsales?pid=19116>

- [5] I. Richardson, "Vcodex White Paper: Video compression patents," Vcodex Ltd., 2008, pp. 3–6.

- [6] F. Mueller. (2012). Motorola calls \$4 billion royalty figure "misleading" but doesn't dispute the number per se. [Online]. Available: <http://www.fosspatents.com/2012/04/motorola-calls-4-billion-royalty-figure.html>

- [7] J. Bankoski, P. Wilkins, and Y. Xu, "Technical overview of VP8, An open source video codec for the Web," in *Proc. 2011 IEEE Int. Conf. Multimedia and Expo (ICME)*, pp. 1–6.

- [8] H. Alvestrand and A. Grange, "VP8 as RTCWEB mandatory to implement." IETF, Internet-Draft/Informational draft-alvestrand-rtcweb-vp8-00, webtrc, Oct. 15, 2012.

- [9] H. Alvestrand, A. Grange, J. Luther, M. Raad, and L. Bivolarski, "Google Inc.'s response to the CfP

on Internet video technologies," ISO/IEC JTC1/SC 29/WG11/M29693, Vienna, Austria, July 2013.

- [10] E. Protalinski. (2013). Google adds its free and open-source VP9 video codec to latest Chrome build. [Online]. Available: <http://thenextweb.com/google/2013/07/01/google-adds-its-free-and-open-source-vp9-video-codec-to-latest-chrome-build/>

- [11] "Call for proposals (CfP) for Internet video coding technologies," ISO/IEC JTC1/SC29/WG11 N12204, 2011.

- [12] R. Wang, Q. Yu, H. Lv, L. Chen, X. Zhang, S. Ma, T. Huang, and W. Gao, "IVC CE2: Internet Video Coding Test Model (ITM) performance improvements," ISO/IEC JTC1/SC 29/WG11/M29130, Incheon, South Korea, Apr. 2013.

- [13] "Call for proposals (CfP) for Internet video coding technologies," ISO/IEC JTC1/SC29/WG11 N13546, 2013.

- [14] "Resolutions of 105th meeting," ISO/IEC JTC1/SC29/WG11 N13648, 2013.

- [15] "Resolutions of 104th meeting," ISO/IEC JTC1/SC29/WG11 N13455, 2013.

[SP]

[dates **AHEAD**]

Please send calendar submissions to:
Dates Ahead, c/o Jessica Barragué
IEEE Signal Processing Magazine
445 Hoes Lane
Piscataway, NJ 08855 USA
e-mail: j.barrague@ieee.org
(Colored reference title indicates
SP-sponsored conference.)

2013**[DECEMBER]**

2013 Fifth International Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP)
15–18 December, Saint Martin, France.
General Cochairs: Aleksandar Dogandži and Martin Haard
URL: <http://www.stevens.edu/camsap2013/>

2014**[MARCH]**

IEEE World Forum on Internet of Things (IOT)
6–8 March, Seoul, South Korea.
General Chair: Roberto Minerva
General Cochairs: Jaiyong Lee, Edward Cole, Stephen Dukes, Paul Hartmann, Abbas Jamalipour, Jean Philippe Vasseur, John Vig, Yen-Kuang Chen, and Mengchu Zhou
URL: <http://sites.ieee.org/wf-iot/>

Digital Object Identifier 10.1109/MSP.2013.2285948
Date of publication: 5 December 2013

Data Compression Conference (DCC)
26–28 March, Salt Lake City, Utah.
URL: <http://cs.brandeis.edu/~dcc/index.html>

[APRIL]

IEEE International Symposium on Biomedical Imaging (ISBI)
28 April–2 May, Beijing, China.
General Chairs: Ge Wang and Bin He
URL: <http://biomedicalimaging.org/2014/>

[MAY]

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
4–9 May, Florence, Italy.
General Cochairs: Fulvio Gini and Marco Luise
URL: <http://www.icassp2014.org/>

IEEE Radar Conference (RadarCon)
19–23 May, Cincinnati, Ohio.
General Chair: Brian Rigling
URL: <http://www.radarcon2014.org/>

6th International Symposium on Communications, Control, and Signal Processing (ISCCSP)
21–23 May, Athens, Greece.
Honorary Chair: Anastasios Venetsanopoulos
General Cochairs: Sanjit K. Mitra and Thanos Stouraitis
URL: <http://isccsp2014.upatras.gr/>

[JUNE]

IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)
22–25 June, A Caruña, Spain.
URL: <http://www.gtec.udc.es/sam2014/>

IEEE Statistical Signal Processing Workshop (SSP)
29 June–2 July, Gold Coast, Australia.

[JULY]

2nd IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)
9–13 July, Xi'an, China.
General Chairs: Mingyi He and Kung Yao
URL: <http://www.chinasip2014.org/CfP.htm>

IEEE International Conference on Multimedia and Expo (ICME)
14–18 July, Chengdu, China.
General Cochairs: Touradj Ebrahimi, Shipeng Li, Houjun Wang, and Jie Yang
URL: <http://www.icme2014.org/>

[DECEMBER]

6–9 December, South Lake Tahoe, California.
IEEE Spoken Language Technology Workshop (SLT)
General Chairs: Murat Akbacak and John Hansen

2015**[APRIL]**

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
19–24 April, Brisbane, Australia.
General Cochairs: Vaughan Clarkson and Jonathan Manton
URL: <http://icassp2015.org/>

Mos Kaveh and Li Deng

The Discipline of Signal Processing: Part 2

EDITOR'S INTRODUCTION

In this issue of *IEEE Signal Processing Magazine*, we continue the exercise of compiling in a few words the reflections from signal processing experts on the following three questions:

- How would you explain what signal processing is to somebody with no significant background and knowledge in technology disciplines?
- What does signal processing mean to you? (to be answered in technical terms, as preferred/needed)
- Where do you see the field of signal processing ten years from now?

Our goal was to use these questions as conduits to elicit reflections about what is signal processing and to where it is

headed. In this issue of the magazine, we present the thoughts from Mos Kaveh and Li Deng. As I mentioned in the November 2013 issue, we want to make this article an exercise of perspectives and reflection, so please feel free to share with me your own thoughts on the matter. I hope that you, our readers, will find this column interesting and you will enjoy reading this perspective from the experts in the signal processing field.

Andres Kwasinski (axkeec@rit.edu) is an associate professor at the Rochester Institute of Technology and area editor of columns and forum of *IEEE Signal Processing Magazine*.

MOS KAVEH

(mos@umn.edu)

- Professor of electrical and computer engineering, University of Minnesota.
- Associate dean for research and planning for the College of Science and Engineering, University of Minnesota.
- President, IEEE Signal Processing Society (2010–2011).
- Vice president, awards and membership, IEEE Signal Processing Society (2003–2007).
- Vice president, publications, IEEE Signal Processing Society (1993–1996).
- General chair, 1993 International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

In Shakespeare's *Romeo and Juliet*, Juliet famously muses "What's in a name? That which we call a rose, by any other name would smell as sweet." Good question, Juliet. Many have been wondering whether signal processing is the "right" name for our field and profession in the 21st century. Mind you, the name is as meaningful and sweet for those of us who

have and continue to work in signal processing or in areas closely allied with it. Such questioning has more to do with the branding of what we do so others—students, general public, media, and government and business decision makers—appreciate the breadth, impact, and

KEY IN THE DIVERSIFICATION OF SIGNAL PROCESSING IS THE NOTION AND CHARACTER OF A SIGNAL ITSELF, WHICH NOW IS ANY REPRESENTATION OF INFORMATION.

continued critical relevance of signal processing in the information- and data-driven world we live in.

Over the past four decades, the field has evolved and diversified tremendously in its fundamentals and applications. Signal processing is everywhere; it is science and application agnostic, dynamic, and synergistic. Its tools and the resulting technologies ubiquitous and enablers

of how we live, communicate, play, and discover. While initiated and nurtured in electrical engineering, the discipline has been richly influenced by theories and tools of mathematics, statistics, speech and language, computer and computational sciences, physics, and geophysics.

Key in the diversification of signal processing is the notion and character of a signal itself, which now is any representation of information. The first sentence in the field of interest of the IEEE Signal Processing Society (<http://www.signalprocessing.org/about-sps/scope-mission/>) states "Signal processing is the enabling technology for the generation, transformation, extraction, and interpretation of information." Today, a signal processing professional is likely to be using signals/data and related models to solve problems in audio, speech, image, video, communications, energy systems, social systems and networks, finance, genomics, brain function, astronomy, to name a few.

Alas, signal processing is not a principal go-to discipline for talent and expertise when employment is offered, and collaboration is sought on data analytics,

Digital Object Identifier 10.1109/MSP.2013.2282792

Date of publication: 5 December 2013

REFLECTIONS continued

such as modeling, mining, and classification. The emergence of data science, primarily through the morphing of statistics, elements of computer and information sciences and business analytics, is creating additional challenges, but also opportunities for the signal processing community. The challenge remains in the brand. Not only does “signal processing inside” characterize lack of visibility of the field and its technologies in the broader community, scientific or otherwise, the vibrant subdisciplines of signal processing that are contributing significantly to data acquisition and analytics for small and big data, are doubly hidden.

So say I am at a party, and someone asks me about my field of research. Signal processing, I say. Faced with a blank stare, and a long sip of his drink, I open up. One of my recent students and I worked with a radiologist to develop algorithms to help with the detection of certain abnormalities from magnetic resonance imaging of the brain. Another student and I came up with a way of automatically mitigating the interference from wind turbines on Doppler weather radar signals, so that meteorologists would not get confused in classifying the severity of a storm, and another student...The conversation becomes really interesting after this technical prelude.

Signal processing is thriving. Submissions to the publications and conferences of the IEEE Signal Processing Society have been increasing, apparently with no end in sight. To more effectively brand what we do, we have no choice but to connect with applications that are meaningful to the public. And, to have real impact beyond our own circles, we must actively engage and collaborate with domain experts, for example, in biology, medicine, energy, and business. We must engage such domain experts in our publications and conferences, and, in turn, communicate results in influential venues outside of our own.

LI DENG

(deng@microsoft.com)

- Principal researcher of Microsoft Research, and affiliate professor of the University of Washington (since 2000).

- Assistant, tenured associate, and full professor of the University of Waterloo, Canada (1989–1999).
- Member, Board of Governors of the IEEE Signal Processing Society (2008–2010).
- Editor-in-chief, *IEEE Signal Processing Magazine* (2010–2012).
- Editor-in-chief, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2012–2014).
- General chair, ICASSP 2013.

How would you explain what signal processing is to somebody with no significant background and knowledge in technology disciplines?

Signal processing is in the everyday lives of you and me. Each time you turn on a cell phone, signal processing is working for you. Each time you start your

TO MORE EFFECTIVELY
BRAND WHAT WE DO,
WE HAVE NO CHOICE
BUT TO CONNECT
WITH APPLICATIONS
THAT ARE MEANINGFUL
TO THE PUBLIC.

car engine, you rely on signal processing. The type of “signals” in signal processing is very broad: audio, music, speech, language, text, image, graphics, video, sensor measurements, communication, geophysical, sonar, radar, biological, chemical, molecular, genomic, medical, data, etc. What can you do with all these sorts of “signals?” Without signal processing, they may be of little use. With signal processing, such rich sets of information can be transformed to enable devices or machines to operate according to your needs or to even exhibit intelligent behavior. Examples are the fast transmission of videos for you to enjoy watching YouTube, where video coding is at play, Kinect in Microsoft Xbox, and conversational devices such as iPhone’s Siri. It is fair to characterize signal processing as applied mathematics for modern information technology.

What does signal processing mean to you?

Signal processing is an enabling technology that allows many forms of signal and information to be made practically useful via many forms of transformation or “processing.” The development of signal processing over the past decade or so has expanded the traditional concepts of both “signal” and “processing” in significant ways. A few years ago, I wrote an editorial [1] elaborating on the expansion in both of these axes. Along the “signal” axis, the traditional signals of audio, speech, image, video, communication, geophysical, sonar, radar, etc. has been extended to include language, text, biological, genomic, and social interactions data. What characterizes such new types of signals is the structure embedded in them (e.g., sequential or graph structure) and the symbolic nature endowing the signals with semantic information (e.g., language and genomic data). Along the “processing” axis, the traditional concepts on digital signal processing are also growing to encompass the elements of interpretation and understanding of the signals. So what modern signal processing means to me is the expanded scope along both of the axes above. In my own areas of speech and language processing research, the traditional disciplines of speech analysis, coding, enhancement, synthesis, and recognition are now reaching higher levels of “processing” that comprise understanding of the meaning embedded in the speech signal, not just getting the text out of the signal, and making decisions on what the appropriate response the conversational device should give based on the understanding of the speech signal.

Where do you see the field of signal processing ten years from now?

Within the next ten years, I expect a great deal of and growing interplay between the community of signal processing and those from artificial intelligence, machine learning, computer science, and

applied mathematics (e.g., optimization). As an example, we just announced that our Society's *IEEE Transactions on Audio, Speech, and Language Processing* is merging with the ACM counterpart, with the main goal of expanding language processing with greater strength from computer science researchers. As a result of this merger, the number of editorial board members has now increased to contain more computer scientists including those from major technology companies such as Google, Microsoft, IBM, and Facebook. In particular, machine learning will be more and more deeply ingrained within signal processing technology. In a recently published paper that I wrote together with X. Li, now a Facebook scientist, during a period of close to three years [2], we analyzed the mutual influence between machine learning and speech recognition over the past 30 years,

and reached the conclusion that the major effects of such positive influence are just beginning. An important recent

AN IMPORTANT RECENT EXAMPLE IS THE SPECTACULAR SUCCESS OF DEEP LEARNING IN SPEECH RECOGNITION, AS A RESULT OF TIMELY COLLABORATION BETWEEN SIGNAL PROCESSING AND MACHINE-LEARNING RESEARCHERS.

example is the spectacular success of deep learning in speech recognition, as a result of timely collaboration between signal

processing and machine-learning researchers [3], [4]. In ten years, I believe signal processing will become more ubiquitous, with intelligent behavior emerging from signal processing algorithms/techniques developed likely side-by-side with machine learning and artificial intelligence methods and technology.

REFERENCES

- [1] L. Deng, "Expanding the scope of signal processing," *IEEE Signal Processing Mag.*, vol. 25, no. 3, pp. 2–4, May 2008.
- [2] L. Deng and X. Li, "Machine learning paradigms for speech recognition: An overview," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 21, no. 5, pp. 1060–1089, May 2013.
- [3] G. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Processing Mag.*, vol. 29, no. 6, pp. 82–97, 2012.
- [4] G. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large vocabulary speech recognition," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 20, no. 1, pp. 30–42, 2012.

SP



Innovation doesn't just happen. Read first-person accounts of IEEE members who were there.

IEEE Global History Network
www.ieeeahn.org

Photo: NASA

 **IEEE**

advertisers **INDEX**

The Advertisers' Index contained in this issue is compiled as a service to our readers and advertisers: the publisher is not liable for errors or omissions although every effort is made to ensure its accuracy. Be sure to let our advertisers know you found them through *IEEE Signal Processing Magazine*.

ADVERTISER	PAGE	URL	PHONE
EE Live!	3	eeliveshow.com	
Mathworks	CVR 4	www.mathworks.com/accelerate	+1 508 647 7040

advertising **SALES OFFICES**

James A. Vick
Sr. Director, Advertising
Phone: +1 212 419 7767;
Fax: +1 212 419 7589
jv.ieeemedia@ieee.org

Marion Delaney
Advertising Sales Director
Phone: +1 415 863 4717;
Fax: +1 415 863 4717
md.ieeemedia@ieee.org

Susan E. Schneiderman
Business Development Manager
Phone: +1 732 562 3946;
Fax: +1 732 981 1855
ss.ieeemedia@ieee.org

Product Advertising
MIDATLANTIC

Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, PA, DE, MD, DC, KY, WV

**NEW ENGLAND/SOUTH CENTRAL/
EASTERN CANADA**

Jody Estabrook
Phone: +1 774 283 4528;
Fax: +1 774 283 4527
je.ieeemedia@ieee.org
ME, VT, NH, MA, RI, CT, AR, LA, OK, TX
Canada: Quebec, Nova Scotia,
Newfoundland, Prince Edward Island,
New Brunswick

SOUTHEAST

Thomas Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
tf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

MIDWEST/CENTRAL CANADA

Dave Jones
Phone: +1 708 442 5633;
Fax: +1 708 442 7620
dj.ieeemedia@ieee.org
IL, IA, KS, MN, MO, NE, ND,
SD, WI, OH
Canada: Manitoba,
Saskatchewan, Alberta

**MIDWEST/ ONTARIO,
CANADA**

Will Hamilton
Phone: +1 269 381 2156;
Fax: +1 269 381 2556
wh.ieeemedia@ieee.org
IN, MI, Canada: Ontario

**WEST COAST/MOUNTAIN STATES/
WESTERN CANADA**

Marshall Rubin
Phone: +1 818 888 2407;
Fax: +1 818 888 4907
mr.ieeemedia@ieee.org
AZ, CO, HI, NM, NV, UT, AK, ID, MT,
WY, OR, WA, CA. Canada: British
Columbia

**EUROPE/AFRICA/MIDDLE EAST
ASIA/FAR EAST/PACIFIC RIM**

Louise Smith
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
les.ieeemedia@ieee.org
Europe, Africa, Middle East
Asia, Far East, Pacific Rim, Australia,
New Zealand

Recruitment Advertising**MIDATLANTIC**

Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, CT, PA, DE, MD, DC, KY, WV

NEW ENGLAND/EASTERN CANADA

Liza Reich
Phone: +1 212 419 7578;
Fax: +1 212 419 7589
e.reich@ieee.org
ME, VT, NH, MA, RI, Canada: Quebec,
Nova Scotia, Prince Edward Island,
Newfoundland, New Brunswick

SOUTHEAST

Cathy Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
cf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

**MIDWEST/SOUTH CENTRAL/
CENTRAL CANADA**

Darcy Giovingo
Phone: +224 616 3034;
Fax: +1 847 729 4269
dg.ieeemedia@ieee.org
AR, IL, IN, IA, KS, LA, MI, MN, MO, NE,
ND, SD, OH, OK, TX, WI, Canada:
Ontario, Manitoba, Saskatchewan, Alberta

**WEST COAST/SOUTHWEST/
MOUNTAIN STATES/ASIA**

Tim Matteson
Phone: +1 310 836 4064;
Fax: +1 310 836 4067
tm.ieeemedia@ieee.org
AZ, CO, HI, NV, NM, UT, CA, AK, ID, MT,
WY, OR, WA, Canada: British Columbia

EUROPE/AFRICA/MIDDLE EAST

Louise Smith
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
les.ieeemedia@ieee.org
Europe, Africa, Middle East

Digital Object Identifier 10.1109/MSP.2013.2290495



LEARNING HAS NO
BOUNDARIES

**YOU KNOW YOUR STUDENTS NEED IEEE INFORMATION.
NOW THEY CAN HAVE IT. AND YOU CAN AFFORD IT.**

IEEE RECOGNIZES THE SPECIAL NEEDS OF SMALLER COLLEGES, and wants students to have access to the information that will put them on the path to career success. Now, smaller colleges can subscribe to the same IEEE collections that large universities receive, but at a lower price, based on your full-time enrollment and degree programs.

Find out more—visit www.ieee.org/learning





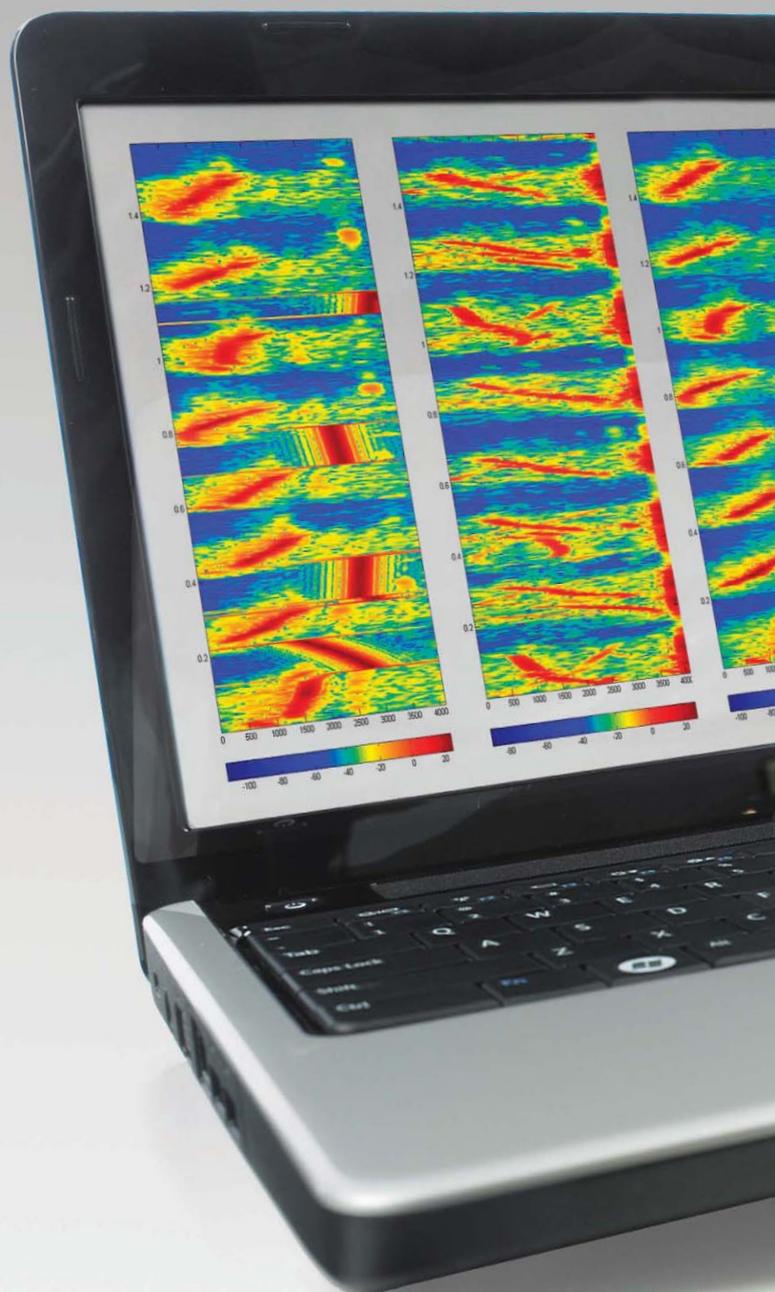
Find it at
mathworks.com/accelerate
datasheet
video example
trial request

RUN MATLAB PROGRAMS IN PARALLEL

with
Parallel Computing Toolbox™

If you can write a FOR-loop, you can write a MATLAB parallel program for high-performance computing. Parallel Computing Toolbox lets you quickly adapt MATLAB programs to run on your multicore computer, GPU or cluster, with features for task-parallel and data-parallel programs.

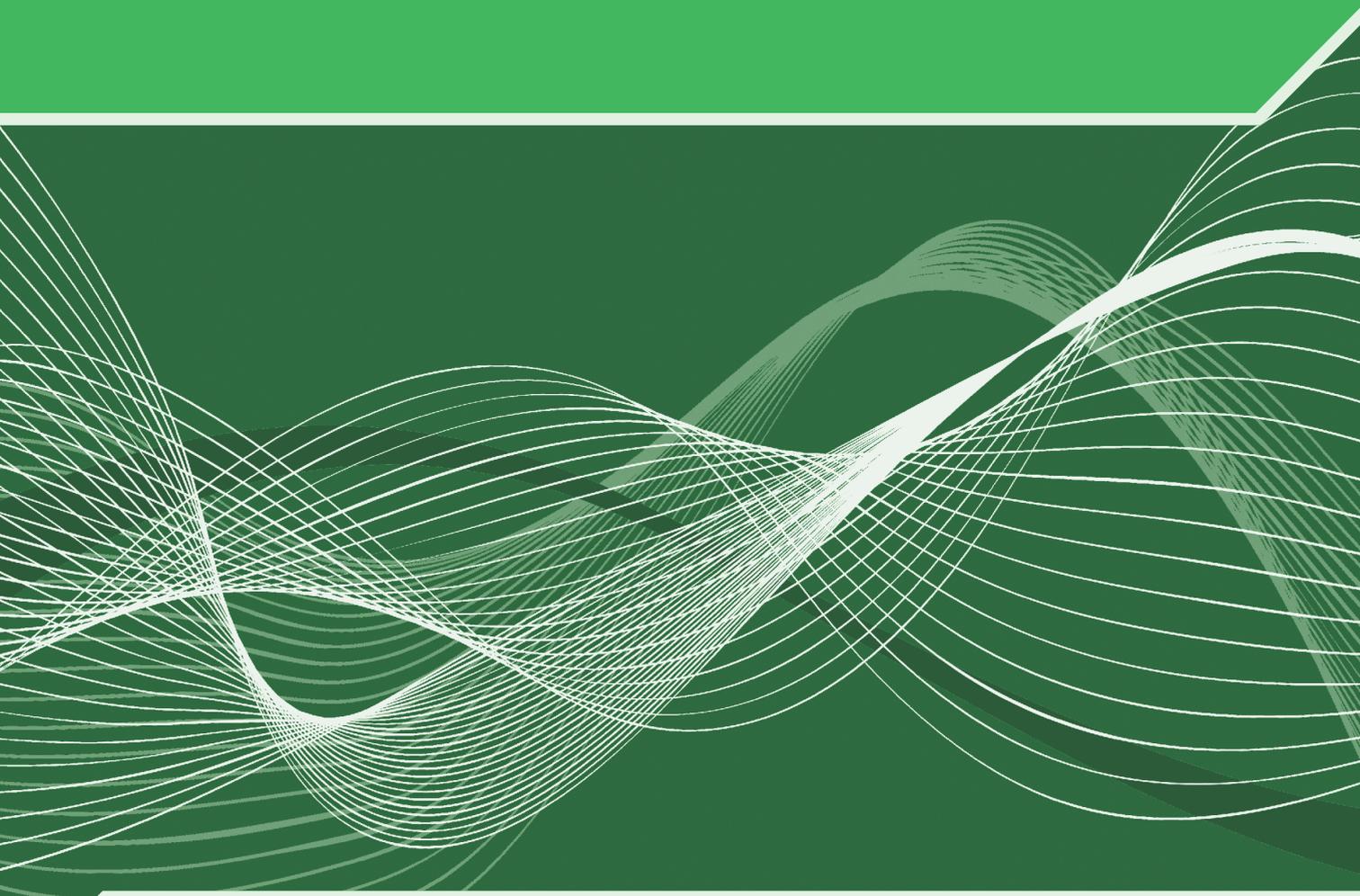
MATLAB®
& SIMULINK®



IEEE SIGNAL PROCESSING SOCIETY

CONTENT GAZETTE

[ISSN 2167-5023]



JANUARY 2014



IEEE Statistical Signal Processing Workshop 2014

29 June – 2 July, Gold Coast, Australia



General co-chairs

Rob Evans
University of Melbourne

Abd-Krim Seghouane
University of Melbourne

Technical co-chairs

Ba-Ngu Vo
Curtin University

Rod Kennedy
Australian National University

Local Arrangements

Vaughan Clarkson
University of Queensland

Tutorials

Kutluyil Doganacy
University of South Australia

Special sessions

Jean-Yves Tournieret
University of Toulouse

Government liaison

Neil Gordon
DSTO

Industry Liaison

Sven Nordholm
Curtin University

Publications

Doug Gray
University of Adelaide

Publicity

Branko Ristic
DSTO

Student activity

Igmar Land
University of South Australia

International Liaison

Steve McLaughlin
University of Edinburgh

The 2014 IEEE Workshop on Statistical Signal Processing (SSP'14) will be held at Jupiters, Gold Coast, Australia from June 29th to July 2nd 2014. The SSP Workshop is a unique meeting attended by members of the IEEE Signal Processing Society together with researchers from allied fields such as statistics, machine learning, communications and bioinformatics. It has the salient feature of having all the contributed and special sessions as poster sessions allowing extensive interaction and networking.

The scientific program of SSP 2014 will include invited plenary talks, regular and special sessions with contributed research papers. All submitted papers are reviewed by experts and only a proportion is accepted to maintain a high quality workshop. All accepted papers are published on IEEE Xplore. The scope of the workshop includes basic theory, methods and algorithms, and applications in the following areas:

Theoretical Topics:

- Adaptive systems and signal processing
- Detection and estimation theory
- Learning theory and pattern recognition
- Multivariate statistical analysis
- System identification and calibration
- Monte Carlo methods
- Network and graph analysis
- Random matrix theory
- Time-frequency and time-scale analysis
- Compressed sensing
- Point process estimation

Application Areas:

- Bioinformatics and genomics
- Array processing, radar and sonar
- Communication systems and networks
- Sensor networks
- Information forensics and security
- Medical imaging
- Biomedical signal processing
- Preventive, social network analysis
- Smart grid and industrial applications
- Geoscience
- Astrophysics

Submission of papers: Prospective authors are invited to submit full-length four-page papers, including figures and references. Papers are submitted electronically at SSP2014.org

Special session proposals: Special session proposals are submitted through the SSP2014.org. They must include a topical title, rational, session outline, contact information and the list of invited papers. Special session authors are referred to the workshop website for additional information regarding submissions.

Plenary speakers: Louis Scharf (Colorado State), Arnaud Doucet (Oxford), Michael Unser (EPFL), Victor Solo (UNSW), Brian Anderson (ANU), Steven Tingay (Curtin)

Important Dates:

- | | |
|---|---------------------------------|
| • Deadline for special session proposals submission | December 1 st , 2013 |
| • Special sessions selected | January 2 nd , 2014 |
| • Deadline for 4-page paper submissions | February 2 nd , 2014 |
| • Notification of paper acceptance | March 30 th , 2014 |
| • Camera ready upload | April 10 th , 2014 |
| • Start of conference | June 29 th , 2014 |

Venue: The Gold Coast is a coastal city located in the South East of Queensland, Australia. The city is 94 km (58 mi) south of the state capital Brisbane. With more than 300 days of sunshine each year you have the opportunity to explore more than 70 km of unspoiled coastline and beaches as well as 100,000 hectares of world heritage rainforest. The Gold Coast is one of the leading tourist destinations, boasting 30 different beaches each with their own personality, from world-famous surf breaks and cosy coastal settings in the south to family fun swimming spots and surf-life saving action in the north. At your fingertips is every on-water activity imaginable! With its sunny subtropical climate, surfing beaches, canal and waterway systems, its high-rise dominated skyline, theme parks, nightlife and rainforest hinterland, Gold Coast will turn your stay into lifetime memories.

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



NOVEMBER 15, 2013

VOLUME 61

NUMBER 22

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

- Analysis of Kalman Filter Approximations for Nonlinear Measurements <http://dx.doi.org/10.1109/TSP.2013.2279367> *M. R. Morelande and Á. F. García-Fernández* 5477
- A Spatio-Temporal Array Processing for Passive Localization of Radio Transmitters <http://dx.doi.org/10.1109/TSP.2013.2278515> *J. Bosse, A. Ferréol, and P. Larzabal* 5485
- Direct Optimization of the Dictionary Learning Problem <http://dx.doi.org/10.1109/TSP.2013.2278158> *A. Rakotomamonjy* 5495
- Blind Signal Separation via Tensor Decomposition With Vandermonde Factor: Canonical Polyadic Decomposition <http://dx.doi.org/10.1109/TSP.2013.2276416> *M. Sørensen and L. De Lathauwer* 5507
- Source Separation of Convulsive and Noisy Mixtures Using Audio-Visual Dictionary Learning and Probabilistic Time-Frequency Masking <http://dx.doi.org/10.1109/TSP.2013.2277834> *Q. Liu, W. Wang, P. J. B. Jackson, M. Barnard, J. Kittler, and J. Chambers* 5520
- Orthogonal Frequency Division Multiplexing With Index Modulation <http://dx.doi.org/10.1109/TSP.2013.2279771> *E. Başar, Ü. Aygözü, E. Panayırçı, and H. V. Poor* 5536

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$537.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



MIMO Broadcast DFE Transceivers With QoS Constraints: Min-Power and Max-Rate Solutions http://dx.doi.org/10.1109/TSP.2013.2279769	<i>C. -H. Liu and P. P. Vaidyanathan</i>	5550
Frequency-Domain Approach to Relay Beamforming With Adaptive Decision Delay for Frequency-Selective Channels http://dx.doi.org/10.1109/TSP.2013.2279365	<i>T. Wang, B. P. Ng, and M. H. Er</i>	5563
Denosing Based on Multivariate Stochastic Volatility Modeling of Multiwavelet Coefficients http://dx.doi.org/10.1109/TSP.2013.2279077 ..	<i>S. H. Fouladi, M. Hajiramezani, H. Amindavar, J. A. Ritcey, and P. Arabshahi</i>	5578
Multi-Portfolio Optimization: A Potential Game Approach http://dx.doi.org/10.1109/TSP.2013.2277839	<i>Y. Yang, F. Rubio, G. Scutari, and D. P. Palomar</i>	5590
Ambiguity Function Shaping for Cognitive Radar Via Complex Quartic Optimization http://dx.doi.org/10.1109/TSP.2013.2273885	<i>A. Aubry, A. De Maio, B. Jiang, and S. Zhang</i>	5603
Sparse and Non-Negative BSS for Noisy Data http://dx.doi.org/10.1109/TSP.2013.2279358	<i>J. Rapin, J. Bobin, A. Larue, and J.-L. Starck</i>	5620
Statistical Inference in Large Antenna Arrays Under Unknown Noise Pattern http://dx.doi.org/10.1109/TSP.2013.2280443	<i>J. Vinogradova, R. Couillet, and W. Hachem</i>	5633
Joint Precoding and Multivariate Backhaul Compression for the Downlink of Cloud Radio Access Networks http://dx.doi.org/10.1109/TSP.2013.2280111	<i>S. -H. Park, O. Simeone, O. Sahin, and S. Shamai</i>	5646
Efficient Solutions for Weighted Sum Rate Maximization in Multicellular Networks With Channel Uncertainties http://dx.doi.org/10.1109/TSP.2013.2278815	<i>M. F. Hanif, L.-N. Tran, A. Tölli, M. Juntti, and S. Glisic</i>	5659
Distributed Adaptive Networks: A Graphical Evolutionary Game-Theoretic View http://dx.doi.org/10.1109/TSP.2013.2280444	<i>C. Jiang, Y. Chen, and K. J. R. Liu</i>	5675
Rank Regularization and Bayesian Inference for Tensor Completion and Extrapolation http://dx.doi.org/10.1109/TSP.2013.2278516	<i>J. A. Bazerque, G. Mateos, and G. B. Giannakis</i>	5689
Sparse Conjoint Analysis Through Maximum Likelihood Estimation http://dx.doi.org/10.1109/TSP.2013.2278529	<i>E. Tsakonas, J. Jaldén, N. D. Sidiropoulos, and B. Ottersten</i>	5704
Bayesian Simultaneous Sparse Approximation With Smooth Signals http://dx.doi.org/10.1109/TSP.2013.2280441	<i>M. Luessi, S. D. Babacan, R. Molina, and A. K. Katsaggelos</i>	5716
A Class of Bounded Component Analysis Algorithms for the Separation of Both Independent and Dependent Sources http://dx.doi.org/10.1109/TSP.2013.2280115	<i>A. T. Erdogan</i>	5730
A Fast and Accurate Reconstruction Algorithm for Compressed Sensing of Complex Sinusoids http://dx.doi.org/10.1109/TSP.2013.2280125	<i>L. Hu, J. Zhou, Z. Shi, and Q. Fu</i>	5744
Online Bayesian Inference in Some Time-Frequency Representations of Non-Stationary Processes http://dx.doi.org/10.1109/TSP.2013.2280128	<i>R. G. Everitt, C. Andrieu, and M. Davy</i>	5755
Multiple Observations HMM Learning by Aggregating Ensemble Models http://dx.doi.org/10.1109/TSP.2013.2280179	<i>N. Asadi, A. Mirzaei, and E. Haghshenas</i>	5767
RSP-Based Analysis for Sparsest and Least ℓ_1 -Norm Solutions to Underdetermined Linear Systems http://dx.doi.org/10.1109/TSP.2013.2281030	<i>Y.-B. Zhao</i>	5777
Analysis-by-Synthesis Quantization for Compressed Sensing Measurements http://dx.doi.org/10.1109/TSP.2013.2280445	<i>A. Shirazinia, S. Chatterjee, and M. Skoglund</i>	5789

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



NOVEMBER 15, 2013

VOLUME 61

NUMBER 22

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

Statistical Signal Processing

- Analysis of Kalman Filter Approximations for Nonlinear Measurements <http://dx.doi.org/10.1109/TSP.2013.2278529> *M. R. Morelande and Á. F. García-Fernández* 5477
- Sparse Conjoint Analysis Through Maximum Likelihood Estimation <http://dx.doi.org/10.1109/TSP.2013.2280441> *E. Tsakonas, J. Jaldén, N. D. Sidiropoulos, and B. Ottersten* 5704
- Bayesian Simultaneous Sparse Approximation With Smooth Signals <http://dx.doi.org/10.1109/TSP.2013.2279077> *M. Luessi, S. D. Babacan, R. Molina, and A. K. Katsaggelos* 5716
- Denosing Based on Multivariate Stochastic Volatility Modeling of Multiwavelet Coefficients <http://dx.doi.org/yes> *S. H. Fouladi, M. Hajiramezanali, H. Amindavar, J. A. Ritcey, and P. Arabshahi* 5578

Adaptive Signal Processing

- Distributed Adaptive Networks: A Graphical Evolutionary Game-Theoretic View <http://dx.doi.org/10.1109/TSP.2013.2280128> *C. Jiang, Y. Chen, and K. J. R. Liu* 5675

Digital and Multirate Signal Processing

- Online Bayesian Inference in Some Time-Frequency Representations of Non-Stationary Processes <http://dx.doi.org/10.1109/TSP.2013.2278158> *R. G. Everitt, C. Andrieu, and M. Davy* 5755
- Direct Optimization of the Dictionary Learning Problem <http://dx.doi.org/10.1109/TSP.2013.2281030> *A. Rakotomamonjy* 5495
- RSP-Based Analysis for Sparsest and Least ℓ_1 -Norm Solutions to Underdetermined Linear Systems <http://dx.doi.org/10.1109/TSP.2013.2280179> *Y.-B. Zhao* 5777

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), non-members \$537.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



<i>Multidimensional Signal Processing</i>	
Rank Regularization and Bayesian Inference for Tensor Completion and Extrapolation http://dx.doi.org/10.1109/TSP.2013.2277834	5689
..... J. A. Bazerque, G. Mateos, and G. B. Giannakis	
<i>Machine Learning</i>	
Multiple Observations HMM Learning by Aggregating Ensemble Models http://dx.doi.org/10.1109/TSP.2013.2280443	5767
..... N. Asadi, A. Mirzaei, and E. Haghshenas	
Source Separation of Convolutional and Noisy Mixtures Using Audio-Visual Dictionary Learning and Probabilistic Time-Frequency Masking http://dx.doi.org/10.1109/TSP.2013.2273885	5520
..... Q. Liu, W. Wang, P. J. B. Jackson, M. Barnard, J. Kittler, and J. Chambers	
Sparse and Non-Negative BSS for Noisy Data http://dx.doi.org/10.1109/TSP.2013.2276416	5620
..... J. Rapin, J. Bobin, A. Larue, and J.-L. Starck	
A Class of Bounded Component Analysis Algorithms for the Separation of Both Independent and Dependent Sources http://dx.doi.org/10.1109/TSP.2013.2280445	5730
..... A. T. Erdogan	
<i>Sensor Array and Multichannel Processing</i>	
Statistical Inference in Large Antenna Arrays Under Unknown Noise Pattern http://dx.doi.org/10.1109/TSP.2013.2279771	5633
..... J. Vinogradov, R. Couillet, and W. Hachem	
Ambiguity Function Shaping for Cognitive Radar Via Complex Quartic Optimization http://dx.doi.org/10.1109/TSP.2013.2279769	5603
..... A. Aubry, A. De Maio, B. Jiang, and S. Zhang	
<i>Signal Processing for Communications</i>	
Blind Signal Separation via Tensor Decomposition With Vandermonde Factor: Canonical Polyadic Decomposition http://dx.doi.org/10.1109/TSP.2013.2278815	5507
..... M. Sørensen and L. De Lathauwer	
Analysis-by-Synthesis Quantization for Compressed Sensing Measurements http://dx.doi.org/10.1109/TSP.2013.2278515	5789
..... A. Shirazinia, S. Chatterjee, and M. Skoglund	
Orthogonal Frequency Division Multiplexing With Index Modulation http://dx.doi.org/10.1109/TSP.2013.2279365	5536
..... E. Başar, Ü. Aygözü, E. Panayırçı, and H. V. Poor	
<i>MIMO Communications & Signal Processing</i>	
MIMO Broadcast DFE Transceivers With QoS Constraints: Min-Power and Max-Rate Solutions http://dx.doi.org/10.1109/TSP.2013.2280111	5550
..... C.-H. Liu and P. P. Vaidyanathan	
Efficient Solutions for Weighted Sum Rate Maximization in Multicellular Networks With Channel Uncertainties http://dx.doi.org/10.1109/TSP.2013.2277839	5659
..... M. F. Hanif, L.-N. Tran, A. Tölli, M. Juntti, and S. Glisic	
<i>Signal Processing for Sensor Networks</i>	
A Spatio-Temporal Array Processing for Passive Localization of Radio Transmitters http://dx.doi.org/10.1109/TSP.2013.2280125	5485
..... J. Bosse, A. Ferréol, and P. Larzabal	
<i>Signal Processing for Wireless Networks</i>	
Frequency-Domain Approach to Relay Beamforming With Adaptive Decision Delay for Frequency-Selective Channels http://dx.doi.org/	5563
..... T. Wang, B. P. Ng, and M. H. Er	
Joint Precoding and Multivariate Backhaul Compression for the Downlink of Cloud Radio Access Networks http://dx.doi.org/ ..	5646
..... S.-H. Park, O. Simeone, O. Sahin, and S. Shamai	
<i>Other Areas and Applications</i>	
Multi-Portfolio Optimization: A Potential Game Approach http://dx.doi.org/	5590
..... Y. Yang, F. Rubio, G. Scutari, and D. P. Palomar	
A Fast and Accurate Reconstruction Algorithm for Compressed Sensing of Complex Sinusoids http://dx.doi.org/	5744
..... L. Hu, J. Zhou, Z. Shi, and Q. Fu	

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



DECEMBER 1, 2013

VOLUME 61

NUMBER 23

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

- Regularized Covariance Matrix Estimation in Complex Elliptically Symmetric Distributions Using the Expected Likelihood Approach—Part 1: The Over-Sampled Case <http://dx.doi.org/10.1109/TSP.2013.2272924> *Y. I. Abramovich and O. Besson* 5807
- Regularized Covariance Matrix Estimation in Complex Elliptically Symmetric Distributions Using the Expected Likelihood Approach—Part 2: The Under-Sampled Case <http://dx.doi.org/10.1109/TSP.2013.2285511> *O. Besson and Y. I. Abramovich* 5819
- Cooperative Wireless Sensor Network Positioning via Implicit Convex Feasibility <http://dx.doi.org/10.1109/TSP.2013.2279770> *M. R. Gholami, L. Tetrushvili, E. G. Ström, and Y. Censor* 5830
- Beamforming Design for Multiuser Two-Way Relaying: A Unified Approach via Max-Min SINR <http://dx.doi.org/10.1109/TSP.2013.2281032> *Z. Fang, X. Wang, and X. Yuan* 5841
- A Variational Bayesian Learning Approach for Nonlinear Acoustic Echo Control <http://dx.doi.org/10.1109/TSP.2013.2281021> *S. Malik and G. Enzner* 5853

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), non-members \$537.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



Network-Wide Distributed Carrier Frequency Offsets Estimation and Compensation via Belief Propagation http://dx.doi.org/10.1109/TSP.2013.2281037	<i>J. Du and Y.-C. Wu</i>	5868
Exploration of Lattice Reduction Aided Soft-Output MIMO Detection on a DLP/ILP Baseband Processor http://dx.doi.org/10.1109/TSP.2013.2279773	<i>U. Ahmad, M. Li, R. Appeltans, H. Duy Nguyen, A. Amin, A. Dejonghe, L. Van der Perre, R. Lauwereins, and S. Pollin</i>	5878
Jointly Optimal High-Order Error Feedback and Realization for Roundoff Noise Minimization in 1-D and 2-D State-Space Digital Filters http://dx.doi.org/10.1109/TSP.2013.2281300	<i>T. Hinamoto, A. Doi, and W.-S. Lu</i>	5893
Fast and Accurate Algorithms for Re-Weighted ℓ_1 -Norm Minimization http://dx.doi.org/10.1109/TSP.2013.2279362	<i>M. S. Asif and J. Romberg</i>	5905
A Max-Product EM Algorithm for Reconstructing Markov-Tree Sparse Signals From Compressive Samples http://dx.doi.org/10.1109/TSP.2013.2277833	<i>Z. Song and A. Dogandžić</i>	5917
On the Pareto-Optimal Beam Structure and Design for Multi-User MIMO Interference Channels http://dx.doi.org/10.1109/TSP.2013.2281784	<i>J. Park and Y. Sung</i>	5932
PETRELS: Parallel Subspace Estimation and Tracking by Recursive Least Squares From Partial Observations http://dx.doi.org/10.1109/TSP.2013.2282910	<i>Y. Chi, Y. C. Eldar, and R. Calderbank</i>	5947
Parameter Estimation For Multivariate Generalized Gaussian Distributions http://dx.doi.org/10.1109/TSP.2013.2282909	<i>F. Pascal, L. Bombrun, J.-Y. Tournet, and Y. Berthoumieu</i>	5960
Simultaneous Information and Power Transfer for Broadband Wireless Systems http://dx.doi.org/10.1109/TSP.2013.2281026	<i>K. Huang and E. G. Larsson</i>	5972
Atomic Norm Denoising With Applications to Line Spectral Estimation http://dx.doi.org/10.1109/TSP.2013.2273443	<i>B. N. Bhaskar, G. Tang, and B. Recht</i>	5987
Non-Linear Distributed Average Consensus Using Bounded Transmissions http://dx.doi.org/10.1109/TSP.2013.2282912	<i>S. Dasarathan, C. Tepedelenliolu, M. Banavar, and A. Spanias</i>	6000
Multi-Observer Privacy-Preserving Hidden Markov Models http://dx.doi.org/10.1109/TSP.2013.2282911 ...	<i>H. X. Nguyen and M. Roughan</i>	6010
Random Field Reconstruction With Quantization in Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2280442	<i>I. Nevat, G. W. Peters, and I. B. Collings</i>	6020
L -infinity Norm Design of Linear-Phase Robust Broadband Beamformers Using Constrained Optimization http://dx.doi.org/10.1109/TSP.2013.2283463	<i>R. C. Nongpiur and D. J. Shpak</i>	6034
Spectrum- and Energy-Efficient OFDM Based on Simultaneous Multi-Channel Reconstruction http://dx.doi.org/10.1109/TSP.2013.2282920	<i>L. Dai, J. Wang, Z. Wang, P. Tsiaklakis, and M. Moonen</i>	6047
A Unified Approach to Structured Covariances: Fast Generalized Sliding Window RLS Recursions for Arbitrary Basis http://dx.doi.org/10.1109/TSP.2013.2278150	<i>R. Merched</i>	6060
Unicast Multi-Antenna Relay Beamforming With Per-Antenna Power Control: Optimization and Duality http://dx.doi.org/10.1109/TSP.2013.2281778	<i>M. Dong, B. Liang, and Q. Xiao</i>	6076
Limited Feedback of Precoder and Bit Loading for MIMO Systems: A Joint Design http://dx.doi.org/10.1109/TSP.2013.2281785	<i>H.-C. Chen and Y.-P. Lin</i>	6091
Robust Adaptive Beamforming for General-Rank Signal Model With Positive Semi-Definite Constraint via POTDC http://dx.doi.org/10.1109/TSP.2013.2281301	<i>A. Khabbazibasmenj and S. A. Vorobyov</i>	6103
Improved Source Number Detection and Direction Estimation With Nested Arrays and ULAs Using Jackknifing http://dx.doi.org/10.1109/TSP.2013.2283462	<i>K. Han and A. Nehorai</i>	6118

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



DECEMBER 1, 2013

VOLUME 61

NUMBER 23

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

Statistical Signal Processing

- A Max-Product EM Algorithm for Reconstructing Markov-Tree Sparse Signals From Compressive Samples
<http://dx.doi.org/10.1109/TSP.2013.2277833> Z. Song and A. Dogandžić 5917
- Parameter Estimation For Multivariate Generalized Gaussian Distributions
<http://dx.doi.org/10.1109/TSP.2013.2282909> F. Pascal, L. Bombrun, J.-Y. Tournet, and Y. Berthoumieu 5960
- PETRELS: Parallel Subspace Estimation and Tracking by Recursive Least Squares From Partial Observations
<http://dx.doi.org/10.1109/TSP.2013.2282910> Y. Chi, Y. C. Eldar, and R. Calderbank 5947
- Atomic Norm Denoising With Applications to Line Spectral Estimation
<http://dx.doi.org/10.1109/TSP.2013.2273443> B. N. Bhaskar, G. Tang, and B. Recht 5987
- Multi-Observer Privacy-Preserving Hidden Markov Models
<http://dx.doi.org/10.1109/TSP.2013.2282911> ... H. X. Nguyen and M. Roughan 6010

Adaptive Signal Processing

- A Unified Approach to Structured Covariances: Fast Generalized Sliding Window RLS Recursions for Arbitrary Basis
<http://dx.doi.org/10.1109/TSP.2013.2278150> R. Merched 6060
- A Variational Bayesian Learning Approach for Nonlinear Acoustic Echo Control
<http://dx.doi.org/10.1109/TSP.2013.2281021> S. Malik and G. Enzner 5853

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), non-members \$537.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



<i>Digital and Multirate Signal Processing</i>	
Jointly Optimal High-Order Error Feedback and Realization for Roundoff Noise Minimization in 1-D and 2-D State-Space Digital Filters http://dx.doi.org/10.1109/TSP.2013.2281300	<i>T. Hinamoto, A. Doi, and W.-S. Lu</i> 5893
Fast and Accurate Algorithms for Re-Weighted ℓ_1 -Norm Minimization http://dx.doi.org/10.1109/TSP.2013.2279362	<i>M. S. Asif and J. Romberg</i> 5905
<i>Machine Learning</i>	
Random Field Reconstruction With Quantization in Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2280442	<i>I. Nevat, G. W. Peters, and I. B. Collings</i> 6020
<i>Sensor Array and Multichannel Processing</i>	
Regularized Covariance Matrix Estimation in Complex Elliptically Symmetric Distributions Using the Expected Likelihood Approach—Part 1: The Over-Sampled Case http://dx.doi.org/10.1109/TSP.2013.2272924	<i>Y. I. Abramovich and O. Besson</i> 5807
Regularized Covariance Matrix Estimation in Complex Elliptically Symmetric Distributions Using the Expected Likelihood Approach—Part 2: The Under-Sampled Case http://dx.doi.org/10.1109/TSP.2013.2285511	<i>O. Besson and Y. I. Abramovich</i> 5819
Robust Adaptive Beamforming for General-Rank Signal Model With Positive Semi-Definite Constraint via POTDC http://dx.doi.org/10.1109/TSP.2013.2281301	<i>A. Khabbazibasmenj and S. A. Vorobyov</i> 6103
L -infinity Norm Design of Linear-Phase Robust Broadband Beamformers Using Constrained Optimization http://dx.doi.org/10.1109/TSP.2013.2283463	<i>R. C. Nongpiur and D. J. Shpak</i> 6034
Improved Source Number Detection and Direction Estimation With Nested Arrays and ULAs Using Jackknifing http://dx.doi.org/10.1109/TSP.2013.2283462	<i>K. Han and A. Nehorai</i> 6118
<i>Signal Processing for Communications</i>	
Network-Wide Distributed Carrier Frequency Offsets Estimation and Compensation via Belief Propagation http://dx.doi.org/10.1109/TSP.2013.2281037	<i>J. Du and Y.-C. Wu</i> 5868
Spectrum- and Energy-Efficient OFDM Based on Simultaneous Multi-Channel Reconstruction http://dx.doi.org/10.1109/TSP.2013.2282920	<i>L. Dai, J. Wang, Z. Wang, P. Tsiaflakis, and M. Moonen</i> 6047
Beamforming Design for Multiuser Two-Way Relaying: A Unified Approach via Max-Min SINR http://dx.doi.org/10.1109/TSP.2013.2281032	<i>Z. Fang, X. Wang, and X. Yuan</i> 5841
<i>MIMO Communications & Signal Processing</i>	
On the Pareto-Optimal Beam Structure and Design for Multi-User MIMO Interference Channels http://dx.doi.org/10.1109/TSP.2013.2281784	<i>J. Park and Y. Sung</i> 5932
Limited Feedback of Precoder and Bit Loading for MIMO Systems: A Joint Design http://dx.doi.org/10.1109/TSP.2013.2281785	<i>H.-C. Chen and Y.-P. Lin</i> 6091
<i>Signal Processing for Sensor Networks</i>	
Non-Linear Distributed Average Consensus Using Bounded Transmissions http://dx.doi.org/10.1109/TSP.2013.2282912	<i>S. Dasarathan, C. Tepedelenliolu, M. Banavar, and A. Spanias</i> 6000
Cooperative Wireless Sensor Network Positioning via Implicit Convex Feasibility http://dx.doi.org/10.1109/TSP.2013.2279770	<i>M. R. Gholami, L. Tetrushvili, E. G. Ström, and Y. Censor</i> 5830
<i>Signal Processing for Wireless Networks</i>	
Unicast Multi-Antenna Relay Beamforming With Per-Antenna Power Control: Optimization and Duality http://dx.doi.org/10.1109/TSP.2013.2281778	<i>M. Dong, B. Liang, and Q. Xiao</i> 6076
<i>Implementation of Signal Processing Systems</i>	
Exploration of Lattice Reduction Aided Soft-Output MIMO Detection on a DLP/ILP Baseband Processor http://dx.doi.org/10.1109/TSP.2013.2279773	<i>U. Ahmad, M. Li, R. Appeltans, H. Duy Nguyen, A. Amin, A. Dejonghe, L. Van der Perre, R. Lauwereins, and S. Pollin</i> 5878
Simultaneous Information and Power Transfer for Broadband Wireless Systems http://dx.doi.org/10.1109/TSP.2013.2281026	<i>K. Huang and E. G. Larsson</i> 5972

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



NOVEMBER 2013

VOLUME 21

NUMBER 11

ITASD8

(ISSN 1558-7916)

SPECIAL SECTION ON LARGE-SCALE OPTIMIZATION FOR AUDIO, SPEECH, AND LANGUAGE PROCESSING

EDITORIAL

Introduction to the Special Section on Large-Scale Optimization for Audio, Speech, and Language Processing
<http://dx.doi.org/10.1109/TASL.2013.2283631> *D. Kanevsky, X. He, G. Heigold, H. Li, and S. J. Wright* 2229

PAPERS

Optimization Algorithms and Applications for Speech and Language Processing <http://dx.doi.org/10.1109/TASL.2013.2283777>
..... *S. J. Wright, D. Kanevsky, L. Deng, X. He, G. Heigold, and H. Li* 2231

Second Order Methods for Optimizing Convex Matrix Functions and Sparse Covariance Clustering
<http://dx.doi.org/10.1109/TASL.2013.2263142> *G. M. Chin, J. Nocedal, P. A. Olsen, and S. J. Rennie* 2244

A Difference of Convex Functions Approach to Large-Scale Log-Linear Model Estimation <http://dx.doi.org/10.1109/TASL.2013.2271592> ..
..... *T. Tsiligkaridis, E. Marcheret, and V. Goel* 2255

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 1558-7916) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$307.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



Optimization Techniques to Improve Training Speed of Deep Neural Networks for Large Speech Tasks http://dx.doi.org/10.1109/TASL.2013.2284378	<i>T. N. Sainath, B. Kingsbury, H. Soltau, and B. Ramabhadran</i>	2267
Active-Set Newton Algorithm for Overcomplete Non-Negative Representations of Audio http://dx.doi.org/10.1109/TASL.2013.2263144	<i>T. Virtanen, J. F. Gemmeke, and B. Raj</i>	2277
Large Vocabulary Speech Recognition on Parallel Architectures http://dx.doi.org/10.1109/TASL.2013.2271591	<i>P. Cardinal, P. Dumouchel, and G. Boulianne</i>	2290

REGULAR PAPERS

Room Reverberation Reconstruction: Interpolation of the Early Part Using Compressed Sensing http://dx.doi.org/10.1109/TASL.2013.2273662	<i>R. Mignot, L. Daudet, and F. Ollivier</i>	2301
Improving Graph-Based Dependency Parsing Models With Dependency Language Models http://dx.doi.org/10.1109/TASL.2013.2273715 ..	<i>M. Zhang, W. Chen, X. Duan, and R. Zhang</i>	2313
Robust Ultra-Low Latency Soft-Decision Decoding of Linear PCM Audio http://dx.doi.org/10.1109/TASL.2013.2273716	<i>F. Pflug and T. Fingscheidt</i>	2324
Scalable Speech Coding for IP Networks: Beyond iLBC http://dx.doi.org/10.1109/TASL.2013.2274694	<i>K. Seto and T. Ogunfunmi</i>	2337
Diffused Sensing for Sharp Directive Beamforming http://dx.doi.org/10.1109/TASL.2013.2274695	<i>K. Niwa, Y. Hioka, K. Furuya, and Y. Haneda</i>	2346
Cross Pattern Coherence Algorithm for Spatial Filtering Applications Utilizing Microphone Arrays http://dx.doi.org/10.1109/TASL.2013.2277928	<i>S. Delikaris-Manias and V. Pulkki</i>	2356
Robust SVD-Based Audio Watermarking Scheme With Differential Evolution Optimization http://dx.doi.org/10.1109/TASL.2013.2277929 ..	<i>B. Lei, I. Y. Soon, and E. L. Tan</i>	2368
Distributional Semantic Models for Affective Text Analysis http://dx.doi.org/10.1109/TASL.2013.2277931	<i>N. Malandrakis, A. Potamianos, E. Iosif, and S. Narayanan</i>	2379
Passive Temporal Offset Estimation of Multichannel Recordings of an Ad-Hoc Microphone Array http://dx.doi.org/10.1109/TASL.2013.2286921	<i>Pasi Pertilä, Matti S. Hämäläinen, and Mikael Mieskolainen</i>	2393
Convergence Analysis of Narrowband Feedback Active Noise Control System With Imperfect Secondary Path Estimation http://dx.doi.org/10.1109/TASL.2013.2277934	<i>L. V. Wang, W.-S. Gan, A. W. H. Khong, and S. M. Kuo</i>	2403
Robust Segments Detector for De-Synchronization Resilient Audio Watermarking http://dx.doi.org/10.1109/TASL.2013.2279312	<i>C.-M. Pun and X.-C. Yuan</i>	2412
Quality Measure Functions for Calibration of Speaker Recognition Systems in Various Duration Conditions http://dx.doi.org/10.1109/TASL.2013.2279332	<i>M. I. Mandasari, R. Saeidi, M. McLaren, and D. A. van Leeuwen</i>	2425
Acoustic Modeling With Hierarchical Reservoirs http://dx.doi.org/10.1109/TASL.2013.2280209	<i>F. Triefenbach, A. Jalalvand, K. Demuynck, and J. -P. Martens</i>	2439
Unsupervised Spoken Language Understanding for a Multi-Domain Dialog System http://dx.doi.org/10.1109/TASL.2013.2280212	<i>D. Lee, M. Jeong, K. Kim, S. Ryu, and G. G. Lee</i>	2451

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



NOVEMBER 2013

VOLUME 21

NUMBER 11

ITASD8

(ISSN 1558-7916)

SPECIAL SECTION ON LARGE-SCALE OPTIMIZATION FOR AUDIO, SPEECH, AND LANGUAGE PROCESSING

EDITORIAL

Introduction to the Special Section on Large-Scale Optimization for Audio, Speech, and Language Processing http://dx.doi.org/10.1109/TASL.2013.2283631	<i>D. Kanevsky, X. He, G. Heigold, H. Li, and S. J. Wright</i>	2229
---	--	------

PAPERS

Optimization Algorithms and Applications for Speech and Language Processing http://dx.doi.org/10.1109/TASL.2013.2283777	<i>S. J. Wright, D. Kanevsky, L. Deng, X. He, G. Heigold, and H. Li</i>	2231
Second Order Methods for Optimizing Convex Matrix Functions and Sparse Covariance Clustering http://dx.doi.org/10.1109/TASL.2013.2263142	<i>G. M. Chin, J. Nocedal, P. A. Olsen, and S. J. Rennie</i>	2244
A Difference of Convex Functions Approach to Large-Scale Log-Linear Model Estimation http://dx.doi.org/10.1109/TASL.2013.2271592 ..	<i>T. Tsiligkaridis, E. Marcheret, and V. Goel</i>	2255
Optimization Techniques to Improve Training Speed of Deep Neural Networks for Large Speech Tasks http://dx.doi.org/10.1109/TASL.2013.2284378	<i>T. N. Sainath, B. Kingsbury, H. Soltau, and B. Ramabhadran</i>	2267

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 1558-7916) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$307.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



Active-Set Newton Algorithm for Overcomplete Non-Negative Representations of Audio http://dx.doi.org/10.1109/TASL.2013.2263144	2277
..... <i>T. Virtanen, J. F. Gemmeke, and B. Raj</i>	
Large Vocabulary Speech Recognition on Parallel Architectures http://dx.doi.org/10.1109/TASL.2013.2271591	2290
..... <i>P. Cardinal, P. Dumouchel, and G. Boulianne</i>	

REGULAR PAPERS

Room Acoustics and Acoustic System Modeling

Room Reverberation Reconstruction: Interpolation of the Early Part Using Compressed Sensing http://dx.doi.org/10.1109/TASL.2013.2273662	2301
..... <i>R. Mignot, L. Daudet, and F. Ollivier</i>	

Loudspeaker and Microphone Array Signal Processing

Diffused Sensing for Sharp Directive Beamforming http://dx.doi.org/10.1109/TASL.2013.2274695	2346
..... <i>K. Niwa, Y. Hioka, K. Furuya, and Y. Haneda</i>	

Passive Temporal Offset Estimation of Multichannel Recordings of an Ad-Hoc Microphone Array http://dx.doi.org/10.1109/TASL.2013.2274695	2393
..... <i>Pasi Pertilä, Matti S. Hämäläinen, and Mikael Mieskolainen</i>	

Active Noise Control

Convergence Analysis of Narrowband Feedback Active Noise Control System With Imperfect Secondary Path Estimation http://dx.doi.org/10.1109/TASL.2013.2277934	2403
..... <i>L. V. Wang, W.-S. Gan, A. W. H. Khong, and S. M. Kuo</i>	

Spatial and Multichannel Audio

Cross Pattern Coherence Algorithm for Spatial Filtering Applications Utilizing Microphone Arrays http://dx.doi.org/10.1109/TASL.2013.2277928	2356
..... <i>S. Delikaris-Manias and V. Pulkki</i>	

Audio Coding

Robust Ultra-Low Latency Soft-Decision Decoding of Linear PCM Audio http://dx.doi.org/10.1109/TASL.2013.2273716	2324
..... <i>F. Pflug and T. Fingscheidt</i>	

Content-Based Music Processing

Robust Segments Detector for De-Synchronization Resilient Audio Watermarking http://dx.doi.org/10.1109/TASL.2013.2279312	2412
..... <i>C.-M. Pun and X.-C. Yuan</i>	

Audio Processing Systems

Robust SVD-Based Audio Watermarking Scheme With Differential Evolution Optimization http://dx.doi.org/10.1109/TASL.2013.2277929	2368
..... <i>B. Lei, I. Y. Soon, and E. L. Tan</i>	

Speech Coding

Scalable Speech Coding for IP Networks: Beyond iLBC http://dx.doi.org/10.1109/TASL.2013.2274694	2337
..... <i>K. Seto and T. Ogunfunmi</i>	

Acoustic Modeling for Automatic Speech Recognition

Acoustic Modeling With Hierarchical Reservoirs http://dx.doi.org/10.1109/TASL.2013.2280209	2439
..... <i>F. Triefenbach, A. Jalalvand, K. Demuyneck, and J. -P. Martens</i>	

Speaker Recognition and Characterization

Quality Measure Functions for Calibration of Speaker Recognition Systems in Various Duration Conditions http://dx.doi.org/10.1109/TASL.2013.2279332	2425
..... <i>M. I. Mandasari, R. Saeidi, M. McLaren, and D. A. van Leeuwen</i>	

Spoken Language Understanding

Distributional Semantic Models for Affective Text Analysis http://dx.doi.org/10.1109/TASL.2013.2277931	2379
..... <i>N. Malandrakis, A. Potamianos, E. Iosif, and S. Narayanan</i>	

Unsupervised Spoken Language Understanding for a Multi-Domain Dialog System http://dx.doi.org/10.1109/TASL.2013.2280212	2451
..... <i>D. Lee, M. Jeong, K. Kim, S. Ryu, and G. G. Lee</i>	

Human Language Acquisition, Development, and Learning

Improving Graph-Based Dependency Parsing Models With Dependency Language Models http://dx.doi.org/10.1109/TASL.2013.2273715	2313
..... <i>M. Zhang, W. Chen, X. Duan, and R. Zhang</i>	

IEEE TRANSACTIONS ON IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



OCTOBER 2013

VOLUME 22

NUMBER 10

IIPRE4

(ISSN 1057-7149)

PAPERS

Non-Uniform Deblurring in HDR Image Reconstruction http://dx.doi.org/10.1109/TIP.2013.2257809	3739
..... C. S. Vijay, C. Paramanand, A. N. Rajagopalan, and R. Chellappa	
Towards Online Iris and Periocular Recognition Under Relaxed Imaging Constraints http://dx.doi.org/10.1109/TIP.2013.2260165	3751
..... C.-W. Tan and A. Kumar	
Cluster-Based Co-Saliency Detection http://dx.doi.org/10.1109/TIP.2013.2260166	3766
..... H. Fu, X. Cao, and Z. Tu	
Hierarchical Super-Resolution-Based Inpainting http://dx.doi.org/10.1109/TIP.2013.2261308	3779
..... O. L. Meur, M. Ebdelli, and C. Guillemot	
On the Method of Logarithmic Cumulants for Parametric Probability Density Function Estimation http://dx.doi.org/10.1109/TIP.2013.2262285	3791
..... V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia	
Two-Dimensional Maximum Local Variation Based on Image Euclidean Distance for Face Recognition http://dx.doi.org/10.1109/TIP.2013.2262286	3807
..... Q. Gao, F. Gao, H. Zhang, X.-J. Hao, and X. Wang	
Luma-Chroma Space Filter Design for Subpixel-Based Monochrome Image Downsampling http://dx.doi.org/10.1109/TIP.2013.2262288 ..	3818
..... L. Fang, O. C. Au, N.-M. Cheung, A. K. Katsaggelos, H. Li, and F. Zou	
Image Size Invariant Visual Cryptography for General Access Structures Subject to Display Quality Constraints http://dx.doi.org/10.1109/TIP.2013.2262290	3830
..... K.-H. Lee and P.-L. Chiu	
An Efficient Algorithm for Multiphase Image Segmentation with Intensity Bias Correction http://dx.doi.org/10.1109/TIP.2013.2262291	3842
..... H. Zhang, X. Ye, and Y. Chen	
Monocular Human Motion Tracking by Using DE-MC Particle Filter http://dx.doi.org/10.1109/TIP.2013.2263146	3852
..... M. Du, X. Nan, and L. Guan	
Sparse Texture Active Contour http://dx.doi.org/10.1109/TIP.2013.2263147	3866
..... Y. Gao, S. Bouix, M. Shenton, and A. Tannenbaum	
Video Texture Synthesis with Multi-frame LBP-TOP and Diffeomorphic Growth Model http://dx.doi.org/10.1109/TIP.2013.2263148	3879
..... Y. Guo, G. Zhao, Z. Zhou, and M. Pietikäinen	
A Modified Model of the Just Noticeable Depth Difference and Its Application to Depth Sensation Enhancement http://dx.doi.org/10.1109/TIP.2013.2263150	3892
..... S.-W. Jung	
Sparse Tensor Discriminant Analysis http://dx.doi.org/10.1109/TIP.2013.2264678	3904
..... Z. Lai, Y. Xu, J. Yang, J. Tang, and D. Zhang	
Intra Prediction Based on Markov Process Modeling of Images http://dx.doi.org/10.1109/TIP.2013.2264679	3916
..... F. Kamisli	



Spline-Based Deforming Ellipsoids for Interactive 3D Bioimage Segmentation http://dx.doi.org/10.1109/TIP.2013.2264680	3926
..... <i>R. Delgado-Gonzalo, N. Chenouard, and M. Unser</i>	
Near-Optimal Compressed Sensing Guarantees for Total Variation Minimization http://dx.doi.org/10.1109/TIP.2013.2264681	3941
..... <i>D. Needell and R. Ward</i>	
Bandlimited Reconstruction of Multidimensional Images from Irregular Samples http://dx.doi.org/10.1109/TIP.2013.2265880	3950
..... <i>X. Xu, W. Ye, and A. Entezari</i>	
Incorporating Anatomical Side Information into PET Reconstruction Using Nonlocal Regularization http://dx.doi.org/10.1109/TIP.2013.2265881	3961
..... <i>V.-G. Nguyen and S.-J. Lee</i>	
DCT/DST-Based Transform Coding for Intra Prediction in Image/Video Coding http://dx.doi.org/10.1109/TIP.2013.2265882	3974
..... <i>V.-G. Nguyen and S.-J. Lee</i>	
A No-Reference Perceptual Based Contrast Enhancement Metric for Ocean Scenes in Fog http://dx.doi.org/10.1109/TIP.2013.2265884	3982
..... <i>K. B. Gibson and T. Q. Nguyen</i>	
Compressive Blind Image Deconvolution http://dx.doi.org/10.1109/TIP.2013.2266100	3994
..... <i>B. Amizic, L. Spinoulas, R. Molina, and A. K. Katsaggelos</i>	
Sparse Feature Fidelity for Perceptual Image Quality Assessment http://dx.doi.org/10.1109/TIP.2013.2266579	4007
..... <i>H.-W. Chang, H. Yang, Y. Gan, and M.-H. Wang</i>	
Object-Level Image Segmentation Using Low Level Cues http://dx.doi.org/10.1109/TIP.2013.2268973	4019
..... <i>H. Zhu, J. Zheng, J. Cai, and N. M. Thalmann</i>	
A Combined Microphone and Camera Calibration Technique With Application to Acoustic Imaging http://dx.doi.org/10.1109/TIP.2013.2268974	4028
..... <i>M. Legg and S. Bradley</i>	
Pose Estimation With Segmentation Consistency http://dx.doi.org/10.1109/TIP.2013.2268975	4040
..... <i>H. Lu, X. Shao, and Y. Xiao</i>	
Noise-Resistant Local Binary Pattern With an Embedded Error-Correction Mechanism http://dx.doi.org/10.1109/TIP.2013.2268976	4049
..... <i>J. Ren, X. Jiang, and J. Yuan</i>	
Low Latency Secondary Transforms for Intra/Inter Prediction Residual http://dx.doi.org/10.1109/TIP.2013.2270087	4061
..... <i>A. Saxena and F. C. Fernandes</i>	
Reducing Integrability Error of Color Tensor Gradients for Image Fusion http://dx.doi.org/10.1109/TIP.2013.2270108	4072
..... <i>R. Montagna and G. D. Finlayson</i>	
A Curve Evolution Approach for Unsupervised Segmentation of Images With Low Depth of Field http://dx.doi.org/10.1109/TIP.2013.2270110	4086
..... <i>J. Mei, Y. Si, and H. Gao</i>	
C4: A Real-Time Object Detection Framework http://dx.doi.org/10.1109/TIP.2013.2270111	4096
..... <i>J. Wu, N. Liu, C. Geyer, and J. M. Rehg</i>	
DSIM: A DisSIMilarity-Based Image Clutter Metric for Targeting Performance http://dx.doi.org/10.1109/TIP.2013.2270112	4108
..... <i>D. Xu and Z. Shi</i>	
3D Superalloy Grain Segmentation Using a Multichannel Edge-Weighted Centroidal Voronoi Tessellation Algorithm http://dx.doi.org/10.1109/TIP.2013.2270113	4123
..... <i>Y. Cao, L. Ju, Y. Zhou, and S. Wang</i>	
A Decoupled Approach to Illumination-Robust Optical Flow Estimation http://dx.doi.org/10.1109/TIP.2013.2270374	4136
..... <i>A. Kumar, F. Tung, A. Wong, and D. A. Clausi</i>	
EDICS-Editors Information Classification Scheme	4148
Information for Authors	4149

IEEE TRANSACTIONS ON IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



OCTOBER 2013

VOLUME 22

NUMBER 10

IIPRE4

(ISSN 1057-7149)

PAPERS

<i>Image and Video Sensing and Acquisition</i>		
Near-Optimal Compressed Sensing Guarantees for Total Variation Minimization http://dx.doi.org/10.1109/TIP.2013.2264681	<i>D. Needell and R. Ward</i>	3941
<i>Statistical-Model Based Methods</i>		
On the Method of Logarithmic Cumulants for Parametric Probability Density Function Estimation http://dx.doi.org/10.1109/TIP.2013.2262285	<i>V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia</i>	3791
Two-Dimensional Maximum Local Variation Based on Image Euclidean Distance for Face Recognition http://dx.doi.org/10.1109/TIP.2013.2262286	<i>Q. Gao, F. Gao, H. Zhang, X.-J. Hao, and X. Wang</i>	3807
Sparse Tensor Discriminant Analysis http://dx.doi.org/10.1109/TIP.2013.2264678	<i>Z. Lai, Y. Xu, J. Yang, J. Tang, and D. Zhang</i>	3904
<i>Perception and Quality Models for Images and Video</i>		
Sparse Feature Fidelity for Perceptual Image Quality Assessment http://dx.doi.org/10.1109/TIP.2013.2266579	<i>H.-W. Chang, H. Yang, Y. Gan, and M.-H. Wang</i>	4007
<i>Partial Differential Equation Based Processing of Images and Video</i>		
Sparse Texture Active Contour http://dx.doi.org/10.1109/TIP.2013.2263147	<i>Y. Gao, S. Bouix, M. Shenton, and A. Tannenbaum</i>	3866
A Decoupled Approach to Illumination-Robust Optical Flow Estimation http://dx.doi.org/10.1109/TIP.2013.2270374	<i>A. Kumar, F. Tung, A. Wong, and D. A. Clausi</i>	4136
<i>Restoration and Enhancement</i>		
Cluster-Based Co-Saliency Detection http://dx.doi.org/10.1109/TIP.2013.2260166	<i>H. Fu, X. Cao, and Z. Tu</i>	3766
Hierarchical Super-Resolution-Based Inpainting http://dx.doi.org/10.1109/TIP.2013.2261308	<i>O. L. Meur, M. Ebdelli, and C. Guillemot</i>	3779
Luma-Chroma Space Filter Design for Subpixel-Based Monochrome Image Downsampling http://dx.doi.org/10.1109/TIP.2013.2262288	<i>L. Fang, O. C. Au, N.-M. Cheung, A. K. Katsaggelos, H. Li, and F. Zou</i>	3818



A No-Reference Perceptual Based Contrast Enhancement Metric for Ocean Scenes in Fog http://dx.doi.org/10.1109/TIP.2013.2265884	3982
..... <i>K. B. Gibson and T. Q. Nguyen</i>	
DSIM: A DisSIMilarity-Based Image Clutter Metric for Targeting Performance http://dx.doi.org/10.1109/TIP.2013.2270112	4108
..... <i>D. Xu and Z. Shi</i>	
<i>Formation and Reconstruction</i>	
Bandlimited Reconstruction of Multidimensional Images from Irregular Samples http://dx.doi.org/10.1109/TIP.2013.2265880	3950
..... <i>X. Xu, W. Ye, and A. Entezari</i>	
Compressive Blind Image Deconvolution http://dx.doi.org/10.1109/TIP.2013.2266100	3994
..... <i>B. Amizic, L. Spinoulas, R. Molina, and A. K. Katsaggelos</i>	
<i>Lossy Coding of Images and Video</i>	
Intra Prediction Based on Markov Process Modeling of Images http://dx.doi.org/10.1109/TIP.2013.2264679	3916
..... <i>F. Kamisli</i>	
DCT/DST-Based Transform Coding for Intra Prediction in Image/Video Coding http://dx.doi.org/10.1109/TIP.2013.2265882	3974
..... <i>A. Saxena and F. C. Fernandes</i>	
Low Latency Secondary Transforms for Intra/Inter Prediction Residual http://dx.doi.org/10.1109/TIP.2013.2270087	4061
..... <i>A. Saxena and F. C. Fernandes</i>	
<i>Image and Video Processing for Watermarking and Security</i>	
Image Size Invariant Visual Cryptography for General Access Structures Subject to Display Quality Constraints http://dx.doi.org/10.1109/TIP.2013.2262290	3830
..... <i>K.-H. Lee and P.-L. Chiu</i>	
<i>Image Scanning and Capture</i>	
Non-Uniform Deblurring in HDR Image Reconstruction http://dx.doi.org/10.1109/TIP.2013.2257809	3739
..... <i>C. S. Vijay, C. Paramanand, A. N. Rajagopalan, and R. Chellappa</i>	
<i>Color and Multispectral Imaging</i>	
Reducing Integrability Error of Color Tensor Gradients for Image Fusion http://dx.doi.org/10.1109/TIP.2013.2270108	4072
..... <i>R. Montagna and G. D. Finlayson</i>	
<i>Stereoscopic and Multiview Processing and Display</i>	
A Modified Model of the Just Noticeable Depth Difference and Its Application to Depth Sensation Enhancement http://dx.doi.org/10.1109/TIP.2013.2263150	3892
..... <i>S. W. Jung</i>	
<i>Acoustic and Ultrasound Imaging</i>	
A Combined Microphone and Camera Calibration Technique With Application to Acoustic Imaging http://dx.doi.org/10.1109/TIP.2013.2268974	4028
..... <i>M. Legg and S. Bradley</i>	
<i>Microscopic Imaging</i>	
3D Superalloy Grain Segmentation Using a Multichannel Edge-Weighted Centroidal Voronoi Tessellation Algorithm http://dx.doi.org/10.1109/TIP.2013.2270113	4123
..... <i>Y. Cao, L. Ju, Y. Zhou, and S. Wang</i>	
<i>Tomographic Imaging</i>	
Incorporating Anatomical Side Information into PET Reconstruction Using Nonlocal Regularization http://dx.doi.org/10.1109/TIP.2013.2265881	3961
..... <i>V.-G. Nguyen and S.-J. Lee</i>	
<i>Magnetic Resonance Imaging</i>	
An Efficient Algorithm for Multiphase Image Segmentation with Intensity Bias Correction http://dx.doi.org/10.1109/TIP.2013.2262291	3842
..... <i>H. Zhang, X. Ye, and Y. Chen</i>	
<i>Region, Boundary, and Shape Analysis</i>	
Spline-Based Deforming Ellipsoids for Interactive 3D Bioimage Segmentation http://dx.doi.org/10.1109/TIP.2013.2264680	3926
..... <i>R. Delgado-Gonzalo, N. Chenouard, and M. Unser</i>	
Object-Level Image Segmentation Using Low Level Cues http://dx.doi.org/10.1109/TIP.2013.2268973	4019
..... <i>H. Zhu, J. Zheng, J. Cai, and N. M. Thalmann</i>	
Pose Estimation With Segmentation Consistency http://dx.doi.org/10.1109/TIP.2013.2268975	4040
..... <i>H. Lu, X. Shao, and Y. Xiao</i>	
A Curve Evolution Approach for Unsupervised Segmentation of Images With Low Depth of Field http://dx.doi.org/10.1109/TIP.2013.2270110	4086
..... <i>J. Mei, Y. Si, and H. Gao</i>	

<i>Image and Video Interpretation and Understanding</i>	
Monocular Human Motion Tracking by Using DE-MC Particle Filter http://dx.doi.org/10.1109/TIP.2013.2263146	3852
..... M. Du, X. Nan, and L. Guan	
Noise-Resistant Local Binary Pattern With an Embedded Error-Correction Mechanism http://dx.doi.org/10.1109/TIP.2013.2268976	4049
..... J. Ren, X. Jiang, and J. Yuan	
C4: A Real-Time Object Detection Framework http://dx.doi.org/10.1109/TIP.2013.2270111	4096
..... J. Wu, N. Liu, C. Geyer, and J. M. Rehg	
<i>Image and Video Biometric Analysis</i>	
Towards Online Iris and Periocular Recognition Under Relaxed Imaging Constraints http://dx.doi.org/10.1109/TIP.2013.2260165	3751
..... C.-W. Tan and A. Kumar	
<i>Image and Video Synthesis, Rendering, and Visualization</i>	
Video Texture Synthesis with Multi-frame LBP-TOP and Diffeomorphic Growth Model http://dx.doi.org/10.1109/TIP.2013.2263148	3879
..... Y. Guo, G. Zhao, Z. Zhou, and M. Pietikäinen	
EDICS-Editor's Information Classification Scheme	4148
Information for Authors	4149

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

OCTOBER 2013

VOLUME 8

NUMBER 10

ITIFA6

(ISSN 1556-6013)

SPECIAL ISSUE ON INTELLIGENT VIDEO SURVEILLANCE FOR PUBLIC SECURITY AND PERSONAL PRIVACY

EDITORIAL

Guest Editorial Special Issue on Intelligent Video Surveillance for Public Security and Personal Privacy
<http://dx.doi.org/10.1109/TIFS.2013.2279945> *N. Babaguchi, A. Cavallaro, R. Chellappa, F. Dufaux, and L. Wang* 1559

SPECIAL ISSUE PAPERS

Anonymization and Data Privacy

Authenticating Lossy Surveillance Video <http://dx.doi.org/10.1109/TIFS.2013.2279542> *Y. J. Ren, L. O'Gorman, L. J. Wu, F. Chang, T. L. Wood, and J. R. Zhang* 1678

An ROI Privacy Protection Scheme for H.264 Video Based on FMO and Chaos <http://dx.doi.org/10.1109/TIFS.2013.2259819> *F. Peng, X.-W. Zhu, and M. Long* 1688

Biometrics

A New View-Invariant Feature for Cross-View Gait Recognition <http://dx.doi.org/10.1109/TIFS.2013.2252342> *W. Kusakunniran, Q. Wu, J. Zhang, Y. Ma, and H. Li* 1642

Multiple Subcategories Parts-Based Representation for One Sample Face Identification <http://dx.doi.org/10.1109/TIFS.2013.2263498> *X. Zhao, X. Li, Z. Wu, Y. Fu, and Y. Liu* 1654

Face Tracking and Recognition at a Distance: A Coaxial and Concentric PTZ Camera System <http://dx.doi.org/10.1109/TIFS.2013.2261061> .. *U. Park, H.-C. Choi, A. K. Jain, and S.-W. Lee* 1665

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY (ISSN 1556-6013) is published monthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$288.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.

Surveillance

Hardware Acceleration of Background Modeling in the Compressed Domain http://dx.doi.org/10.1109/TIFS.2013.2276753	<i>S. Popa, D. Crookes, and P. Miller</i>	1562
The Large-Scale Crowd Behavior Perception Based on Spatio-Temporal Viscous Fluid Field http://dx.doi.org/10.1109/TIFS.2013.2277773 ..	<i>H. Su, H. Yang, S. Zheng, Y. Fan, and S. Wei</i>	1575
Video Anomaly Search in Crowded Scenes via Spatio-Temporal Motion Context http://dx.doi.org/10.1109/TIFS.2013.2272243	<i>Y. Cong, J. Yuan, and Y. Tang</i>	1590
Attribute Regularization Based Human Action Recognition http://dx.doi.org/10.1109/TIFS.2013.2258152	<i>Z. Zhang, C. Wang, B. Xiao, W. Zhou, and S. Liu</i>	1600
Exploiting Spatio-Temporal Scene Structure for Wide-Area Activity Analysis in Unconstrained Environments http://dx.doi.org/10.1109/TIFS.2013.2277669	<i>N. M. Nayak, Y. Zhu, and A. K. Roy-Chowdhury</i>	1610
Carried Object Detection in Videos Using Color Information http://dx.doi.org/10.1109/TIFS.2013.2279797	<i>G. Tzanidou, I. Zafar, and E. A. Edirisinghe</i>	1620
Transferring Training Instances for Convenient Cross-View Object Classification in Surveillance http://dx.doi.org/10.1109/TIFS.2013.2265089	<i>Z. Zhang, Y. Zhao, Y. Wang, J. Liu, Z. Yao, and J. Tang</i>	1632

EDICS—Editor's Information Classification Scheme		1700
Information for Authors		1701

ANNOUNCEMENTS

Call for Papers—IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING Special Issue on Visual Signal Processing for Wireless Networks		1703
Call for Papers—IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING Special Issue on Signal Processing for Situational Awareness From Networked Sensors and Social Media		1704

CALL FOR PAPERS
IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY
Special Issue on Facial Biometrics in the Wild

Guest Editors

<i>Aly A. Farag</i>	University of Louisville (aly.farag@louisville.edu)
<i>Bir Bhanu,</i>	University of California, Riverside (bhanu@cris.ucr.edu)
<i>Edwin Hancock</i>	University of York, UK (erh@cs.york.ac.uk)
<i>Gerard Medioni</i>	University of Southern California (medioni@usc.edu)
<i>Jie Yang</i>	Shanghai Jiaotong University, China (jieyang@sjtu.edu.cn)

Face recognition is an established field of machine learning and has progressed for over three decades in terms of theory, algorithms and applications. While surveillance systems are in common practice and close-range facial recognition at entry points exist, the field has progressed beyond still image recognition in controlled imaging environments. Facial recognition in the wild connotes general face recognition from still and video imaging, unabated by age, position, illumination and expression (A-PIE) of individuals. Facial recognition in that sense of generality crosses boundaries of applications from security to education and various man-machine interfaces. This special issue will address the latest developments in age models, illumination models, pose-invariant recognition, and the role of expression understanding in face recognition. Both still and video-based imaging are considered, and partial information from either modality is of interest. Hybrid methods (still + video) are of particular interest, and fusion methodologies that handle uncertain information, occlusion and cross-age variations are strongly encouraged. This special issue will focus on the theoretical foundation of facial information modeling and analysis, novel algorithms and systems, and multidisciplinary perspectives and applications of facial biometrics.

These topics, and related subject matters, will be categorized into four broad categories:

- *Facial Information Modeling*: Illumination; Facial Features; Pose Invariance; Age Modeling; Expression Modeling & Synthesis
- *Man-Machine Interface for Facial Biometrics*: Emotion Models; Affective Models; State of mind; detection and analysis; Facial biometric-driven robotics and smart systems; Facial readings for special needs
- *Performance Evaluation of Facial Biometrics Methodologies*: Performance measures of facial detection; Performance measures of facial features for recognition; Overall performance measures; A-PIE overall performance measures; Performance measures with open environment facial biometrics.
- *Facial Biometrics Systems*: Single camera systems (single and multiple object); Video-based facial tracking; Face recognition at a distance; Facial biometrics smart systems and tele-presence; Mobile facial biometric systems; Person re-identification; Super-resolution and low resolution facial biometric systems.

Submission Procedure: Manuscripts are to be submitted according to the Information for Authors at <http://www.signalprocessingsociety.org/publications/periodicals/forensics/forensics-authors-info/> using the IEEE online manuscript system, Manuscript Central. Papers must not have appeared or be under review elsewhere. Manuscripts by the guest editors submitted to this SI will be handled by the EIC of IEEE-TIFS.

Schedule:

Submission deadline: February 15, 2014
First Review: June 1, 2014
Revisions Due: July 15, 2014
Final Decision: September 1, 2014
Final manuscript due: September 30, 2014
Tentative publication date: December 1, 2014



11th IEEE International Conference on Advanced Video and Signal-Based Surveillance

August 26-29, 2014 HanaSquare, Korea University, Seoul, Korea

www.avss2014.org



General Chair:
Hanseok Ko, Korea Univ.

General Co-Chair:
Jin Young Choi, SNU

Program Chair:
Daejin Kim, POSTECH
Peter Tu, GE Global Research

Workshops Chair:
Inso Kweon, KAIST

Industrial Chair:
Sungho Cho, Hanyang Univ.

Local Arrangements Chair:
Sangmin Yoon, Kookmin Univ.

Publicity Chairs:
Moon Gi Kang, Yonsei Univ.

Publications Chair:
Dongil Kim, Samsung-Techwin

Finance/Registration Chair:
Wooil Kim, Incheon Univ.

Web Chair:
Jihyun Kim, Korea Univ.

AVSS is the premier annual international conference in the field of video and signal-based surveillance that brings together experts from academia, industry, and government to advance theories, methods, systems, and applications related to surveillance. AVSS is sponsored by the IEEE and, in particular, by its two societies, the Signal Processing Society (IVMSP TC) and the IEEE Computer Society (PAMI TC).

AVSS will celebrate its 11th anniversary in Seoul in 2014. It has been steadily growing in both stature and attendance, from about 70 attendees in 1998 (Genova, Italy) to 125 in 2010 (Boston, USA), 140 in 2011 (Klagenfurt, Austria), and about 200 in 2013 (Krakow, Poland). A strong attendance is expected in Seoul in 2014. AVSS focuses on underlying theory, methods, systems, and applications of surveillance and invites submissions in areas listed below, especially cross-disciplinary and game-changing ones. The list of topics of interest includes, but is not limited to:

Sensor-Centric Processing

- Sensors (visible/infrared/3D/mm-wave/ audio/radio, etc.)
- Ground, airborne, satellite based (fixed/ mobile/UAV)
- Crowdsourcing (cellular, social networks)
- Calibration and positioning (GPS, etc.)
- Communications and networked sensing

Data Management & Human-Computer Interaction

- Compression and summarization
- Archival, search and retrieval
- Human-computer interfaces
- Visualization algorithms
- Mobile and distributed interaction

Security and Privacy

- Data authenticity
- Privacy in surveillance
- Forensics
- Biometrics (standoff, multi-modal, voice, etc.)
- Cybersecurity for surveillance (wireless, network, computer)

Processing, Detection, Tracking & Recognition

- Modeling and feature selection
- Detection and estimation (change, motion, anomaly, saliency, pattern)
- Data association and (multi) target tracking
- Classification and recognition
- Multi-modal fusion

Analytics, Situation Awareness & Decision Making

- Activity/interaction analysis and monitoring
- Intention
- estimation and situation awareness
- Crowdsourcing-based methods
- Cognitive dynamic systems and bio-inspired methods

Surveillance Systems and Applications

- Hardware and software architectures
- Research prototypes
- Simulators
- Civilian, industrial, and military
- Transportation (road, rail, air, maritime)
- Performance evaluation

Paper Submission

Original contributions in any of the above (or related) areas are being solicited. Prospective authors are invited to submit full-length papers, up to 6 pages long, including results, figures, tables and references by the due date (see below). Detailed submission instructions will be posted on Paper Submission page in due time. Each submission will be double-blind peer-reviewed by at least two experts. The conference proceedings will be published in IEEE Xplore digital library.

Call for Workshops

AVSS-2014 will host workshops on August 26, 2014 prior to the technical program of the conference. Prospective organizers are invited to submit workshop proposals by **February 14, 2014**. Detailed instructions regarding workshop proposal preparation and submission can be found on the [Call for Workshops](#) page.

Important dates:

Workshop proposals:	February 14, 2014
Paper submission:	March 1, 2014
Paper acceptance:	April 18, 2014
Camera ready:	May 2, 2014
Early registration:	June 1, 2014



IEEE TRANSACTIONS ON *MULTIMEDIA*

A PUBLICATION OF
THE IEEE CIRCUITS AND SYSTEMS SOCIETY
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY



<http://www.signalprocessingsociety.org/tmm/>

OCTOBER 2013

VOLUME 15

NUMBER 6

ITMUF8

(ISSN 1520-9210)

SPECIAL SECTION ON SOCIAL MEDIA AS SENSORS

GUEST EDITORIAL

Special Section on Social Media as Sensors <http://dx.doi.org/10.1109/TMM.2013.2264232> *I. Kompatsiaris, D. Gatica-Perez, X. Xie, and J. Luo* 1229

SPECIAL SECTION PAPERS

Learning Crowdsourced User Preferences for Visual Summarization of Image Collections <http://dx.doi.org/10.1109/TMM.2013.2261481> ... *S. Rudinac, M. Larson, and A. Hanjalic* 1231

Tracking Large-Scale Video Remix in Real-World Events <http://dx.doi.org/10.1109/TMM.2013.2264929> *L. Xie, A. Natsev, X. He, J. R. Kender, M. Hill, and J. R. Smith* 1244

Towards Cross-Domain Learning for Social Video Popularity Prediction <http://dx.doi.org/10.1109/TMM.2013.2265079> *S. D. Roy, T. Mei, W. Zeng, and S. Li* 1255

Sensing Trending Topics in Twitter <http://dx.doi.org/10.1109/TMM.2013.2265080> *L. M. Aiello, G. Petkos, C. Martin, D. Corney, S. Papadopoulos, R. Skraba, A. Göker, I. Kompatsiaris, and A. Jaimes* 1268

Travel Recommendation by Mining People Attributes and Travel Group Types From Community-Contributed Photos <http://dx.doi.org/10.1109/TMM.2013.2265077> *Y.-Y. Chen, A.-J. Cheng, and W. H. Hsu* 1283

Inferring Contexts From Facebook Interactions: A Social Publicity Scenario <http://dx.doi.org/10.1109/TMM.2013.2265168> *S. Servia-Rodríguez, A. Fernández-Vilas, R. P. Díaz-Redondo, and J. J. Pazos-Arias* 1296

Quantitative Study of Music Listening Behavior in a Social and Affective Context <http://dx.doi.org/10.1109/TMM.2013.2265078> *Y.-H. Yang and J.-Y. Liu* 1304

Connectivity, Online Social Capital, and Mood: A Bayesian Nonparametric Analysis <http://dx.doi.org/10.1109/TMM.2013.2264274> *D. Phung, S. K. Gupta, T. Nguyen, and S. Venkatesh* 1316



REGULAR ISSUE PAPERS

Audio/Video Analysis and Synthesis

Group Delay Based Methods for Speaker Segregation and its Application in Multimedia Information Retrieval
<http://dx.doi.org/10.1109/TMM.2013.2247391> K. Nathwani, P. Pandit, and R. M. Hegde 1326

Compression and Coding

Kinect-Like Depth Data Compression <http://dx.doi.org/10.1109/TMM.2013.2247584> J. Fu, D. Miao, W. Yu, S. Wang, Y. Lu, and S. Li 1340

Watermarking, Encryption, and Data Hiding

A Collusion-Resistant Conditional Access System for Flexible-Pay-Per-Channel Pay-TV Broadcasting
<http://dx.doi.org/10.1109/TMM.2013.2250493> Z. Wan, J. Liu, R. Zhang, and R. H. Deng 1353

Multimodal Human–Machine Interfaces and Interaction

Localization of Taps on Solid Surfaces for Human-Computer Touch Interfaces <http://dx.doi.org/10.1109/TMM.2013.2264656>
 V. G. Reju, A. W. H. Khong, and A. B. Sulaiman 1365

Indexing, Searching, Retrieving, Query, and Archiving Databases

Joint Spatio-Temporal Alignment of Sequences <http://dx.doi.org/10.1109/TMM.2013.2247390> F. Diego, J. Serrat, and A. M. López 1377

Automatic Training Image Acquisition and Effective Feature Selection From Community-Contributed Photos for Facial
 Attribute Detection <http://dx.doi.org/10.1109/TMM.2013.2250492> Y.-Y. Chen, W.H. Hsu, and H.-Y.M. Liao 1388

Content Understanding and Knowledge Modeling

Exploiting Semantic and Visual Context for Effective Video Annotation <http://dx.doi.org/10.1109/TMM.2013.2250266>
 J. Yi, Y. Peng, and J. Xiao 1400

Multimedia Search and Retrieval

MSIDX: Multi-Sort Indexing for Efficient Content-Based Image Search and Retrieval <http://dx.doi.org/10.1109/TMM.2013.2247989>
 E. Tiakas, D. Rafailidis, A. Dimou, and P. Daras 1415

Multimedia Streaming and Transport

QoE-Driven Cache Management for HTTP Adaptive Bit Rate Streaming Over Wireless Networks
<http://dx.doi.org/10.1109/TMM.2013.2247583> W. Zhang, Y. Wen, Z. Chen, and A. Khisti 1431

Modeling and Analysis of Skype Video Calls: Rate Control and Video Quality <http://dx.doi.org/10.1109/TMM.2013.2247988>
 X. Zhang, Y. Xu, H. Hu, Y. Liu, Z. Guo, and Y. Wang 1446

CORRESPONDENCE

Display Technology for Multimedia

Example-Based Super-Resolution With Soft Information and Decision <http://dx.doi.org/10.1109/TMM.2013.2264654>
 Z. Xiong, D. Xu, X. Sun, and F. Wu 1458

EDICS—Editors Classification Scheme 1466
 Information for Authors 1467

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING


www.ieee.org/sp/index.html

OCTOBER 2013

VOLUME 7

NUMBER 5

IJSTGY

(ISSN 1932-4553)

ISSUE ON LEARNING-BASED DECISION MAKING IN DYNAMIC SYSTEMS UNDER UNCERTAINTY

EDITORIAL

Introduction to the Issue on Learning-Based Decision Making in Dynamic Systems Under Uncertainty
<http://dx.doi.org/10.1109/IJSTSP.2013.2263591> *Q. Zhao, E. K. P. Chong, B. Krishnamachari, A. Leshem, S. Meyn, and V. V. Veeravalli* 743

PAPERS

- Feature Search in the Grassmanian in Online Reinforcement Learning <http://dx.doi.org/10.1109/IJSTSP.2013.2255022> *S. Bhatnagar, V. S. Borkar, and P. K. J.* 746
- Deterministic Sequencing of Exploration and Exploitation for Multi-Armed Bandit Problems
<http://dx.doi.org/10.1109/IJSTSP.2013.2263494> *S. Vakili, K. Liu, and Q. Zhao* 759
- Sequentiality and Adaptivity Gains in Active Hypothesis Testing <http://dx.doi.org/10.1109/IJSTSP.2013.2261279> *M. Naghshvar and T. Javidi* 768
- Multistage Adaptive Estimation of Sparse Signals <http://dx.doi.org/10.1109/IJSTSP.2013.2256105> *D. Wei and A. O. Hero* 783
- Hypothesis Testing in Feedforward Networks With Broadcast Failures <http://dx.doi.org/10.1109/IJSTSP.2013.2258657> *Z. Zhang, E. K. P. Chong, A. Pezeshki, and W. Moran* 797
- Learning-Based Constraint Satisfaction With Sensing Restrictions <http://dx.doi.org/10.1109/IJSTSP.2013.2251604> *A. Checco and D. J. Leith* 811
- Distributed Energy-Aware Diffusion Least Mean Squares: Game-Theoretic Learning <http://dx.doi.org/10.1109/IJSTSP.2013.2266318> *O. N. Gharehshiran, V. Krishnamurthy, and G. Yin* 821
- Distributed Learning and Multiaccess of On-Off Channels <http://dx.doi.org/10.1109/IJSTSP.2013.2261278> *S. Chen and L. Tong* 837
- Winning the Lottery: Learning Perfect Coordination With Minimal Feedback <http://dx.doi.org/10.1109/IJSTSP.2013.2259465> *W. Zame, J. Xu, and M. van der Schaar* 846
- Multiagent Reinforcement Learning Based Spectrum Sensing Policies for Cognitive Radio Networks
<http://dx.doi.org/10.1109/IJSTSP.2013.2259797> *J. Lundén, S. R. Kulkarni, V. Koivunen, and H. V. Poor* 858
- Opportunistic Spectrum Access by Exploiting Primary User Feedbacks in Underlay Cognitive Radio Systems: An Optimality Analysis <http://dx.doi.org/10.1109/IJSTSP.2013.2259464> *K. Wang, L. Chen, and Q. Liu* 869
- Maximizing Quality of Information From Multiple Sensor Devices: The Exploration vs Exploitation Tradeoff
<http://dx.doi.org/10.1109/IJSTSP.2013.2259798> *E. N. Ciftcioglu, A. Yener, and M. J. Neely* 883
- Transmit Power Control Policies for Energy Harvesting Sensors With Retransmissions <http://dx.doi.org/10.1109/IJSTSP.2013.2258656> *A. Aprem, C. R. Murthy, and N. B. Mehta* 895
- Robust Reputation Protocol Design for Online Communities: A Stochastic Stability Analysis
<http://dx.doi.org/10.1109/IJSTSP.2013.2263785> *Y. Zhang and M. van der Schaar* 907



Information for Authors	921
<hr/>	
ANNOUNCEMENTS	
J-STSP Call for Special Issue Proposals	923
Call for Papers—Special Issue on Signal Processing in Smart Electric Power Grid	924
Call for Papers—Special Issue on Visual Signal Processing for Wireless Networks	925
Call for Papers—Special Issue on Signal Processing for Situational Awareness from Networked Sensors and Social Media	926
<hr/>	

IEEE Journal on Selected Topics in Signal Processing

Special Issue on Interactive Media Processing for Immersive Communication

Interpersonal communication is an essential and intrinsic element of the human way of life. Yet, for interpersonal communication among individuals connected via high-speed data networks, the prevalent technologies are limited to exchanging 2D video and audio captured by a single camera and microphone. These technologies fail to provide a level of immersivity necessary for “in the same room” sense of presence, due to limitations such as gaze mismatch and lack of depth perception in the rendered scene.

With the recent advances in sensing technologies and accompanying data analysis tools, one can now acquire a large collection of media data describing both the sender’s physical environment (e.g., texture/depth images and audio captured from arrays of cameras and microphones) and the manner in which the sender is perceiving the presented media (e.g., gaze & head movements). This means that although the viewer’s display capabilities may remain limited (e.g., 2D display, stereo speakers), the sense of immersion can be greatly enhanced through innovative human-centric interaction with the presented media (e.g., gazed-corrected view, seamless view-switching and/or spatial audio corresponding to viewer’s tracked head position, haptic vibrations in response to loud audio events). This enhanced media interaction must be designed within the context of delay- and loss-prone networks for delivery of delay-sensitive data, towards the ultimate goal of improving immersive communication beyond 2D video communication. In particular, the main technical challenges are: i) efficient acquisition and processing of observer’s sensory data, ii) compact/robust representation of media data for network transport given the receiver’s current patterns of media consumption and display limitations, iii) real-time human-centric media interaction for an enriched immersive experience. Further, evaluation methodologies and metrics for immersive communication systems must be subjectively accurate for the range of systems deploying multiple sensory input and output devices.

We invite authors to address aspects of immersive communication related to interactive media processing, such as the following. Please note that submission of pure video compression papers with no direct connection to media interaction is not encouraged.

- **Interactive Visual Communication** (e.g., efficient view switching systems enabling motion parallax, 3D visual representation with desirable properties like flexible decoding, robustness, etc.).
- **Acquisition & Reconstruction of Media Data for Immersivity** (e.g., depth data acquisition and pre-processing, microphone array design for spatial audio, hybrid approaches to depth estimation, gaze/head tracking and prediction).
- **Multi-modal Media Interaction** (e.g., multi-modal responses to detected communication events, visual interaction (defocusing, saliency-based content adaptation) based on gaze patterns and/or pupil size).
- **Streaming/Transport of Immersive Media Data** (e.g., media-specific, delay-sensitive FEC, multiple descriptions for multi-path transmission, multi-modal loss concealment strategies).
- **Applications of Immersive Communication** (e.g., systems for tele-medicine/education, immersive video conferencing).
- **Quality Issues in Immersive Communication** (e.g., evaluation methodologies and metrics for immersive systems).

Prospective authors should visit <http://www.signalprocessingsociety.org/publications/periodicals/jstsp/> for information on paper submission. Manuscripts should be submitted at <http://mc.manuscriptcentral.com/jstsp-ieee>.

Manuscript Submission	First Review Due	Revised Manuscript	Second Review Due	Final Manuscript
April 2, 2014	July 1, 2014	September 1, 2014	November 1, 2014	December 1, 2014

Guest Editors:

Gene Cheung, National Institute of Informatics, Japan (cheung@nii.ac.jp)

Dinei Florencio, Microsoft Research, USA (dinei@microsoft.com)

Patrick Le Callet, University of Nantes, France (patrick.lecallet@univ-nantes.fr)

Chia-Wen Lin, National Tsing-Hwa University, Taiwan (cwlin@ee.nthu.edu.tw)

Enrico Magli, Politecnico di Torino, Italy (enrico.magli@polito.it)



Call for Papers
Journal of Selected Topics in Signal Processing

IEEE

Special Issue on
Signal Processing for Situational Awareness from
Networked Sensors and Social Media

Situational awareness is of great interest in security/surveillance, disaster management, environmental monitoring, etc. At the same time, distributed data sources are common today and generate a wealth of information, e.g., surveillance cameras, smartphones, parking occupancy sensors, RFID (proximity) sensors. However, there exists a gap between the wealth of distributed information captured and the understanding of a scene where the sensors are located. The goal of this special issue is an attempt at bridging this gap by bringing together various signal processing methodologies associated with such data analysis. Since situational awareness methods in practice today often combine sensor measurements with information gleaned from social media, this issue will consider both these information sources.

Challenges in information processing from large sensor networks are many - from the use of single sensors adapting internal or external parameters, to the cooperation of multiple sensors in a networked manner. Also, the amount of data that needs to be analyzed can be prohibitive, leading to various “big data” challenges. These challenges are particularly acute for video data, the main surveillance modality today, due to its sheer volume. Recent technological developments have also led to more efficient and robust sensors that more tightly couple the acquisition and analysis phases (e.g., network cameras, smartphones). This requires the development of algorithms, ideally implemented close to the sensor and which are specifically tuned to the characteristics of the sensors. Seeking solutions to the above challenges cannot be done without considering the usage of resources, such as energy consumption, communication bandwidth, computational power, etc. Finally, a very recent data modality, that is likely to play a significant role in the future, comes from social media in the form of text, images, videos, and sound recordings. The fusion of social media data with traditional surveillance streams is an uncharted, but potentially fertile, territory.

This special issue proposal is innovative because it aims to bring together, in one publication, novel research from diverse areas of signal processing with the common goal of exploiting “high volume data” collected daily for more efficient situational awareness. We will encourage submissions describing different approaches, both in terms of the number of sensors (networked vs. single), their types (network cameras, smart phones, radar, social media, etc.), data processing methodologies, and application domains. The fusion of data captured by different sensors or different platforms will be of great interest for the proposed issue as well. In order to clearly differentiate this issue from others in JSTSP, papers submitted **must** have a significant contribution to advancing the state of the art in situational awareness. Work with a variety of sensor modalities is welcome, including novel multi-modal fusion techniques.

A **list of topics** follows, but we will encourage additional topics falling within the purview of situational awareness.

- **Sensors:** Novel sensors, Sensing from mobile platforms (UAV, UGV, satellite, etc.), Beyond visual perception sensing (radar, hyperspectral, etc.)
- **Sensor networks:** Sensor network processing and control, Distributed detection, localization and tracking, Distributed machine learning for scene analysis, Multiagent coordination for wide-area coverage and analysis, Re-identification and network tracking, Network resource management, Information fusion over distributed sensor networks
- **System-level scene understanding:** Multi-modal fusion, Context-based object description and scene understanding, Event annotation, summarization and visualization, Inferring spatio-temporal, causal and contextual relations between events, Multimedia Analytics
- **Exploitation of social media for situational awareness:** Visual data analysis in large online repositories (e.g., YouTube, Instagram), Context analysis in social media (e.g., Twitter, Facebook, etc.), Fusion of sensor and social data

Prospective authors should visit <http://www.signalprocessingsociety.org/publications/periodicals/jstsp/> for information on paper submission. Manuscripts should be submitted using Manuscript Central at <http://mc.manuscriptcentral.com/jstsp-ieee>.

Dates

Paper Submission	First Review	Revision Due	Final Decision	Camera Ready	Expected Publication
1 March 2014	1 June 2014	1 August 2014	1 October 2014	15 November 2014	First half of 2015

Guest Editors

Amit K. Roy-Chowdhury, University of California, Riverside, USA; amitr@ee.ucr.edu

Christian Micheloni, Università degli Studi di Udine, Italy; christian.micheloni@uniud.it

Janusz Konrad, Boston University, USA; jkonrad@bu.edu

Mohan S. Kankanhalli, National University of Singapore, Singapore; mohan@comp.nus.edu.sg

Pramod Varshney, Syracuse University, USA; varshney@syr.edu

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

NOVEMBER 2013

VOLUME 20

NUMBER 11

ISPLEM

(ISSN 1070-9908)

EDITORIAL

 An Extra Page for Citations in the Signal Processing Letters <http://dx.doi.org/10.1109/LSP.2013.2278719> *A. Scaglione* 1005

LETTERS

- Optimal Multiband Pulse Shaping and Improved Power Spectrum Estimation Using Stopbands Information
<http://dx.doi.org/10.1109/LSP.2013.2278622> *A. R. Forouzan and M. F. Sabahi* 1006
- A Nonlocal Poisson Denoising Algorithm Based on Stochastic Distances <http://dx.doi.org/10.1109/LSP.2013.2277111>
 *A. A. Bindilatti and N. D. A. Mascarenhas* 1010
- Opportunistic Scheduling With BIA Under Block Fading Broadcast Channels <http://dx.doi.org/10.1109/LSP.2013.2278710>
 *W. Jiang, Z. He, K. Niu, L. Guo, and W. Wu* 1014
- Noise-Adaptive LDA: A New Approach for Speech Recognition Under Observation Uncertainty
<http://dx.doi.org/10.1109/LSP.2013.2278556> *D. Kolossa, S. Zeiler, R. Saeidi, and R. F. Astudillo* 1018
- Robust Painting Recognition and Registration for Mobile Augmented Reality <http://dx.doi.org/10.1109/LSP.2013.2279014>
 *N. Martinel, C. Micheloni, and G. L. Foresti* 1022
- An Exact Tree Projection Algorithm for Wavelets <http://dx.doi.org/10.1109/LSP.2013.2278147> *C. Cartis and A. Thompson* 1026
- Destination Assisted Secret Wireless Communication With Cooperative Helpers <http://dx.doi.org/10.1109/LSP.2013.2278835>
 *B. Yang, W. Wang, B. Yao, and Q. Yin* 1030
- Weighted Shape-Based Averaging With Neighborhood Prior Model for Multiple Atlas Fusion-Based Medical Image
 Segmentation <http://dx.doi.org/10.1109/LSP.2013.2279269> *S. Gorthi, M. Bach Cuadra, P.-A. Tercier, A. S. Allal, and J.-P. Thiran* 1034
- Block Sparse Estimator for Grid Matching in Single Snapshot DoA Estimation <http://dx.doi.org/10.1109/LSP.2013.2279124>
 *R. Jagannath and K.V.S. Hari* 1038
- A Blind AM/PM Estimation Method for Power Amplifier Linearization <http://dx.doi.org/10.1109/LSP.2013.2280394>
 *Z. Zhu, X. Huang, M. Caron, and H. Leung* 1042

IEEE SIGNAL PROCESSING LETTERS (ISSN 1070-9908) is published quarterly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$266.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage pending at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE SIGNAL PROCESSING LETTERS, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.

On the Gradient Descent Localization of Radioactive Sources http://dx.doi.org/10.1109/LSP.2013.2279499	<i>H. E. Baidoo-Williams, S. Dasgupta, R. Mudumbai, and E. Bai</i>	1046
Revisiting Inter-Genre Similarity http://dx.doi.org/10.1109/LSP.2013.2280031	<i>B. L. Sturm and F. Gouyon</i>	1050
A Direct Algorithm for 1-D Total Variation Denoising http://dx.doi.org/10.1109/LSP.2013.2278339	<i>L. Condat</i>	1054
Background Subtraction With Video Coding http://dx.doi.org/10.1109/LSP.2013.2280138	<i>Z. Huang, R. Hu, and Z. Wang</i>	1058
A New Algorithm for Fitting a Gaussian Function Riding on the Polynomial Background http://dx.doi.org/10.1109/LSP.2013.2280577	<i>E. Kheirati Roonizi</i>	1062
Lossy Depth Image Compression using Greedy Rate-Distortion Slope Optimization http://dx.doi.org/10.1109/LSP.2013.2280495	<i>I. Schioppa and I. Tabus</i>	1066
Iterative Directional Total Variation Refinement for Compressive Sensing Image Reconstruction http://dx.doi.org/10.1109/LSP.2013.2280571	<i>X. Fei, Z. Wei, and L. Xiao</i>	1070
Improving the Bound on the RIP Constant in Generalized Orthogonal Matching Pursuit http://dx.doi.org/10.1109/LSP.2013.2279977	<i>S. Satpathi, R. L. Das, and M. Chakraborty</i>	1074
A Novel Nonlinear Regression Approach for Efficient and Accurate Image Matting http://dx.doi.org/10.1109/LSP.2013.2274874	<i>Q. Zhu, Z. Zhang, Z. Song, Y. Xie, and L. Wang</i>	1078
Third-Order Kalman Filter: Tuning and Steady-State Performance http://dx.doi.org/10.1109/LSP.2013.2277668	<i>H. Shu, E. P. Simon, and L. Ros</i>	1082
Rao and Wald Tests for Distributed Targets Detection With Unknown Signal Steering http://dx.doi.org/10.1109/LSP.2013.2277371	<i>W. Liu, W. Xie, and Y. Wang</i>	1086
A New Modified SLM Scheme for Wireless OFDM Systems Without Side Information http://dx.doi.org/10.1109/LSP.2013.2278286	<i>J. Ji and G. Ren</i>	1090
Kernel Based Multiple Cue Adaptive Appearance Model For Robust Real-time Visual Tracking http://dx.doi.org/10.1109/LSP.2013.2278400	<i>F. Zeng, X. Liu, Z. Huang, and Y. Ji</i>	1094
High Throughput Stochastic Log-MAP Turbo-Decoder Based on Low Bits Computation http://dx.doi.org/10.1109/LSP.2013.2278853	<i>J. Chen and J. Hu</i>	1098
Set-Valued State Estimation and Attack Detection for Uncertain Descriptor Systems http://dx.doi.org/10.1109/LSP.2013.2278281	<i>A. N. Bishop and A. V. Savkin</i>	1102
Deterministic Construction of Real-Valued Ternary Sensing Matrices Using Optical Orthogonal Codes http://dx.doi.org/10.1109/LSP.2013.2281597	<i>N. Y. Yu and N. Zhao</i>	1106
Wideband Gaussian Source Processing Using a Linear Nested Array http://dx.doi.org/10.1109/LSP.2013.2281514	<i>K. Han and A. Nehorai</i>	1110
Dynamic Diffusion Estimation in Exponential Family Models http://dx.doi.org/10.1109/LSP.2013.2282042	<i>K. Dedecius and V. Sečkárová</i>	1114
Optimal Bayesian Resampling for OFDM Signaling Over Multi-scale Multi-lag Channels http://dx.doi.org/10.1109/LSP.2013.2282339	<i>S. Beygi and U. Mitra</i>	1118
A New Variable Tap-Length and Step-Size FxLMS Algorithm http://dx.doi.org/10.1109/LSP.2013.2282396	<i>D.-C. Chang and F.-T. Chu</i>	1122
Bayesian Early Mode Decision Technique for View Synthesis Prediction-Enhanced Multiview Video Coding http://dx.doi.org/10.1109/LSP.2013.2281607	<i>S. Khattak, R. Hamzaoui, T. Maugey, S. Ahmad, and P. Frossard</i>	1126
On the Fisher Information Matrix for Multivariate Elliptically Contoured Distributions http://dx.doi.org/10.1109/LSP.2013.2281914	<i>O. Besson and Y. I. Abramovich</i>	1130
QoS Feasibility in MIMO Broadcast Channels With Widely Linear Transceivers http://dx.doi.org/10.1109/LSP.2013.2282186	<i>C. Hellings, M. Joham, and W. Utschick</i>	1134
The Use of Spatially Random Base Stations in Cloud Radio Access Networks http://dx.doi.org/10.1109/LSP.2013.2282157	<i>Z. Ding and H. V. Poor</i>	1138

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

www.ieee.org/sp/index.html

NOVEMBER 2013

VOLUME 20

NUMBER 11

ISPLEM

(ISSN 1070-9908)

EDITORIAL

An Extra Page for Citations in the Signal Processing Letters <http://dx.doi.org/10.1109/LSP.2013.2278719> A. Scaglione 1005

LETTERS

*Biomedical Signal Processing*Weighted Shape-Based Averaging With Neighborhood Prior Model for Multiple Atlas Fusion-Based Medical Image Segmentation <http://dx.doi.org/10.1109/LSP.2013.2279269> S. Gorthi, M. Bach Cuadra, P.-A. Tercier, A. S. Allal, and J.-P. Thiran 1034*Design and Implementation of Signal Processing Systems*High Throughput Stochastic Log-MAP Turbo-Decoder Based on Low Bits Computation <http://dx.doi.org/10.1109/LSP.2013.2278853> J. Chen and J. Hu 1098*Digital and Multirate Signal Processing*An Exact Tree Projection Algorithm for Wavelets <http://dx.doi.org/10.1109/LSP.2013.2278147> C. Cartis and A. Thompson 1026A Direct Algorithm for 1-D Total Variation Denoising <http://dx.doi.org/10.1109/LSP.2013.2278339> L. Condat 1054Rao and Wald Tests for Distributed Targets Detection With Unknown Signal Steering <http://dx.doi.org/10.1109/LSP.2013.2277371> W. Liu, W. Xie, and Y. Wang 1086*Image and Multidimensional Signal Processing*A Nonlocal Poisson Denoising Algorithm Based on Stochastic Distances <http://dx.doi.org/10.1109/LSP.2013.2277111> A. A. Bindilatti and N. D. A. Mascarenhas 1010Robust Painting Recognition and Registration for Mobile Augmented Reality <http://dx.doi.org/10.1109/LSP.2013.2279014> N. Martinel, C. Micheloni, and G. L. Foresti 1022Lossy Depth Image Compression using Greedy Rate-Distortion Slope Optimization <http://dx.doi.org/10.1109/LSP.2013.2280495> I. Schioppa and I. Tabus 1066A Novel Nonlinear Regression Approach for Efficient and Accurate Image Matting <http://dx.doi.org/10.1109/LSP.2013.2274874> Q. Zhu, Z. Zhang, Z. Song, Y. Xie, and L. Wang 1078Kernel Based Multiple Cue Adaptive Appearance Model For Robust Real-time Visual Tracking <http://dx.doi.org/10.1109/LSP.2013.2278400> F. Zeng, X. Liu, Z. Huang, and Y. Ji 1094Bayesian Early Mode Decision Technique for View Synthesis Prediction-Enhanced Multiview Video Coding <http://dx.doi.org/10.1109/LSP.2013.2281607> S. Khattak, R. Hamzaoui, T. Maugey, S. Ahmad, and P. Frossard 1126IEEE SIGNAL PROCESSING LETTERS (ISSN 1070-9908) is published quarterly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997.**IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$266.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage pending at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE SIGNAL PROCESSING LETTERS, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.

Machine Learning and Statistical Signal Processing

Revisiting Inter-Genre Similarity http://dx.doi.org/10.1109/LSP.2013.2280031	<i>B. L. Sturm and F. Gouyon</i>	1050
Background Subtraction With Video Coding http://dx.doi.org/10.1109/LSP.2013.2280138	<i>Z. Huang, R. Hu, and Z. Wang</i>	1058
Iterative Directional Total Variation Refinement for Compressive Sensing Image Reconstruction http://dx.doi.org/10.1109/LSP.2013.2280571	<i>X. Fei, Z. Wei, and L. Xiao</i>	1070
Improving the Bound on the RIP Constant in Generalized Orthogonal Matching Pursuit http://dx.doi.org/10.1109/LSP.2013.2279977	<i>S. Satpathi, R. L. Das, and M. Chakraborty</i>	1074

Multimedia Signal Processing

A New Variable Tap-Length and Step-Size FxLMS Algorithm http://dx.doi.org/10.1109/LSP.2013.2282396	<i>D.-C. Chang and F.-T. Chu</i>	1122
---	----------------------------------	------

Sensor Array and Multichannel Signal Processing

Block Sparse Estimator for Grid Matching in Single Snapshot DoA Estimation http://dx.doi.org/10.1109/LSP.2013.2279124	<i>R. Jagannath and K.V.S. Hari</i>	1038
Deterministic Construction of Real-Valued Ternary Sensing Matrices Using Optical Orthogonal Codes http://dx.doi.org/10.1109/LSP.2013.2281597	<i>N. Y. Yu and N. Zhao</i>	1106
Wideband Gaussian Source Processing Using a Linear Nested Array http://dx.doi.org/10.1109/LSP.2013.2281514 ...	<i>K. Han and A. Nehorai</i>	1110

Signal Processing for Communications

Opportunistic Scheduling With BIA Under Block Fading Broadcast Channels http://dx.doi.org/10.1109/LSP.2013.2278710	<i>W. Jiang, Z. He, K. Niu, L. Guo, and W. Wu</i>	1014
Destination Assisted Secret Wireless Communication With Cooperative Helpers http://dx.doi.org/10.1109/LSP.2013.2278835	<i>B. Yang, W. Wang, B. Yao, and Q. Yin</i>	1030
Third-Order Kalman Filter: Tuning and Steady-State Performance http://dx.doi.org/10.1109/LSP.2013.2277668	<i>H. Shu, E. P. Simon, and L. Ros</i>	1082
Optimal Bayesian Resampling for OFDM Signaling Over Multi-scale Multi-lag Channels http://dx.doi.org/10.1109/LSP.2013.2282339	<i>S. Beygi and U. Mitra</i>	1118
QoS Feasibility in MIMO Broadcast Channels With Widely Linear Transceivers http://dx.doi.org/10.1109/LSP.2013.2282186	<i>C. Hellings, M. Joham, and W. Utschick</i>	1134
The Use of Spatially Random Base Stations in Cloud Radio Access Networks http://dx.doi.org/10.1109/LSP.2013.2282157	<i>Z. Ding and H. V. Poor</i>	1138

Speech Processing

Noise-Adaptive LDA: A New Approach for Speech Recognition Under Observation Uncertainty http://dx.doi.org/10.1109/LSP.2013.2278556	<i>D. Kolossa, S. Zeiler, R. Saeidi, and R. F. Astudillo</i>	1018
---	--	------

Statistical and Adaptive Signal Processing

Optimal Multiband Pulse Shaping and Improved Power Spectrum Estimation Using Stopbands Information http://dx.doi.org/10.1109/LSP.2013.2278622	<i>A. R. Forouzan and M. F. Sabahi</i>	1006
A Blind AM/PM Estimation Method for Power Amplifier Linearization http://dx.doi.org/10.1109/LSP.2013.2280394	<i>Z. Zhu, X. Huang, M. Caron, and H. Leung</i>	1042
On the Gradient Descent Localization of Radioactive Sources http://dx.doi.org/10.1109/LSP.2013.2279499	<i>H. E. Baidoo-Williams, S. Dasgupta, R. Mudumbai, and E. Bai</i>	1046
A New Algorithm for Fitting a Gaussian Function Riding on the Polynomial Background http://dx.doi.org/10.1109/LSP.2013.2280577	<i>E. Kheirati Roonizi</i>	1062
A New Modified SLM Scheme for Wireless OFDM Systems Without Side Information http://dx.doi.org/10.1109/LSP.2013.2278286	<i>J. Ji and G. Ren</i>	1090
Set-Valued State Estimation and Attack Detection for Uncertain Descriptor Systems http://dx.doi.org/10.1109/LSP.2013.2278281	<i>A. N. Bishop and A. V. Savkin</i>	1102
Dynamic Diffusion Estimation in Exponential Family Models http://dx.doi.org/10.1109/LSP.2013.2282042	<i>K. Dedecius and V. Sečkarová</i>	1114
On the Fisher Information Matrix for Multivariate Elliptically Contoured Distributions http://dx.doi.org/10.1109/LSP.2013.2281914	<i>O. Besson and Y. I. Abramovich</i>	1130

IEEE SignalProcessing

MAGAZINE

[VOLUME 31 NUMBER 1 JANUARY 2014]

SIGNAL AND IMAGE PROCESSING IN HYPERSPECTRAL REMOTE SENSING

ADVANCES, CHALLENGES,
AND OPPORTUNITIES

IMAGE INPAINTING

SECURING DIGITAL
REPUTATION

THE SIGNAL PROCESSING
DISCIPLINE: PART 2

IEEE
Signal Processing Society

IEEE

CONTENTS

[VOLUME 31 NUMBER 1]

SPECIAL SECTION—SIGNAL AND IMAGE PROCESSING IN HYPERSPECTRAL REMOTE SENSING

22 FROM THE GUEST EDITORS

Wing-Kin Ma, José M. Bioucas-Dias, Jocelyn Chanussot, and Paul Gader

24 DETECTION ALGORITHMS IN HYPERSPECTRAL IMAGING SYSTEMS

Dimitris Manolakis, Eric Truslow, Michael Pieper, Thomas Cooley, and Michael Brueggeman

34 HYPERSPECTRAL TARGET DETECTION

Nasser M. Nasrabadi

45 ADVANCES IN HYPERSPECTRAL IMAGE CLASSIFICATION

Gustavo Camps-Valls, Devis Tuia, Lorenzo Bruzzone, and Jón Atli Benediktsson

55 MANIFOLD-LEARNING-BASED FEATURE EXTRACTION FOR CLASSIFICATION OF HYPERSPECTRAL DATA

Dalton Lunga, Saurabh Prasad, Melba M. Crawford, and Okan Ersoy

67 A SIGNAL PROCESSING PERSPECTIVE ON HYPERSPECTRAL UNMIXING

Wing-Kin Ma, José M. Bioucas-Dias, Tsung-Han Chan, Nicolas Gillis, Paul Gader, Antonio J. Plaza, ArulMurugan Ambikapathi, and Chong-Yung Chi

82 NONLINEAR UNMIXING OF HYPERSPECTRAL IMAGES

Nicolas Dobigeon, Jean-Yves Tourneret, Cédric Richard, José Carlos M. Bermudez, Stephen McLaughlin, and Alfred O. Hero

95 ENDMEMBER VARIABILITY IN HYPERSPECTRAL ANALYSIS

Alina Zare and K.C. Ho

105 COMPRESSIVE CODED APERTURE SPECTRAL IMAGING

Gonzalo R. Arce, David J. Brady, Lawrence Carin, Henry Arguello, and David S. Kittle

116 SPARSITY AND STRUCTURE IN HYPERSPECTRAL IMAGING

Rebecca M. Willett, Marco F. Duarte, Mark A. Davenport, and Richard G. Baraniuk

FEATURE

127 IMAGE INPAINTING

Christine Guillemot and Olivier Le Meur

COLUMNS

4 FROM THE EDITOR

Risky Long-Term Research: Two Success Stories in Audio Signal Processing
Abdelhak Zoubir

6 PRESIDENT'S MESSAGE

A Lasting Journey
K.J. Ray Liu

8 READER'S CHOICE

Top Downloads in IEEE *Xplore*

15 SPECIAL REPORTS

Fresh Approaches Promise Wireless Quality and Reliability Improvements
John Edwards

Automotive Industry Is a Key Component to the Success of the DSP Sector
Ron Schneiderman

145 STANDARDS IN A NUTSHELL

Royalty-Free Video Coding Standards in MPEG
Kiho Choi and Euee S. Jang

149 APPLICATIONS CORNER

Securing Digital Reputation in Online Social Media
Yuhong Liu and Yan (Lindsay) Sun

157 REFLECTIONS

The Discipline of Signal Processing: Part 2
Mos Kaveh and Li Deng

DEPARTMENTS

10 SOCIETY NEWS

New Members-at-Large, Directors-at-Large, and Class of Distinguished Lecturers for 2014

156 DATES AHEAD

ChinaSIP
001011 国际会议

2nd IEEE China Summit & International Conference on Signal and Information Processing

IEEE Signal Processing Society

Xi'an International Conference Center

9-13 July 2014 Xi'an, CHINA

Organizing Committee

Honorary Chair:

Jianguo HUANG, IEEE Xi'an Section

General Chairs:

Mingyi HE, NW Polytech Univ, China
Kung YAO, UCLA, USA

Technical Program Committee Chairs:

Jingdong CHEN, NW Polytech Univ, China
Xuelong LI, Chinese Academy Science
Z Jane WANG, UBC, Canada

Finance Chair:

Yan FENG, NW Polytech Univ, China

Industry Forum Chair:

Li DENG, Microsoft Research, USA

Special/Invited Session Chairs:

Hongwei LIU, Xidian Univ, China
Hong ZHAO, Univ Alberta, Canada

Education & Research Panel Chairs:

Changshui ZHANG, Tsinghua Univ, China
Guisheng LIAO, Xidian Univ, China

Publicity Chairs:

Xuanqin MOU, Xi'an Jiaotong Univ, China
Yanning ZHANG, NW Polytech Univ, China

Publication Chairs:

Yanguy FAN, NW Polytech Univ, China
Guizhong LIU, Xi'an Jiaotong Univ, China

Summer School & Tutorial Chairs:

Kehu YANG, Xidian Univ, China
Lei XIE, NW Polytech Univ, China

Local Arrangement Chair:

Bo LI, NW Polytech Univ, China

Demo & Exhibition Chair:

Ling WANG, NW Polytech Univ, China

Registration Chairs:

Chengbing HE, NW Polytech Univ, China
Shaohui MEI, NW Polytech Univ, China

International Liaisons:

David ANDERSON, Georgia Tech, USA
Chengchew LIM, Adelaide Univ, Australia
Shoji MAKINO, Tsukuba Univ, Japan
Changyun WEN, NTU, Singapore

Steering Committee

Chair:

Min WU, Univ Maryland, USA

Members-at-Large:

Yun HE, Tsinghua Univ, China
Jiwu HUANG, Sun-Yat-Sen Univ, China
Xingzhao LIU, Shanghai Jiaotong U, China

SPS VP Conference:

Wan-Chi SIU

SPS VP Technical Direction:

Charles Bouman

SPS Region-10 Director:

Ta-Sung LEE

IEEE China Operation Director:

Ning HUA

ChinaSIP'14

The 2nd IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP'14) will be held in Xi'an, China from 9th to 13th, July 2014. ChinaSIP is a new annual summit and international conference held in China for domestic and international scientists, researchers, and developers to report and discuss their latest progress in all theoretical and technological aspects of signal and information processing. ChinaSIP is a unique platform developed by IEEE signal processing society (SPS) to help colleagues in China engage with the global community, by building a broad technical community with domestic and international colleagues, providing a ladder for China's signal and information processing professionals to further advance on the global stage. Meanwhile, it offers global colleagues opportunities to network and develop international collaborations. The conference will feature world-class speakers, tutorials, exhibits, and over decade lectures and poster sessions. The regular technical program tracks and topics include (but not limited to):

Tracks

- Speech, Audio, and Language
- Image/Video Coding, Transmission and Processing
- Multimedia SIP and Emerging Applications
- Pattern Recognition for SIP
- Machine Learning and Big Data Analysis for SIP
- SIP for Bioinformatics & Bio/Medicine
- Information Forensics & Security (Including Biometrics)
- SIP for Communications and Networking
- SIP Theory, Methods, and Applications
- Remote Sensing Data Processing
- Multichannel and Array Signal Processing

Features

Keynote and invited talks by global experts
Demos and showcase of recent journal articles
Summer school and hands-on tutorials
Education & research panel

Education & research panel
Industry forum and exhibits
Professional development and networking
with global leaders

The Proceedings of ChinaSIP'14 will be included in the IEEE Xplore database and Indexed by EI Compendex.

Xi'an The ancient capital city of 13 dynasties known as Chang'an and with more than 3,100 years of history, records the great changes of the Chinese nation just like a living history book. Specialists and scholars regard it as an inexhaustible treasure house of literature while heads of state from many countries and people from all over the world turn their eyes to this tourist attraction, trying to broaden their knowledge of Chinese civilization.

Paper Submission The official language of the conference is English. Prospective authors are invited to submit up to 4 pages in length (with an optional 5th page containing only references). Any accepted paper included in the final program is expected to have at least one author or qualified proxy attend and present the paper at the conference. More information can be found at the conference Web or obtained via conference email.

Web www.chinasip2014.org

Email chinasip2014@163.com

Supporters A variety of support opportunities for ChinaSIP are outlined at conference web. By supporting ChinaSIP, your organization will be highly visible and demonstrate its stature in the exciting field of SIP to a broad audience of researchers, educators, and practitioners from industry, government, and academia both within and beyond China. The participation and support from your organization will significantly enhance the quality of the event, and in turn benefit your organization through raising broad awareness of your technical contributions (products or R&D work) and attracting talents from domestic and global SIP communities.

Important Dates:

Deadline for Full Paper submission	10 Feb 2014
Notification of Acceptance	15 April 2014
Deadline for Camera Ready Paper Submission	10 May 2014
Deadline for Authors' Registration	15 May 2014



2014 IEEE International Conference on Image Processing

CNIT La Défense - Paris, France
Oct 27-30, 2014

web site : www.icip2014.com

CALL FOR PAPERS



The International Conference on Image Processing (ICIP), sponsored by the IEEE Signal Processing Society, is the premier forum for the presentation of technological advances and research results in the fields of theoretical, experimental, and applied image and video processing. ICIP 2014, the twenty-first in the series that has been held annually since 1994, brings together leading engineers and scientists in image and video processing from around the world. Research frontiers in fields ranging from traditional image-processing applications to evolving multimedia and video technologies are regularly advanced by results first reported in ICIP technical sessions. Topics include, but are not limited to:

- **Image/video coding and transmission:** Still-image and video coding, stereoscopic and 3-D coding, distributed source coding, source/channel coding, image/video transmission over wireless networks
- **Image/video processing:** Image and video filtering, restoration and enhancement, image segmentation, video segmentation and tracking, morphological processing, stereoscopic and 3-D processing, feature extraction and analysis, interpolation and super-resolution, motion detection and estimation, color and multispectral processing, biometrics
- **Image formation:** Biomedical imaging, remote sensing, geophysical and seismic imaging, optimal imaging, synthetic-natural hybrid image systems
- **Image scanning, display, and printing:** Scanning and sampling, quantization and halftoning, color reproduction, image representation and rendering, display and printing systems, image-quality assessment
- **Image/video storage, retrieval, and authentication:** Image and video databases, image and video search and retrieval, multimodality image/video indexing and retrieval, authentication and watermarking
- **Applications:** Biomedical sciences, mobile imaging, geosciences and remote sensing, astronomy and space exploration, document image processing and analysis, other applications.

Paper Submission: Prospective authors are invited to submit papers of not more than four (4) pages for technical content including figures and possible references, and with one additional optional 5th page containing only references. Papers will be accepted by electronic submission only at: www.icip2014.com.

Submission deadline	January 31, 2014
Notification of acceptance	April 30, 2014
Camera-ready papers	May 15, 2014
Author registration	May 25, 2014

Tutorials: Tutorials will be held on October, 27, 2014. Brief proposals should be submitted by November 29, 2013 at the conference web site. Proposals for tutorials must include a title, an outline of the tutorial and its motivation, a short description of the material to be covered, contact information including name, affiliation, email, and mailing address for each presenter, and a two-page CV for each presenter.

Please see: www.icip2014.com for details regarding Paper Submission, "no-show" policy, special sessions, and tutorials.

Organizing Committee

General Chairs

Béatrice Pesquet-Popescu
and James Fowler

Vice-General Chair

Frédéric Dufaux

Technical Program Chairs

Pascal Frossard
and Marc Antonini

Treasurer

Michel Kieffer

Publicity Chair

Jean-Luc Dugelay

Special Session Chairs

Adriana Dumitras
and Ricardo de Queiroz

Tutorial Chairs

Mihaela van der Schaar
and Josiane Zerubia

Plenary Chair

Christine Guillemot

Awards Chairs

Hamid Krim
and Marco Cagnazzo

Publication Chair

Jocelyn Chanussot

Registration Chair

Yannick Berthoumie

IEEE Student Activities Chairs

Nelly Pustelnik
and Mihai Ciuc

Electronic Media Chair

Cyril Concolato

Exhibition Chairs

Jean LeFeuvre
and Constantin Vertan

Local Arrangements Chairs

Chloé Clavel, Maria Trocan
and Bogdan Ionescu

Industry Liaisons

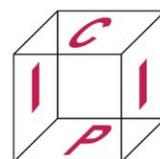
Jacques Blanc-Talon,
Frédéric Barbaresco
and Aljoscha Smolic

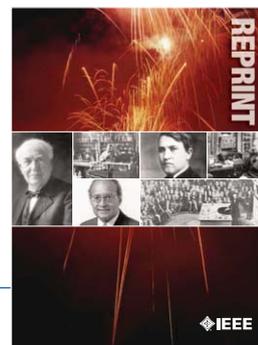
US Liaison

Sheila Hemami

Asia Liaisons

Shipeng Li
and Oscar Au





IEEE ORDER FORM FOR REPRINTS

Purchasing IEEE Papers in Print is easy, cost-effective and quick.

Complete this form, tear it out, and either fax it (24 hours a day) to 732-981-8062 or mail it back to us.

PLEASE FILL OUT THE FOLLOWING

Author: _____

Publication Title: _____

Paper Title: _____

RETURN THIS FORM TO:
 IEEE Publishing Services
 445 Hoes Lane
 Box 1331
 Piscataway, NJ 08855-1331
Call Reprint Department at (732) 562-3941
for questions regarding this form
(732) 981-8062 - FAX

PLEASE SEND ME

50 100 200 300 400 500 or _____ (in multiples of 50) reprints.

YES NO Self-covering/title page required. COVER PRICE: \$74 per 100, \$39 per 50.

\$58.00 Air Freight must be added for all orders being shipped outside the U.S.

\$18.50 must be added for all USA shipments to cover the cost of UPS shipping and handling.

PAYMENT

Check enclosed. Payable on a bank in the USA.

Charge my: Visa Mastercard Amex Diners Club

Account # _____ Exp. date _____

Cardholder's Name (please print): _____

Bill me (you must attach a purchase order) Purchase Order Number _____

Send Reprints to: _____ Bill to address, if different: _____

Because information and papers are gathered from various sources, there may be a delay in receiving your reprint request. This is especially true with postconference publications. Please provide us with contact information if you would like notification of a delay of more than 12 weeks.

Telephone: _____ Fax: _____ Email Address: _____

2011 REPRINT PRICES (without covers)

Number of Text Pages

	1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44	45-48
50	\$126	\$211	\$243	\$245	\$285	\$340	\$371	\$408	\$440	\$477	\$510	\$543
100	\$242	\$423	\$476	\$492	\$570	\$680	\$742	\$817	\$885	\$953	\$1021	\$1088

Larger quantities can be ordered. Email reprints@ieee.org with specific details.

Tax Applies on shipments of regular reprints to CA, DC, FL, MI, NJ, NY, OH and Canada (GST Registration no. 12534188).

Prices are based on black & white printing. Please call us for full color price quote, if applicable.

Authorized Signature: _____ Date: _____



2014 IEEE MEMBERSHIP APPLICATION

(students and graduate students must apply online)



Start your membership immediately: Join online www.ieee.org/join

Please complete both sides of this form, typing or **printing in capital letters**. Use only English characters and abbreviate only if more than 40 characters and spaces per line. We regret that incomplete applications cannot be processed.

1 Name & Contact Information

Please PRINT your name as you want it to appear on your membership card and IEEE correspondence. As a key identifier for the IEEE database, circle your last/surname.

Male Female Date of birth (Day/Month/Year) ____/____/____

Title First/Given Name Middle Last/Family Surname

▼ Primary Address Home Business (All IEEE mail sent here)

Street Address

City State/Province

Postal Code Country

Primary Phone

Primary E-mail

▼ Secondary Address Home Business

Company Name Department/Division

Street Address City State/Province

Postal Code Country

Secondary Phone

Secondary E-mail

To better serve our members and supplement member dues, your postal mailing address is made available to carefully selected organizations to provide you with information on technical services, continuing education, and conferences. Your e-mail address is **not** rented by IEEE. Please check box **only** if you do not want to receive these postal mailings to the selected address.

2 Attestation

I have graduated from a three- to five-year academic program with a university-level degree. Yes No

This program is in one of the following fields of study:

- Engineering
- Computer Sciences and Information Technologies
- Physical Sciences
- Biological and Medical Sciences
- Mathematics
- Technical Communications, Education, Management, Law and Policy
- Other (please specify): _____

This academic institution or program is accredited in the country where the institution is located. Yes No Do not know

I have _____ years of professional experience in teaching, creating, developing, practicing, or managing within the following field:

- Engineering
- Computer Sciences and Information Technologies
- Physical Sciences
- Biological and Medical Sciences
- Mathematics
- Technical Communications, Education, Management, Law and Policy
- Other (please specify): _____

3 Please Tell Us About Yourself

Select the numbered option that best describes yourself. This information is used by IEEE magazines to verify their annual circulation. Please enter numbered selections in the boxes provided.

A. Primary line of business

1. Computers
2. Computer peripheral equipment
3. Software
4. Office and business machines
5. Test, measurement and instrumentation equipment
6. Communications systems and equipment
7. Navigation and guidance systems and equipment
8. Consumer electronics/appliances
9. Industrial equipment, controls and systems
10. ICs and microprocessors
11. Semiconductors, components, sub-assemblies, materials and supplies
12. Aircraft, missiles, space and ground support equipment
13. Oceanography and support equipment
14. Medical electronic equipment
15. OEM incorporating electronics in their end product (not elsewhere classified)
16. Independent and university research, test and design laboratories and consultants (not connected with a mfg. co.)
17. Government agencies and armed forces
18. Companies using and/or incorporating any electronic products in their manufacturing, processing, research or development activities
19. Telecommunications services, telephone (including cellular)
20. Broadcast services (TV, cable, radio)
21. Transportation services (airline, railroad, etc.)
22. Computer and communications and data processing services
23. Power production, generation, transmission and distribution
24. Other commercial users of electrical, electronic equipment and services (not elsewhere classified)
25. Distributor (reseller, wholesaler, retailer)
26. University, college/other educational institutions, libraries
27. Retired
28. Other _____

B. Principal job function

- | | |
|--|---|
| 1. General and corporate management | 9. Design/development engineering—digital |
| 2. Engineering management | 10. Hardware engineering |
| 3. Project engineering management | 11. Software design/development management |
| 4. Research and development management | 12. Computer science |
| 5. Design engineering management—analogue | 13. Science/physics/mathematics (not elsewhere specified) |
| 6. Design engineering management—digital | 14. Engineering (not elsewhere specified) |
| 7. Research and development engineering | 15. Marketing/sales/purchasing |
| 8. Design/development engineering—analogue | 16. Consulting |
| | 17. Education/teaching |
| | 18. Retired |
| | 19. Other _____ |

C. Principal responsibility

- | | |
|--|-----------------------|
| 1. Engineering and scientific management | 6. Education/teaching |
| 2. Management other than engineering | 7. Consulting |
| 3. Engineering design | 8. Retired |
| 4. Engineering | 9. Other _____ |
| 5. Software: science/mngmnt/engineering | |

D. Title

- | | |
|--|--------------------------------|
| 1. Chairman of the Board/President/CEO | 10. Design Engineering Manager |
| 2. Owner/Partner | 11. Design Engineer |
| 3. General Manager | 12. Hardware Engineer |
| 4. VP Operations | 13. Software Engineer |
| 5. VP Engineering/Dir. Engineering | 14. Computer Scientist |
| 6. Chief Engineer/Chief Scientist | 15. Dean/Professor/Instructor |
| 7. Engineering Management | 16. Consultant |
| 8. Scientific Management | 17. Retired |
| 9. Member of Technical Staff | 18. Other _____ |

Are you now or were you ever a member of IEEE?
 Yes No If yes, provide, if known:

Membership Number	Grade	Year Expired
-------------------	-------	--------------

4 Please Sign Your Application

I hereby apply for IEEE membership and agree to be governed by the IEEE Constitution, Bylaws, and Code of Ethics. I understand that IEEE will communicate with me regarding my individual membership and all related benefits. **Application must be signed.**

Signature _____ Date _____ Over Please

Information for Authors

(Updated March 2012)

The IEEE TRANSACTIONS are published monthly covering advances in the theory and application of signal processing. The scope is reflected in the EDICS: the Editor's Information and Classification Scheme. Please consider the journal with the most appropriate scope for your submission.

Authors are encouraged to submit manuscripts of Regular papers (papers which provide a complete disclosure of a technical premise), or Correspondences (brief items that describe a use for or magnify the meaning of a single technical point, or provide comment on a paper previously published in the TRANSACTIONS). Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere.

Every manuscript must (a) provide a clearly defined statement of the problem being addressed, (b) state why it is important to solve the problem, and (c) give an indication as to how the current solution fits into the history of the problem.

By submission/resubmission of your manuscript to this TRANSACTIONS, you are acknowledging that you accept the rules established for publication of manuscripts, including agreement to pay all overlength page charges, color charges, and any other charges and fees associated with publication of the manuscript. Such charges are not negotiable and cannot be suspended.

New and revised manuscripts should be prepared following the "New Manuscript Submission" guidelines below, and submitted to the online manuscript system ScholarOne Manuscripts. After acceptance, finalized manuscripts should be prepared following the "Final Manuscript Submission Guidelines" below. Do not send original submissions or revisions directly to the Editor-in-Chief or Associate Editors; they will access your manuscript electronically via the ScholarOne Manuscripts system.

New Manuscript Submission. Please follow the next steps.

1. *Account in ScholarOne Manuscripts.* If necessary, create an account in the on-line submission system ScholarOne Manuscripts. Please check first if you already have an existing account which is based on your e-mail address and may have been created for you when you reviewed or authored a previous paper.
2. *Electronic Manuscript.* Prepare a PDF file containing your manuscript in double-spaced format (one full blank line between lines of type) using a font size of 11 points or larger, having a margin of at least 1 inch on all sides. For a regular paper, the manuscript may not exceed 30 double-spaced pages, including title; names of authors and their complete contact information; abstract; text; all images, figures and tables; and all references. Upload your manuscript as a PDF file "manuscript.pdf" to the ScholarOne Manuscripts site, then proofread your submission, confirming that all figures and equations are visible in your document before you "SUBMIT" your manuscript. Proofreading is critical; once you submit your manuscript, the manuscript cannot be changed in any way. You may also submit your manuscript as a PostScript or MS Word file. The system has the capability of converting your files to PDF, however it is your responsibility to confirm that the conversion is correct and there are no font or graphics issues prior to completing the submission process.
3. *Double-Column Version of Manuscript.* You are required to also submit a roughly formatted version of the manuscript in single-spaced, double column IEEE format (10 points for a regular submission or 9 points for a Correspondence) using the IEEE style files (it is allowed to let long equations stick out). *If accepted for publication*, over length page charges are levied beginning with the 11th published page of your manuscript. You are, therefore, advised to be conservative in your submission. This double-column version submitted will serve as a confirmation of the approximate publication length of the manuscript and gives an additional confirmation of your understanding that over length page charges will be paid when billed upon publication. Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOneManuscripts site.
4. *Additional Material for Review.* Please upload pdf versions of all items in the reference list which are not publicly available, such as unpublished (submitted) papers. Other materials for review such as supplementary tables and figures, audio fragments and QuickTime movies may be uploaded as well. Reviewers will be able to view these files only if they have the appropriate software on their computers. Use short filenames without spaces or special characters. When the upload of each file is completed, you will be asked to provide a description of that file.
5. *Submission.* After uploading all files and proofreading them, submit your manuscript by clicking "Submit." A confirmation of the successful submission will open on screen containing the manuscript tracking number and will be followed with an e-mail confirmation to the corresponding and all contributing authors. Once you click "Submit," your manuscript cannot be changed in any way.

6. *Copyright Form and Consent Form.* By policy, IEEE owns the copyright to the technical contributions it publishes on behalf of the interests of the IEEE, its authors, and their employers; and to facilitate the appropriate reuse of this material by others. To comply with the IEEE copyright policies, authors are required to sign and submit a completed "IEEE Copyright and Consent Form" prior to publication by the IEEE.

The IEEE recommends authors to use an effective electronic copyright form (eCF) tool within the ScholarOne Manuscripts system. You will be redirected to the "IEEE Electronic Copyright Form" wizard at the end of your original submission; please simply sign the eCF by typing your name at the proper location and click on the "Submit" button.

Correspondence Items. Correspondence items are short disclosures with a reduced scope or significance that typically describe a use for or magnify the meaning of a single technical point, or provide brief comments on material previously published in the TRANSACTIONS. These items may not exceed 12 pages in double-spaced format (3 pages for Comments), using 11 point type, with margins of 1 inch minimum on all sides, and including: title, names and contact information for authors, abstract, text, references, and an appropriate number of illustrations and/or tables. Correspondence items are submitted in the same way as regular manuscripts (see "New Manuscript Submission" above for instructions).

Manuscript Length. Papers published on or after 1 January 2007 can now be up to 10 pages, and any paper in excess of 10 pages will be subject to over length page charges. The IEEE Signal Processing Society has determined that the standard manuscript length shall be no more than 10 published pages (double-column format, 10 point type) for a regular submission, or 6 published pages (9 point type) for a Correspondence item, respectively. Manuscripts that exceed these limits will incur mandatory over length page charges, as discussed below. Since changes recommended as a result of peer review may require additions to the manuscript, it is strongly recommended that you practice economy in preparing original submissions.

Exceptions to the 30-page (regular paper) or 12-page (Correspondences) manuscript length may, under extraordinary circumstances, be granted by the Editor-in-Chief. However, such exception does not obviate your requirement to pay any and all over length or additional charges that attach to the manuscript.

Resubmission of Previously Rejected Manuscripts. Authors of rejected manuscripts are allowed to resubmit their manuscripts only once. The Signal Processing Society strongly discourages resubmission of rejected manuscripts more than once. At the time of submission, you will be asked whether you consider your manuscript to be a new submission or a resubmission of an earlier rejected manuscript. If you choose to submit a new version of your manuscript, you will be asked to submit supporting documents detailing how your new version addresses all of the reviewers' comments.

Full details of the resubmission process can be found in the Signal Processing Society "Policy and Procedures Manual" at <http://www.signalprocessingsociety.org/about/governance/policy-procedure/>. Also, please refer to the decision letter and your Author Center on the on-line submission system.

Author Misconduct.

Author Misconduct Policy: Plagiarism includes copying someone else's work without appropriate credit, using someone else's work without clear delineation of citation, and the uncited reuse of an authors previously published work that also involves other authors. Plagiarism is unacceptable.

Self-plagiarism involves the verbatim copying or reuse of an authors own prior work without appropriate citation; it is also unacceptable. Self-plagiarism includes duplicate submission of a single journal manuscript to two different journals, and submission of two different journal manuscripts which overlap substantially in language or technical contribution.

Authors may only submit original work that has not appeared elsewhere in a journal publication, nor is under review for another journal publication. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper. If authors have used their own previously published work as a basis for a new submission, they are required to cite the previous work and very briefly indicate how the new submission offers substantively novel contributions beyond those of the previously published work.

It is acceptable for conference papers to be used as the basis for a more fully developed journal submission. Still, authors are required to cite related prior work; the papers cannot be identical; and the journal publication must include novel aspects.

Author Misconduct Procedures: The procedures that will be used by the Signal Processing Society in the investigation of author misconduct allegations are described in the IEEE SPS Policies and Procedures Manual.

Author Misconduct Sanctions: The IEEE Signal Processing Society will apply the following sanctions in any case of plagiarism, or in cases of self-plagiarism that involve an overlap of more than 25% with another journal manuscript:

- 1) immediate rejection of the manuscript in question;
- 2) immediate withdrawal of all other submitted manuscripts by any of the authors, submitted to any of the Society's publications (journals, conferences, workshops), except for manuscripts that also involve innocent co-authors; immediate withdrawal of all other submitted manuscripts by any of the authors, submitted to any of the Society's publications (journals, conferences, workshops), except for manuscripts that also involve innocent co-authors;
- 3) prohibition against each of the authors for any new submissions, either individually, in combination with the authors of the plagiarizing manuscript, or in combination with new co-authors, to all of the Society's publications (journals, conferences, workshops). The prohibition shall continue for one year from notice of suspension.

Further, plagiarism and self-plagiarism may also be actionable by the IEEE under the rules of Member Conduct.

Submission Format.

Authors are encouraged to prepare manuscripts employing the on-line style files developed by IEEE. All manuscripts accepted for publication will require the authors to make final submission employing these style files. The style files are available on the web at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2 under "Template for all Transactions." (LaTeX and MS Word).

Authors using LaTeX: the two PDF versions of the manuscript needed for submission can both be produced by the IEEEtran.cls style file. A double-spaced document is generated by including `\documentclass[11pt,draftcls,onecolumn]{IEEEtran}` as the first line of the manuscript source file, and a single-spaced double-column document for estimating the publication page charges via `\documentclass[10pt,twocolumn,twoside]{IEEEtran}` for a regular submission, or `\documentclass[9pt,twocolumn,twoside]{IEEEtran}` for a Correspondence item.

- **Title page and abstract:** The first page of the manuscript shall contain the title, names and contact information for all authors (full mailing address, institutional affiliations, phone, fax, and e-mail), the abstract, and the EDICS. An asterisk * should be placed next to the name of the Corresponding Author who will serve as the main point of contact for the manuscript during the review and publication processes.

An abstract should have not more than 200 words for a regular paper, or 50 words for a Correspondence item. The abstract should indicate the scope of the paper or Correspondence, and summarize the author's conclusions. This will make the abstract, by itself, a useful tool for information retrieval.

- **EDICS:** All submissions must be classified by the author with an EDICS (Editors' Information Classification Scheme) selected from the list of EDICS published online at <http://www.signalprocessingsociety.org/publications/periodicals/tsp/TSP-EDICS/>
- NOTE: EDICS are necessary to begin the peer review process. Upon submission of a new manuscript, please choose the EDICS categories that best suit your manuscript. Failure to do so will likely result in a delay of the peer review process.
- The EDICS category should appear on the first page—i.e., the title and abstract page—of the manuscript.
- **Illustrations and tables:** Each figure and table should have a caption that is intelligible without requiring reference to the text. Illustrations/tables may be worked into the text of a newly-submitted manuscript, or placed at the end of the manuscript. (However, for the final submission, illustrations/tables must be submitted separately and not interwoven with the text.)

Illustrations in color may be used but, unless the final publishing will be in color, the author is responsible that the corresponding grayscale figure is understandable.

In preparing your illustrations, note that in the printing process, most illustrations are reduced to single-column width to conserve space. This may result in as much as a 4:1 reduction from the original. Therefore, make sure that all words are in a type size that will reduce to a minimum of 9 points or 3/16 inch high in the printed version. Only the major grid lines on graphs should be indicated.

- **Abbreviations:** This TRANSACTIONS follows the practices of the IEEE on units and abbreviations, as outlined in the Institute's published standards. See http://www.ieee.org/portal/cms_docs_iportals/iportals/publications/authors/transjnl/auinfo07.pdf for details.
- **Mathematics:** All mathematical expressions must be legible. Do not give derivations that are easily found in the literature; merely cite the reference.

Final Manuscript Submission Guidelines.

Upon formal acceptance of a manuscript for publication, instructions for providing the final materials required for publication will be sent to the Corresponding Author. Finalized manuscripts should be prepared in LaTeX or MS Word, and are required to use the style files established by IEEE, available at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2.

Instructions for preparing files for electronic submission are as follows:

- Files must be self-contained; that is, there can be no pointers to your system setup.
- Include a header to identify the name of the TRANSACTIONS, the name of the author, and the software used to format the manuscript.
- Do not import graphics files into the text file of your finalized manuscript (although this is acceptable for your initial submission). If submitting on disk, use a separate disk for graphics files.
- Do not create special macros.
- Do not send PostScript files of the text.
- File names should be lower case.
- Graphics files should be separate from the text, and not contain the caption text, but include callouts like "(a)," "(b)."
- Graphics file names should be lower case and named fig1.eps, fig2.tif, etc.
- Supported graphics types are EPS, PS, TIFF, or graphics created using Word, Powerpoint, Excel or PDF. Not acceptable is GIF, JPEG, WMF, PNG, BMP or any other format (JPEG is accepted for author photographs only). The provided resolution needs to be at least 600 dpi (400 dpi for color).
- Please indicate explicitly if certain illustrations should be printed in color; note that this will be at the expense of the author. Without other indications, color graphics will appear in color in the online version, but will be converted to grayscale in the print version.

Additional instructions for preparing, verifying the quality, and submitting graphics are available via http://www.ieee.org/publications_standards/publications/authors/authors_journals.html.

Multimedia Materials.

IEEE Xplore can publish multimedia files and Matlab code along with your paper. Alternatively, you can provide the links to such files in a README file that appears on Xplore along with your paper. For details, please see http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect6 under "Multimedia." To make your work reproducible by others, the TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper.

Page Charges.

Voluntary Page Charges. Upon acceptance of a manuscript for publication, the author(s) or his/her/their company or institution will be asked to pay a charge of \$110 per page to cover part of the cost of publication of the first ten pages that comprise the standard length (six pages, in the case of Correspondences).

Mandatory Page Charges. The author(s) or his/her/their company or institution will be billed \$220 per each page in excess of the first ten published pages for regular papers and six published pages for correspondence items. These are mandatory page charges and the author(s) will be held responsible for them. They are not negotiable or voluntary. The author(s) signifies his willingness to pay these charges simply by submitting his/her/their manuscript to the TRANSACTIONS. The Publisher holds the right to withhold publication under any circumstance, as well as publication of the current or future submissions of authors who have outstanding mandatory page charge debt.

Color Charges. Color figures which appear in color only in the electronic (Xplore) version can be used free of charge. In this case, the figure will be printed in the hardcopy version in grayscale, and the author is responsible that the corresponding grayscale figure is intelligible. Color reproduction in print is expensive, and all charges for color are the responsibility of the author. The estimated costs are as follows. There will be a charge of \$62.50 for each figure; this charge may be subject to change without notification. In addition, there are printing preparation charges which may be estimated as follows: color reproductions on four or fewer pages of the manuscript: a total of approximately \$1045; color reproductions on five pages through eight pages: a total of approximately \$2090; color reproductions on nine through 12 pages: a total of approximately \$3135, and so on. Payment of fees on color reproduction is not negotiable or voluntary, and the author's agreement to publish the manuscript in the TRANSACTIONS is considered acceptance of this requirement.

To find the Information for Authors for IEEE Signal Processing Letters, the IEEE Journal of Selected Topics in Signal Processing or the IEEE Signal Processing Magazine, please refer to the IEEE Signal Processing website at www.signalprocessingsociety.org.



THE EIGHTH IEEE SENSOR ARRAY AND MULTICHANNEL SIGNAL PROCESSING WORKSHOP

www.gtec.udc.es/sam2014

June 22-25, 2014, A Coruña, Spain



CALL FOR PAPERS

TECHNICAL PROGRAM

The SAM Workshop is a major IEEE Signal Processing Society event devoted to sensor array and multichannel signal processing. The organizing committee invites the international community to present and discuss state-of-the-art developments in the field. SAM 2014 will feature plenary talks by leading researchers in the field as well as poster sessions with presentations by the participants.

Welcome to A Coruña! – The workshop will be organized in Hotel Hesperia Finisterre, located in the heart of A Coruña, Spain, a modern city looking out onto the Atlantic Ocean. A Coruña is an open city that embraces all its visitors with warmth and according to its motto “A Coruña is the city where nobody is a stranger”. It is blessed with a beach promenade that completely surrounds it, the Hércules Tower, which is a World Heritage site, and many local and regional attractions. The workshop will take place during the Midsummer Festival that will make SAM 2014 an unforgettable experience.

RESEARCH AREAS

Authors are invited to submit contributions in the following areas:

- Adaptive beamforming
- Array processing for biomedical applications
- Array processing for communications
- Blind source separation and channel identification
- Computational and optimization techniques
- Compressive sensing and sparsity-based signal processing
- Detection and estimation
- Direction-of-arrival estimation
- Intelligent systems and knowledge-based signal processing
- Microphone and loudspeaker array applications
- MIMO radar
- MIMO systems and space-time coding
- Multi-channel imaging
- Multi-sensor processing for smart grid and energy
- Non-Gaussian, nonlinear, and non-stationary models
- Performance evaluations with experimental data
- Radar and sonar array processing
- Sensor networks
- Synthetic aperture techniques
- Space-time adaptive processing
- Statistical modeling for sensor arrays
- Waveform diverse sensors and systems

Submission of papers – Full-length four-page papers will be accepted only electronically at www.gtec.udc.es/sam2014.

Special poster session proposals – They should be submitted by e-mail to the Technical Program Chairs and include a topical title, rationale, session outline, contact information, and list of invited speakers.

IMPORTANT DATES

Special Session Proposals
November 29th, 2013

Submission of Papers
January 10th, 2014

Notification of Acceptance
February 28th, 2014

Final Manuscript Submission
March 28th, 2014

Advance Registration
April 18th, 2014

GENERAL CO-CHAIRS

Mónica F. Bugallo, *Stony Brook University, USA*
monica.bugallo@stonybrook.edu
Luis Castedo, *Universidade da Coruña, Spain*
luis@udc.es

TECHNICAL CO-CHAIRS

Aleksandar Dogandžić, *Iowa State University, USA*
ald@iastate.edu
Marius Pesavento, *University of Darmstadt, Germany*
pesavento@nt.tu-darmstadt.de

FINANCE CHAIR

Olivier Besson, *ISAE, France*
olivier.besson@isae.fr

PUBLICATIONS CHAIR

Maria Sabrina Greco, *University of Pisa, Italy*
m.greco@iet.unipi.it

LOCAL ORGANIZATION CHAIR

Roberto López Valcarce, *Universidade de Vigo, Spain*
valcarce@gts.uvigo.es

TECHNICAL PROGRAM COMMITTEE

Rick Blum
Lehigh University, USA

Volkan Cevher
EPFL, Switzerland

Chong-Yung Chi
National Tsing Hua University, Taiwan

Antonio De Maio
Università Studi Napoli "Federico II", Italy

Jean Pierre Delmas
Telecom SudParis, France

Kutluyil Dogancay
University of South Australia, Australia

Mounir Ghogho
University of Leeds, UK

Martin Haardt
Technische Univ. Ilmenau, Germany

Dominic K. C. Ho
Univ. of Missouri - Columbia, USA

Yingbo Hua
University of California, Riverside, USA

Andreas Jakobsson
Lund University, Sweden

Michael Joham
Technische Univ. München, Germany

Shawn Kraut
MIT Lincoln Laboratory, USA

Miguel A. Lagunas
CTTC-UPC, Barcelona, Spain

Pascal Larzabal
Ecole Normale Supér. de Cachan, France

Geert Leus
Delft Univ. of Technology, Netherlands

Marco Lops
University of Cassino, Italy

Christoph Mecklenbräuker
Technische Universität Wien, Austria

Xavier Mestre
CTTC, Barcelona, Spain

Arye Nehorai
Washington University in St. Louis, USA

Daniel Palomar
Hong Kong Univ. of Science & Technology

Athina Petropulu
Rutgers University, USA

Ali Sayed
University of California, Los Angeles, USA

Philip Schniter
The Ohio State University, USA

Erchin Serpedin
Texas A&M University, USA

Shahram Shahbazpanahi
Univ. Ontario Inst. of Technology, Canada

Ananthram Swami
Army Research Laboratory, USA

A. Lee Swindlehurst
University of California at Irvine, USA

Joseph Tabrikian
Ben-Gurion University of the Negev, Israel

Sergiy Vorobyov
University of Alberta, Canada

Kainam Thomas Wong
The Hong Kong Polytechnic University

Ami Wiesel
The Hebrew Univ. of Jerusalem, Israel

Peter Willet
University of Connecticut, USA

Abdelhak M. Zoubir
Technische Univ. Darmstadt, Germany



