

IEEE Signal Processing MAGAZINE

[VOLUME 32 NUMBER 4 JULY 2015]

ART INVESTIGATION

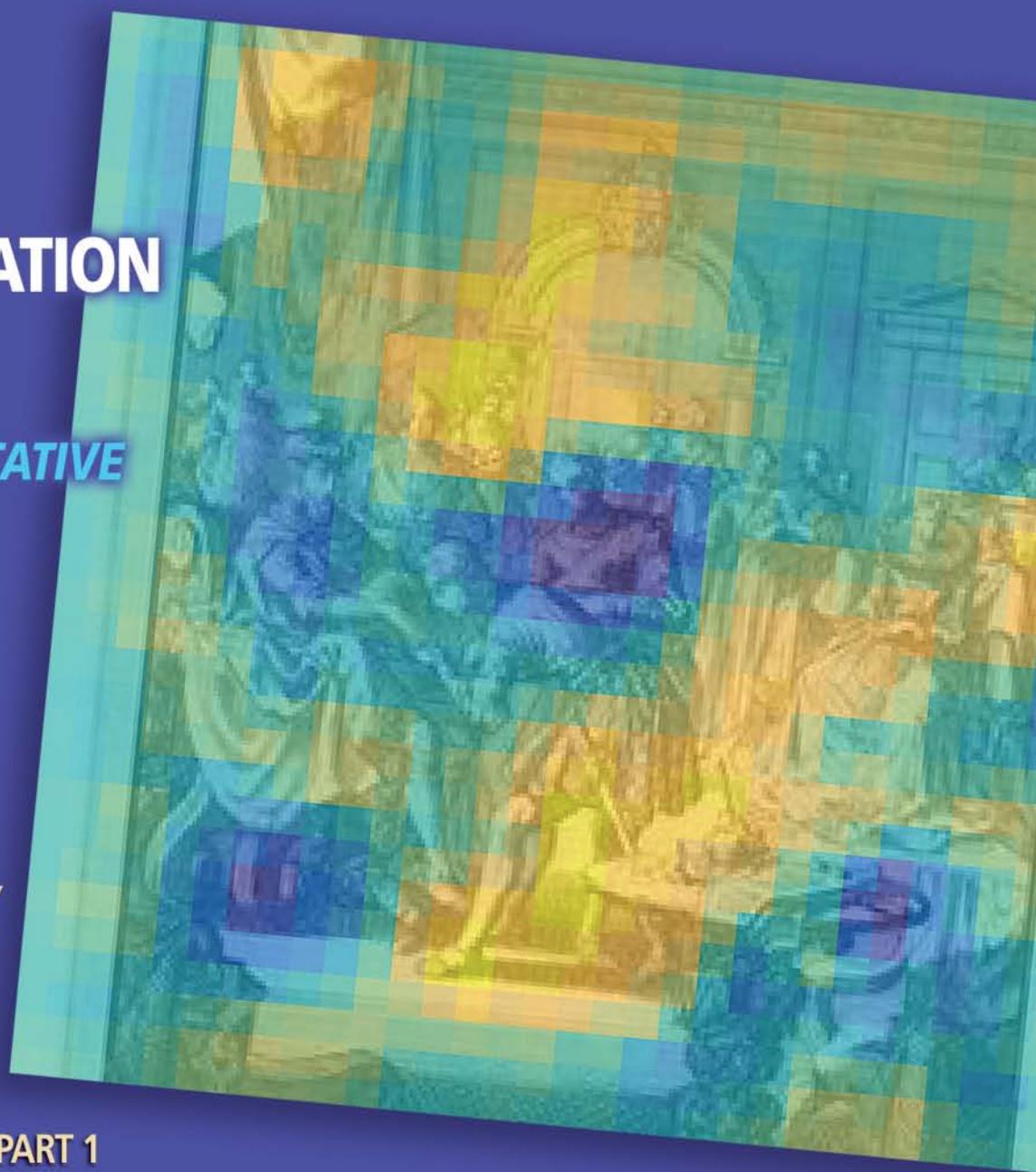
BRIDGING
ART HISTORY
AND QUANTITATIVE
SCIENCES

IMAGE ANALYSIS:
IN PLANT
PHENOTYPING

LECTURE NOTES
ON GAME THEORY
APPLICATIONS

AUTOMATION
IN ANESTHESIA

THE IEEE SIGNAL
PROCESSING CUP: PART 1





USB & Ethernet Programmable ATTENUATORS

New Models up to 120 dB!

0-30, 60, 90, 110 & 120 dB 0.25 dB Step 1 MHz to 6 GHz* from **\$395**

Mini-Circuits' new programmable attenuators offer precise attenuation from 0 up to 120 dB, supporting even more applications and greater sensitivity level measurements! Now available in models with maximum attenuation of 30, 60, 90, 110, and 120 dB with 0.25 dB attenuation steps, they provide the widest range of level control in the industry with accurate, repeatable performance for a variety of applications including fading simulators, handover system evaluation, automated test equipment and more! Our unique designs maintain linear attenuation change per dB over

the entire range of attenuation settings, while USB, Ethernet and RS232 control options allow setup flexibility and easy remote test management. Supplied with user-friendly GUI control software, DLLs for programmers† and everything you need for immediate use right out of the box, Mini-Circuits programmable attenuators offer a wide range of solutions to meet your needs and fit your budget. Visit minicircuits.com for detailed performance specs, great prices, and off the shelf availability. Place your order today for delivery as soon as tomorrow!

RoHS compliant

Models	Attenuation Range	Attenuation Accuracy	Step Size	USB Control	Ethernet Control	RS232 Control	Price Qty. 1-9
RUDAT-6000-30	0-30 dB	±0.4 dB	0.25 dB	✓	-	✓	\$395
RCDAT-6000-30	0-30 dB	±0.4 dB	0.25 dB	✓	✓	-	\$495
RUDAT-6000-60	0-60 dB	±0.3 dB	0.25 dB	✓	-	✓	\$625
RCDAT-6000-60	0-60 dB	±0.3 dB	0.25 dB	✓	✓	-	\$725
RUDAT-6000-90	0-90 dB	±0.4 dB	0.25 dB	✓	-	✓	\$695
RCDAT-6000-90	0-90 dB	±0.4 dB	0.25 dB	✓	✓	-	\$795
NEW RUDAT-6000-110	0-110 dB	±0.45 dB	0.25 dB	✓	-	✓	\$895
NEW RCDAT-6000-110	0-110 dB	±0.45 dB	0.25 dB	✓	✓	-	\$995
NEW RUDAT-4000-120	0-120 dB	±0.5 dB	0.25 dB	✓	-	✓	\$895
NEW RCDAT-4000-120	0-120 dB	±0.5 dB	0.25 dB	✓	✓	-	\$995

*120 dB models specified from 1-4000 MHz.

†No drivers required. DLL objects provided for 32/64-bit Windows® and Linux® environments using ActiveX® and .NET® frameworks.



www.minicircuits.com P.O. Box 350166, Brooklyn, NY 11235-0003 (718) 934-4500 sales@minicircuits.com

523 Rev D

CONTENTS

[VOLUME 32 NUMBER 4]

SPECIAL SECTION—SIGNAL PROCESSING FOR ART INVESTIGATION

14 FROM THE GUEST EDITORS

Patrice Abry, Andrew G. Klein,
William A. Sethares, and
C. Richard Johnson, Jr.

18 MULTISCALE ANISOTROPIC TEXTURE ANALYSIS AND CLASSIFICATION OF PHOTOGRAPHIC PRINTS

Patrice Abry, Stéphane G. Roux,
Herwig Wendt, Paul Messier,
Andrew G. Klein, Nicolas Tremblay,
Pierre Borgnat, Stéphane Jaffard,
Béatrice Vedel, Jim Coddington,
and Lee Ann Daffner

28 HUNTING FOR PAPER MOLDMATES AMONG REMBRANDT'S PRINTS

C. Richard Johnson, Jr., William A.
Sethares, Margaret Holben Ellis,
and Saira Haqqi

38 AUTOMATIC THREAD-LEVEL CANVAS ANALYSIS

Laurens van der Maaten
and Robert Erdmann

46 TOWARD DISCOVERY OF THE ARTIST'S STYLE

Nanne van Noord, Ella Hendriks,
and Eric Postma

55 QUANTITATIVE CANVAS WEAVE ANALYSIS USING 2-D SYNCHROSQUEEZED TRANSFORMS

Haizhao Yang, Jianfeng Lu,
William P. Brown, Ingrid Daubechies,
and Lexing Ying

64 ANCIENT COIN CLASSIFICATION USING REVERSE MOTIF RECOGNITION

Hafeez Anwar, Sebastian Zambanini,
Martin Kampel, and Klaus Vondrovec

75 MULTIMEDIA ANALYSIS AND ACCESS OF ANCIENT MAYA EPIGRAPHY

Rui Hu, Gulcan Can, Carlos Pallán
Gayol, Guido Krempel, Jakub Spotak,
Gabrielle Vail, Stephane
Marchand-Maillet, Jean-Marc Odobez,
and Daniel Gatica-Perez

85 COMPUTERIZED FACE RECOGNITION IN RENAISSANCE PORTRAIT ART

Ramya Srinivasan, Conrad Rudolph,
and Amit Roy-Chowdhury

95 CHALLENGES IN CONTENT-BASED IMAGE INDEXING OF CULTURAL HERITAGE COLLECTIONS

David Picard, Philippe-Henri Gosselin,
and Marie-Claude Gaspard

103 PDE-BASED GRAPH SIGNAL PROCESSING FOR 3-D COLOR POINT CLOUDS

François Lozes, Abderrahim Elmoataz,
and Olivier Lézoray

112 DIGITAL IMAGE PROCESSING OF THE GHENT ALTARPIECE

Aleksandra Pižurica, Ljiljana Platiša,
Tijana Ružić, Bruno Cornelis,
Ann Dooms, Maximiliaan Martens,
Hélène Dubois, Bart Devolder,
Marc De Mey, and Ingrid Daubechies

COLUMNS

4 FROM THE EDITOR

Art, Engineering,
and Community
Min Wu

6 PRESIDENT'S MESSAGE

SigPort: A Paper Repository for
Signal Processing
Alex Acero

8 SPECIAL REPORTS

Human–Machine Interfaces:
Methods of Control
John Edwards

123 SP EDUCATION

Undergraduate Students
Compete in the IEEE Signal
Processing Cup: Part 1
Kin-Man Lam, Carlos Óscar S.
Sorzano, Zhilin Zhang,
and Patrizio Campisi

126 APPLICATIONS CORNER

Image Analysis: The New
Bottleneck in Plant Phenotyping
Massimo Minervini, Hanno Scharf,
and Sotirios A. Tsaftaris

132 LECTURE NOTES

Understanding Game Theory via
Wireless Power Control
Giacomo Bacci, Luca Sanguinetti,
and Marco Luise

138 LIFE SCIENCES

Signal Processing and Automation
in Anesthesia
Guy A. Dumont

145 SP TIPS & TRICKS

Fast, Accurate, and Guaranteed
Stable Sliding Discrete
Fourier Transform
Chun-Su Park

150 STANDARDS IN A NUTSHELL

The High Efficiency Image File
Format Standard
Miska M. Hannuksela, Jani Lainema,
and Vinod K. Malamal Vadakital

DEPARTMENTS

12 SOCIETY NEWS

New Society Officers for 2016 and
Election of Directors-at-Large and
Members-at-Large

160 DATES AHEAD

Digital Object Identifier 10.1109/MSP.2015.2419312

IEEE SIGNAL PROCESSING magazine

EDITOR-IN-CHIEF

Min Wu—University of Maryland, College Park
United States

AREA EDITORS**Feature Articles**

Shuguang Robert Cui—Texas A&M University,
United States

Special Issues

Wade Trappe—Rutgers University, United States

Columns and Forum

Gwenaël Doërr—Technicolor Inc.,
France

Kenneth Lam—Hong Kong Polytechnic
University, Hong Kong SAR of China

e-Newsletter

Christian Debes—TU Darmstadt and
AGT International, Germany

Social Media and Outreach

Andres Kwasinski—Rochester Institute of
Technology, United States

EDITORIAL BOARD

A. Enis Cetin—Bilkent University, Turkey

Patrick Flandrin—ENS Lyon, France

Mounir Ghogho—University of Leeds,
United Kingdom

Lina Karam—Arizona State University,
United States

Bastiaan Kleijn—Victoria University
of Wellington, New Zealand and Delft
University, The Netherlands

Hamid Krim—North Carolina State University,
United States

Ying-Chang Liang—Institute for Infocomm
Research, Singapore

Sven Lončarić—University of Zagreb, Croatia

Brian Lovell—University of Queensland, Australia

Henrique (Rico) Malvar—Microsoft Research,
United States

Stephen McLaughlin—Heriot-Watt University,
Scotland

Athina Petropulu—Rutgers University, United
States

Peter Ramadge—Princeton University, United
States

Shigeki Sagayama—Meiji University, Japan

Eli Saber—Rochester Institute of Technology,
United States

Erchin Serpedin—Texas A&M University,
United States

Shihab Shamma—University of Maryland, United
States

Hing Cheung So—City University of Hong Kong,
Hong Kong

Isabel Trancoso—INESC-ID/Instituto Superior
Técnico, Portugal

Michail K. Tsatsanis—Entropic Communications

Pramod K. Varshney—Syracuse University,
United States

Z. Jane Wang—The University of British Columbia,
Canada

Gregory Wornell—Massachusetts Institute of
Technology, United States

Dapeng Wu—University of Florida, United States

**ASSOCIATE EDITORS—
COLUMNS AND FORUM**

Rodrigo Capobianco Guido —

São Paulo State University

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

Saeid Sanei, — University of Surrey,
United Kingdom

Azadeh Vosoughi — University of Central Florida

Danilo Mandic — Imperial College,
United Kingdom

Roberto Togneri — The University of Western
Australia

Gail Rosen — Drexel University

ASSOCIATE EDITORS—e-NEWSLETTER

Csaba Benedek—Hungarian Academy of Sciences,
Hungary

Paolo Braca—NATO Science and Technology
Organization, Italy

Quan Ding—University of California, San
Francisco, United States

Marco Guerriero—General Electric Research,
United States

Yang Li—Harbin Institute of Technology, China

Yuhong Liu—Penn State University at Altoona,
United States

Andreas Merentitis—University of Athens, Greece

IEEE SIGNAL PROCESSING SOCIETY

Alex Acero—*President*

Rabab Ward—*President-Elect*

Carlo S. Regazzoni—*Vice President, Conferences*

Konstantinos (Kostas) N. Plataniotis—*Vice*

President, Membership

Thrasyloulos (Thrasos) N. Pappas—*Vice President,*

Publications

Charles Bouman —*Vice President,*

Technical Directions

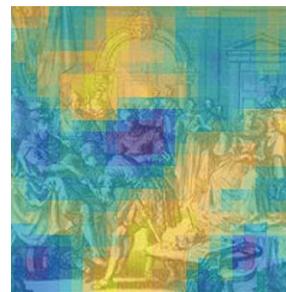
IEEE SIGNAL PROCESSING SOCIETY STAFF

Denise Hurley—Senior Manager of Conferences
and Publications

Rebecca Wollman—Publications Administrator

COVER

COURTESY OF RIJKSMUSEUM/NANNE VAN NOORD

**IEEE PERIODICALS
MAGAZINES DEPARTMENT**

Jessica Barragué
Managing Editor

Geraldine Krolin-Taylor
Senior Managing Editor

Mindy Belfer
Advertising Sales Coordinator

Felicia Spagnoli
Advertising Production Manager

Janet Dudar
Senior Art Director

Gail A. Schnitzer
Associate 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>.

SCOPE: IEEE Signal Processing Magazine publishes tutorial-style articles on signal processing research and applications, as well as columns and forums on issues of interest. Its coverage ranges from fundamental principles to practical implementation, reflecting the multidimensional facets of interests and concerns of the community. Its mission is to bring up-to-date, emerging and active technical developments, issues, and events to the research, educational, and professional communities. It is also the main Society communication platform addressing important issues concerning all members.

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\$213.00 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 ©2015 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.2015.2419313



Now... 2 Ways to Access the IEEE Member Digital Library

With **two great options** designed to meet the needs—and budget—of every member, the IEEE Member Digital Library provides full-text access to any IEEE journal article or conference paper in the IEEE *Xplore*® digital library.

Simply choose the subscription that's right for you:

IEEE Member Digital Library

Designed for the power researcher who needs a more robust plan. Access all the IEEE content you need to explore ideas and develop better technology.

- 25 article downloads every month

IEEE Member Digital Library Basic

Created for members who want to stay up-to-date with current research. Access IEEE content and rollover unused downloads for 12 months.

- 3 new article downloads every month

Get the latest technology research.

Try the IEEE Member Digital Library—FREE!

www.ieee.org/go/trymdl



IEEE Member Digital Library is an exclusive subscription available only to active IEEE members.

[from the **EDITOR**]Min Wu
Editor-in-Chief
minwu@umd.edu

Art, Engineering, and Community

I wrote this editorial on my way back from ICASSP 2015, which was held in Brisbane, Australia, 19–24 April. I, along with *IEEE Signal Processing Magazine's* (SPM's) area editors, presented to the magazine's Editorial Board and the IEEE Signal Processing Society's (SPS's) Publication Board the progress our magazine has made and the plans being carried out in the first part of this year. We also discussed ideas on how we can bring SPM to the next level and best serve our broad community.

One idea that we are working on is to expand ways to enhance the engagement and participation of our readers. Thanks to the efforts of Andres Kwasinski, our area editor for social media and outreach, SPM is now present on several major social media channels (including LinkedIn, Twitter, and Facebook)—so please find us on your favorite social media platform, and bring along your friends and colleagues in cyberspace.

Inspired by *Science's* efforts in recent years to bring young scientists' voices to a wide audience, we are launching a pilot effort through social media platforms to invite you—our readers and members—to offer insight toward two questions about the career and essence of signal processing. See “Share Your Answers” for the questions and links to these discussions. Please concisely explain your thoughts, and provide your contact information (anonymous answers will not be considered). Our editorial team will select enlightening answers and publish excerpts in a future issue of SPM or *Inside Signal Processing e-Newsletter*. Whether you are a student, an educator, a researcher, or a practitioner, readers at all stages of their careers are welcome!

Digital Object Identifier 10.1109/MSP.2015.2425151
Date of publication: 15 June 2015

You may notice several summarizing highlights in this issue of SPM as well as in previous issues. We are experimenting with these highlights to bring your attention to various initiatives and resources that the SPS boards and committees are working on, especially those in electronic forms. For example, in this issue, you will see highlights of informative resources from the first quarter of the monthly e-Newsletters—in case you may have missed them (page 7)—and documents other colleagues may be talking about on the newly opened SigPort repository (page 16). We hope you will find these highlights and summaries helpful, as they point to more in-depth resources you will find in online venues.

SHARE YOUR ANSWERS

- 1) With signal processing training, what do you consider a successful career?
- 2) What is the most unexpected example of signal processing in our daily lives?

We want to hear from you! Visit:

- LinkedIn: <http://linkd.in/1aEgGXd>
- Facebook: <http://www.facebook.com/ieeespm>.

You will also find a pictorial summary accompanying the guest editorial of this special issue (page 14). These representative graphics, selected with input from the guest editors, are assembled to give you an at-a-glance view of the 11 articles in this intriguing special issue where signal processing meets art history. I'd like to extend my sincere thanks to the guest editors for their hard work to help bridge these two communities!

Speaking of art and engineering, I made a brief stop during my return trip transiting through Sydney, Australia, to visit one of the ultimate embodiments of art and engineering—the Opera House in Sydney Harbour. The elegance of this iconic landmark designed by the Danish architect Jørn Utzon is unparalleled—whether viewed up close from land or from the ferries traveling on the water, the curves and shape of the Opera House are always inviting and refreshing.

On a tour offered by the Opera House, I learned the dramatic story behind its vision and construction: the diverging architectural assessments among even the best-known architects on this daring idea versus a “circus tent,” the skepticism and challenges on whether the curved roof could actually be built and how, the budget concerns and delay toward completion, and the resignation of the architect... yet ultimately, overcoming obstacles and controversy, the Opera House came to life from idea to icon, as a triumph of incredible creativity, perseverance, and collective effort by many!

As our signal processing community is exploring new ways to raise its visibility and strengthen its impact, we are surrounded by many ideas from devoted volunteers. Not all new ideas come easily accepted or are straightforward to implement. Perhaps the history behind the Sydney Opera House offers insights and encouragement for our community to embrace new ideas that help bring us closer and stronger, with a brighter future together!

SP



GVA AMPLIFIERS

NOW DC* to 12 GHz up to 1W Output Power from 94¢^{ea.} (qty. 1000)

GVA amplifiers now offer more options and more capabilities to support your needs. The new **GVA-123+** provides ultra-wideband performance with flat gain from 0.01 to 12 GHz, and new model **GVA-91+** delivers output power up to 1W with power added efficiency up to 47%! These new MMIC amplifiers are perfect solutions for many applications from cellular to satellite and more! The GVA series now covers bands from DC to 12 GHz with

various combinations of gain, P1dB, IP3, and noise figure to fit your application. Based on high-performance InGaP HBT technology, these amplifiers are unconditionally stable and designed for a single 5V supply in tiny SOT-89 packages. All models are in stock for immediate delivery! Visit minicircuits.com for detailed specs, performance data, export info, **free X-parameters**, and everything you need to choose your GVA today!

US patent 6,943,629

*Low frequency cut-off determined by coupling cap.
For GVA-60+, GVA-62+, GVA-63+, and GVA-123+ low cut off at 10 MHz.
For GVA-91+, low cut off at 869 MHz.

NOTE: GVA-62+ may be used as a replacement for RFMD SBB-4089Z
GVA-63+ may be used as a replacement for RFMD SBB-5089Z
See model datasheets for details

FREE X-Parameters-Based
Non-Linear Simulation Models for ADS



<http://www.modelithics.com/mvp/Mini-Circuits.asp>

Mini-Circuits®

www.minicircuits.com P.O. Box 350166, Brooklyn, NY 11235-0003 (718) 934-4500 sales@minicircuits.com

458 rev P

[president's MESSAGE]

Alex Acero
2014–2015 SPS President
a.acero@ieee.org



SigPort: A Document Repository for Signal Processing

The Signal Processing Repository (SigPort) is an online repository of manuscripts, reports, technical white papers, theses, and supporting materials. Created and supported by the IEEE Signal Processing Society (SPS), SigPort collects technical material of interest to the broad signal and information processing community, with categories covering each of the Society's technical committees. SigPort documents can be accessed for free at <http://www.sigport.org>.

Some of our SPS members said that they wanted a mechanism to disseminate their research quickly while their IEEE journal submission was being reviewed. The popular [arXiv.org](http://arxiv.org) Web site allows authors to upload e-prints in some areas within computer science, but it does not have an area for signal processing. Much like arXiv, SigPort hosts material to help individuals obtain early and broad exposure to their work. SigPort provides a time stamp for each uploaded document; a unique URL is assigned to the document, designating it as part of the IEEE SPS SigPort as well as for easy referencing.

Also similar to arXiv, SigPort papers are not peer reviewed. Authors retain all the rights to their documents and can submit them later to journals, conferences, books, etc., since submissions to

the SigPort repository are not as restricted as formal publications. We expect a majority of the e-prints to be submitted to one of the Society's journals for publication, but some works may remain purely as e-prints and will never be published in a peer-reviewed journal. SigPort documents are visible to citation and search engines



such as Google Scholar, which will combine citations to the e-print with citations to the final journal article.

SigPort can also archive research drafts, white papers, posters, presentation slides, lecture notes, student papers, auxiliary materials, and more.

Since SigPort was developed from scratch, we've built it with the latest Web technologies. SigPort documents can receive immediate private comments as feedback, public endorsement from colleagues who "like" your work, and potential highlights as recommended by an international editorial board. SigPort's

"thumbs-up" capabilities can help readers identify relevant documents, much like star ratings and popularity ratings have been helping users choose products in today's e-commerce Web. SigPort also has images, animations, and the search and browsing capabilities we expect from modern Web sites.

Maintaining the service does incur some cost, but for now the SPS bears the expense. We're exploring various mechanisms, including modest author fees, to ensure its long-term financial sustainability.

The SPS is always looking for new ways to deliver research, knowledge, and information to its members. If you have any ideas about how to keep our Society at the forefront of providing content, please don't hesitate to reach out to us. If you have questions or suggestions about SigPort, you can contact the team at sigportsps@gmail.com.

I hope that you'll take the time to check out some of our SigPort highlighted documents on page 16 of this issue of *IEEE Signal Processing Magazine* and even submit one of your own—make sure to use coupon code SPM0915 to upload your document(s) to SigPort for free. Please note that this coupon will expire on 30 September 2015.

Our Society and community is what we make it. Let's keep learning and sharing together.

[SP]

Digital Object Identifier 10.1109/MSP.2015.2425152

Date of publication: 15 June 2015

Inside Signal Processing eNewsletter

Highlights • Winter & Spring 2015

Signal Processing Contests Recent years have seen a rapid increase in the contests related to signal and information processing. These include those held at conferences and workshops sponsored or co-sponsored by SPS as well as many held on the Kaggle data science platform.

❖ A Kaggle contest on public health is to predict West Nile virus outbreak in a major U.S. city. [May 2015 Contest News]

❖ 2015 IEEE Signal Processing Cup is a competition on heart rate monitoring resilient to distortions introduced by physical exercises. SP cup encourages undergraduate students to work as a team to solve a challenging real-world problem using signal-processing techniques. Stay tuned for more in-depth analysis in an upcoming issue of SPM.

❖ A Kaggle contest on driver safety considered using telematic data to measure a driver's behavior qualitatively and quantitatively. Participants were asked to develop a "telematic fingerprint" capable of distinguishing when a trip was driven by a given driver. Feature engineering from the raw GPS data proved very important in this contest. [February 2015 Contest News]

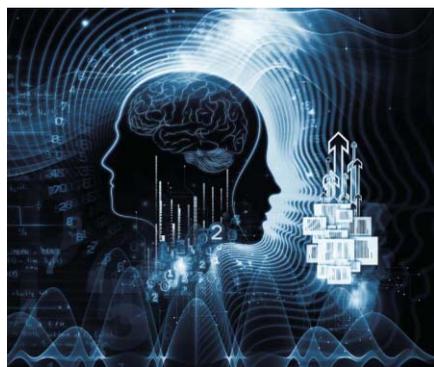
❖ 2015 ISBI Challenges: eight topics related biomedical imaging were selected. Each challenge

generally consists of a contest and a workshop held during ISBI 2015. [December 2014 Contest News]

❖ 2014 MLSP contest: organized as part of the IEEE SPS' MLSP workshop, was on schizophrenia classification. The contest was hosted in Kaggle and attracted 341 participants. Feature extraction and selection played an important role. [December 2014 TC News]

Find an interesting contest? Please let the eNews team know to share it with the community.

Recent Patents eNewsletter editors scanned through recently issued patents and organized those related to signal processing in focused topic areas.



- ❖ Audio Coding (January 2015)
- ❖ Infrared Imaging (February 2015)
- ❖ Deep Learning (March 2015)
- ❖ Wavelet Analysis (April 2015)
- ❖ Gait Recognition (May 2015)

❖ New Patent Policy related to Standards by IEEE (March 2015)

Distributed Deep Learning Network

Deep learning is showing a strong promise for a number of applications. An important issue to address is the computational complexity in training a deep learning network. The big data lab at Impetus Inc. in India, led by Dr. Vijay Srinivas Agneeswaran, has built a prototype of a distributed deep learning network over Apache Spark, whereby the training phase of the network has been parallelized and consequently the training time reduced. Learn about this in the February 2015 Trends section.

TIPS Read eNewsletter for monthly updates on a variety of topics. Enter keywords in the search box or choose "Previous Issues" at the eNewsletter website.



Inside Signal Processing eNewsletters welcomes community inputs and provides monthly updates on society and technical committee news; conference and publication opportunities, new patents, books, and Ph.D. theses; research opportunities and activities in industry consortia, local chapters, and government programs. Contact Area Editor for eNews Christian Debes <cdebes AT agtinternational.com> for more information. Bookmark <www.signalprocessingsociety.org/newsletter/>.

Human–Machine Interfaces: Methods of Control

If you are worried that artificial intelligence enabled systems are well on their way toward assuming total command of the planet, you can take some heart in the fact that there is still a great deal of important research being done in human–machine interfaces (HMIs), much of it involving signal processing. Making certain that various types of systems do precisely what their human masters demand lies at the heart of most HMI research.

The current HMI field is very competitive, and academic, government, and commercial researchers are working hard to create advanced technologies that are both useful and marketable. The major trends driving the sector include an ever-increasing demand for enhanced user efficiency; rapid growth in information technology and telecom sectors; and a continuing expansion of electronic, mobile, computer, and electromechanical applications.

COMMAND BY GESTURE

Smartphones have become increasingly affordable and more widely used over the past several years. Yet smartphones and their applications are difficult to control in situations where the user lacks direct access to the touchscreen, such as while driving a car, cooking a meal, or exercising. While voice recognition technology promises a partial solution to the problem, such systems are far from foolproof and particularly unreliable in noisy environments.



[FIG1] University of Washington researchers have created a new type of low-power wireless sensing technology that promises to allow users to “train” their smartphones to recognize and respond to specific hand gestures made near the phone. (Photo courtesy of the University of Washington.)

In an effort aimed at creating an alternate “hands off” control technology, University of Washington researchers have created a new type of low-power wireless sensing technology that promises to allow users to “train” their smartphones to recognize and respond to specific hand gestures (Figure 1). The new SideSwipe technology developed in the labs of Matt Reynolds and Shwetak Patel, both associate professors of electrical engineering and of computer science, uses the phone’s own wireless transmissions to sense and recognize nearby hand gestures.

“Current smartphones use a variety of different built-in sensors, such as accelerometers and gyroscopes, that can track the motion of the phone itself,” Reynolds says. “We have created an entirely new type of sensor that uses the reflection of the phone’s own wireless signal to detect nearby gestures, allowing users to interact with their phones even when they are not holding the phone, looking at the display, or touching the screen.”

Whenever someone uses a smartphone for voice or data communication, the device transmits radio signals on a 2G, 3G, or 4G cellular network to communicate with a nearby cellular base station. SideSwipe takes advantage of the fact that when a user’s hand moves through space near the phone, the user’s body reflects some of the transmitted signal back toward the device. SideSwipe uses multiple small antennas to capture changes in the reflected signal and classify the changes to detect the specific type of gesture performed. The result is that

hovering, tapping, and sliding gestures can be associated with various phone commands, such as silencing a ring, changing a song, or muting the speakerphone.

“The GSM signal that we are working with was originally designed for communication, but we are analyzing the signal in a different way, from the perspective of pulling out gestures,” Reynolds remarks. He adds that he and coresearcher Patel were inspired by radar technology. “In the case of radar, you have a controlled emitter of an electromagnetic wave that bounces off an aircraft or a ship or something like that and comes back,” he says. “We realized that the same thing is happening all the time to the cell phone transmission.”

Reynolds says that signal processing is essential to the technology. There are multiple phases of signal processing, ranging from very simple filtering to more sophisticated machine learning, he says. “We used signal smoothing, band-pass filtering to extract the frequency band that has useful gesture information in it, which tends to be a very low-frequency signal,” he explains. “Then we use

a machine-learning technique called the *support vector machine* to do the classification of features into specific gestures.”

When developing their prototype, the researchers added a receiver with four directionally sensitive antenna elements to the smartphone's case. “Then we looked at the signal waveforms that were coming from that receiver and figured out a way of ignoring the fact that the signal was originally designed for communication and instead looked at its envelope to see changes in reflection,” Reynolds says.

SideSwipe leverages the unmodified GSM bursts that inherently exist when someone is using a smartphone. When the user performs a particular type of hand gesture, the antennas pick up the fluctuation in their respective propagation paths. By combining the signals from four antennas, the researchers were able to identify unique patterns for different gestures.

A group of ten study participants tested SideSwipe with 14 different hand gestures, such as tapping, hovering, and sliding, at various positions and distances from the smartphone. The smartphone was calibrated to its user's hand movements prior to each test. The smartphone recognized gestures with about 87% accuracy, Reynolds says.

“We are interested in interaction in cases where you are not holding the phone,” Reynolds states. “If you think about the use of the phone during the day, most of the time people have the phone in their pocket or in a handbag or, let's say, on a table.” Because the SideSwipe sensor is based on low-power receivers and relatively simple signal processing when compared with something like camera video, Reynolds expects that SideSwipe will have only a minimal impact on battery life.

Reynold feels that the technology is still at a preliminary stage. “We are facing a much longer series of research leading to even more efficient ways of extracting gesture information, whether that is different signal processing strategies or new



[FIG2] Jan Scheuermann (right) reaches out with the thought-controlled robot arm to touch Jennifer Collinger's hand. (Photo courtesy of the University of Pittsburgh Medical Center.)

machine-learning algorithms,” he says. “Currently, we are using extra receivers that are built into something like a snap-on phone case, but it is likely that the hardware could eventually be built into the phone itself—that would require the cooperation of the phone manufacturer.”

DOING BY THINKING

Futurists and science fiction writers have long predicted that people one day will be able to control various types of devices by thought alone. Such technology would allow individuals to operate remote vehicles, machinery located inside mines and other dangerous places, and a variety of other simple and complex devices conveniently and across any distance. Even more importantly, a thought-driven HMI would allow people who have lost the use of their limbs to control robotic systems that provide mobility or the ability to grasp and manipulate various types of objects, ranging from eating utensils to doorknobs to light switches.

University of Pittsburgh researchers are investigating an HMI technology that

promises to allow people to operate a robotic arm, capable of mimicking natural arm and hand movements, simply by thinking about whatever task that needs to be performed. Working with Jan Scheuermann, a 55-year-old Pittsburgh woman who has been paralyzed from the neck down since 2003 due to a neurodegenerative condition, the researchers have been able to increase the robotic arm's maneuverability from seven dimensions to ten dimensions over the past three years (Figure 2).

The additional dimensions result from four hand movements—finger abduction, a scoop, thumb extension, and a pinch—allowing Scheuermann to pick up, grasp, and move a range of objects much more precisely. “She can now pinch the fingers, flex them all together, spread the fingers apart, and then move the thumb independently,” says Jennifer

2016-2017 IEEE-USA Government Fellowships



Congressional Fellowships

Seeking U.S. IEEE members interested in spending a year working for a Member of Congress or congressional committee.



Engineering & Diplomacy Fellowship

Seeking U.S. IEEE members interested in spending a year serving as a technical adviser at the U.S. State Department.



USAID Fellowship

Seeking U.S. IEEE members who are interested in serving as advisors to the U.S. government as a USAID Engineering & International Development Fellow.

The application deadline for 2016-2017 Fellowships is 15 January 2016.

For eligibility requirements and application information, go to www.ieeeusa.org/policy/govfel or contact Erica Wissolik by emailing e.wissolik@ieee.org or by calling +1 202 530 8347.



special REPORTS continued

Collinger, an assistant professor in the University of Pittsburgh's Department of Physical Medicine and Rehabilitation, a project lead investigator. "This greatly increased the amount of function in the hand."

In 2012, Scheuermann underwent surgery in the regions of her brain responsible for right arm and hand movements. The operation fitted her with a pair of quarter-inch electrode grids, each containing 96 tiny contact points. After the electrode grids on Scheuermann's brain were linked to a computer, creating a brain-machine interface (BMI), their contact points could detect electrical pulses firing between the brain's neurons.

"In terms of signal processing, I would say our system is maybe a little bit simple," Collinger says. "It is more on the data analysis and computational side where it gets complex." Signal filtering is used to remove as much noise as possible. "Then we use linear regression techniques to find a relationship between the firing rates and the movement parameters we are trying to decode," Collinger explains.

Algorithms decode the firing signals and also identify the patterns associated with a particular arm movement, such as lifting the arm or twisting the wrist. "Our results show that individual motor cortical neurons encode many parameters of movement, that object interaction is an important factor when extracting these signals, and that high-dimensional operation of prosthetic devices can be achieved with simple decoding algorithms," Collinger says.

When developing the algorithms, Collinger and her coresearchers followed a unique approach based on neurobiological principles. They first optimized a population vector algorithm that encodes prosthetic movements on the basis of direction-dependent tuning of the motor cortex neuronal ensemble. The decoded movement was then sent to a shared controller that integrated the user's intent, position feedback, and various constraining task-dependent features to optimally guide the robotic arm movements.



[FIG3] Researchers in the LTS5 of the EPFL are investigating a technology that reads and identifies facial expressions indicating various moods, such as anger, disgust, fear, happiness, sadness, and surprise. (Photo courtesy of EPFL.)

In the lab, the robotic arm is set up next to Scheuermann on a stand. "We positioned it close enough to her to allow her to feed herself, to take a drink, those kinds of things," Collinger says. "I certainly think that it could be mounted to her wheelchair."

A cable connects the electrode grids on Scheuermann's brain to the robotic arm. Yet many BMI researchers foresee the day when user commands are transmitted wirelessly to robotic arms and other electromechanical systems. Steps are already being taken in that direction. Last December, Brown University researchers announced a new high data-rate, low-power wireless brain sensor. The head-mounted, 100-channel transmitter is only 5 cm in its largest dimension and weighs just 46.1 g but can transmit data at up to 200 megabits a second. The technology is currently designed to enable neuroscience research that cannot be accomplished with existing sensors that tether subjects with cabled connections.

During the next research phase, Collinger and her colleagues plan to investigate additional ways of making that arm more controllable. "We plan, for instance, to study whether the incorporation of sensory feedback, such as the touch and feel of an object, can improve neuroprosthetic control," she says.

READING FACIAL EXPRESSIONS

HMIs are advancing to the point where they can function subliminally, potentially saving users from dangerous situations of which they may not even be aware. Showing that HMIs are not only about system

control, researchers in the Signal Processing 5 Laboratory (LTS5) of the Ecole Polytechnique Fédérale de Lausanne (EPFL) have created a technology that reads and identifies human facial expressions indicating various moods, such as anger, disgust, fear, happiness, sadness, and surprise (Figure 3). Such a system could prove useful in several fields, including video game development, medicine, and marketing. The EPFL researchers, however, are most interested applying the technology driver safety systems.

"Certain emotional states of the driver, such as stress, rage, or strong euphoria, affect decision making and coordination skills and may cause discomfort and loss of concentration/attention and may give way to accidents," says principal investigator Jean-Philippe Thiran, an EPFL professor. "In this project, we use computer vision and machine-learning methods to detect these emotional states from the drivers' facial expressions." The first studies involved the emotion of stress and the state of fatigue.

Thiran sees significant potential in the research, which is being sponsored by automaker PSA Peugeot Citroën and automotive equipment supplier Valéo. "Some potential applications are an alert system, for instance, in the case of detecting fatigue, or a countermeasure for the certain emotions that may endanger driving quality, such as the activation of a 'driver calming system' using sounds, lights, or odors that are personalized for the driver."

Tests carried out with a prototype proved promising. "We hope that with additional studies we will be able to define the emotional states that are dangerous for driving and also how to reduce them," Thiran says. The researchers' goals are a more comfortable driving experience and fewer accidents attributable to drivers' emotional states. "What is commonly known as 'road rage' or 'aggressive driving,' for example, is a very common cause of car accidents and it is a state that we hope to detect and take measures against," Thiran remarks.

“Signal processing is at the core of research, as in all imaging systems,” Thiran says. “[The system] includes robust face detection, feature extraction, and classification,” he says, noting that a couple of approaches are used to extract discriminative features and to uncover patterns of different facial expressions. “We investigate approaches based on holistic affine warping and local descriptors,” Thiran says.

Holistic affine warping normalizes face images using the coordinates of the left and right eyes. “The locations of the eye centers are derived from the tracked facial landmarks,” Thiran says. “After applying an affine transform, the eye centers are fixed in the canonical coordinates in the normalized image.” A second approach extracts local descriptors from around the tracked facial landmarks. “It preserves the geometrical information of the facial components and does not introduce additional artifacts,” he explains.

“The techniques we are using are common to most facial analysis systems: a face-tracking system that locates the drivers face from the image captured by the in-car system, a head-pose normalization algorithm that compensates for the angle of view, extraction of relevant

appearance features from the tracked face, and classifiers to detect whether the expression we are interested in is present,” Thiran states.

“Having in-car conditions involves many additional requirements compared to indoor systems or standard human-computer interaction (HCI) systems in which you may restrict the user to face

HMIs ARE ADVANCING TO THE POINT WHERE THEY CAN FUNCTION SUBLIMINALLY, POTENTIALLY SAVING USERS FROM DANGEROUS SITUATIONS OF WHICH THEY MAY NOT EVEN BE AWARE.

the camera at all times,” Thiran says. “As this it not a possibility during driving, we had to come up with a solution to compensate for the up-tilted angle of face-view, that results from the special camera configuration in the car.” The system prototype applied a head-pose correction that used a simple three-dimensional model to project the driver’s face image on a

two-dimensional plane for classification. “We are now working on a facial reconstruction from multicamera system that will enable emotion detection totally regardless of where the driver is looking at,” Thiran says.

Vehicle interior lighting conditions, which change frequently during a journey, also posed a challenge for the researchers. “To tackle that problem, we used near-infrared (NIR) cameras with a special lighting system and adequate filtering so that ambient light effects are at minimum,” Thiran says. “We had to rebuild classifiers for expression using NIR images and also applied model adaptation techniques so that we can make use of existing facial expression databases of color images, which are well annotated and validated.”

The researchers are now working on detecting other expressions on drivers’ faces, such as distraction and lip reading for use in vocal recognition.

AUTHOR

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



Do you know? IEEE Signal Processing Magazine goes social!

Follow us on LinkedIn and Facebook, and share your thoughts on the questions posted there.



<http://linkd.in/1aEgGXd>



Your input may appear as the Community’s Voice in a future issue of the magazine.



www.facebook.com/ieeespm

Digital Object Identifier 10.1109/MSP.2015.2434231

[society NEWS]

New Society Officers for 2016 and Election of Directors-at-Large and Members-at-Large

The Board of Governors (BoG) of the IEEE Signal Processing Society (SPS) elected two new officers who will start their terms on 1 January 2016: Ali H. Sayed and Walter Kellermann. Ali H. Sayed [University of California, Los Angeles (UCLA)] will serve as 2016–2017 SPS president-elect. He succeeds Rabab K. Ward (University of British Columbia) who has held the post of president-elect and will become president in 2016. Walter Kellermann (University of Erlangen-Nuremberg) will serve as 2016–2018 SPS vice president-technical directions. He succeeds Charles A. Bouman (Purdue University), who has held this position since January 2013.

NEW SOCIETY OFFICERS

2016–2017 SPS PRESIDENT-ELECT

ALI H. SAYED



He is a professor and former chair of electrical engineering at UCLA, where he directs the UCLA Adaptive Systems Laboratory (<http://www.ee.ucla.edu/asl>). An author of more than 450 scholarly publications, six books, and five patents, his research involves several areas including adaptation and learning, statistical signal processing, network science, information processing theories, and biologically inspired designs. His work has been recognized with several awards from various organizations including the 2014 Athanasios Papoulis Award from the

Digital Object Identifier 10.1109/MSP.2015.2428310

Date of publication: 15 June 2015

European Association for Signal Processing, the 2013 Meritorious Service Award and the 2012 Technical Achievement Award from the IEEE SPS, the 2005 Terman Award from the American Society for Engineering Education, the 2005 Distinguished Lecturer from the IEEE SPS, the 2003 Kuwait Prize, and the 1996 IEEE Donald G. Fink Prize. He has also been awarded several Best Paper Awards from the IEEE SPS (2002, 2005, 2012, and 2014) and is a Fellow of both the IEEE and the American Association for the Advancement of Science; the publisher of the journal *Science*. He is recognized as a highly cited researcher by Thomson Reuters.

Prof. Sayed has been active in serving the IEEE SPS in various capacities across major boards and committees, which include: editor-in-chief, *IEEE Transactions on Signal Processing* (2003–2005); general chair, ICASSP 2008; vice president-publications (2009–2011); member, Awards Board (2005), Publications Board (2003–2005, 2009–2011), Conference Board (2007–2011), Technical Directions Board (2008–2009), BoG (2007–2011), Signal Processing Theory and Methods Technical Committee (2002–2006), Signal Processing for Communications Technical Committee (2002–2007), Sensor Array and Multichannel Technical Committee (2011–present), Machine Learning for Signal Processing Technical Committee (2012–present), and Big Data Special Interest Group (2015–present). He also served as chair and vice chair of the Signal Processing Theory and Methods Technical Committee during 2007–2009. At the IEEE level, he served as a member of the Technical Activities Board Periodicals Review Committee (2007–2009) and the

IEEE Committee on the Future of Education in Engineering and Computing (2011). Concurrent with his services to the Society, Prof. Sayed was the chair of electrical engineering at UCLA (2005–2010) and has led a large academic department with close to 45 faculty members, more than 15 lecturers and 30 staff members, and close to 1,000 undergraduate and graduate students.

2016–2018 SPS VICE PRESIDENT-TECHNICAL DIRECTIONS

WALTER KELLERMANN



He is a professor of multimedia communications and signal processing at the University of Erlangen-Nuremberg, Germany, and heads the chair's audio research group comprising more than a dozen researchers. He received the Dipl.-Ing. degree in electrical engineering from the University of Erlangen-Nuremberg in 1983 and the Dr.-Ing. degree from the Technical University Darmstadt, Germany, in 1988. From 1989 to 1990, he was a postdoctoral member of technical staff at AT&T Bell Laboratories, Murray Hill, New Jersey. In 1990, he joined Philips Kommunikations Industrie, Nuremberg, Germany. From 1993 to 1999, he was a professor, Fachhochschule Regensburg, where he also became the director of the Institute of Applied Research in 1997. In 1999, he cofounded DSP Solutions, a consulting firm in digital signal processing, and joined the University Erlangen-Nuremberg as a professor and head of the Audio Research Laboratory. From 2002 to 2006, he was the dean of studies, Faculty of Engineering, University

Erlangen-Nuremberg, overseeing organization and quality management for more than 500 courses per semester.

He also continuously consults at the industry level in his technical field, and he acts on various boards and as a consultant to governmental institutions on the national and European level regarding research and university politics. His current research interests include speech signal processing, array signal processing, adaptive filtering, and its applications to acoustic human/machine interfaces.

Prof. Kellermann has authored or co-authored 15 book chapters and more than 200 refereed papers in journals and conference proceedings and more than 50 granted patents. He has served as a guest editor for various journals, as associate editor (and guest editor), *IEEE Transactions on Speech and Audio Processing* (2000–2004), *EURASIP Journal on Applied Signal Processing* and *EURASIP Journal on Signal Processing* (2005–2007); general chair, 2003 International Workshop on Microphone Arrays, 2005 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, and 2014 ITG Conference on Speech Communication; general cochair, 2008 and 2011 International Workshop on Hands-Free Speech Communication and Microphone Arrays; cochair, International Workshop on Acoustic Signal Enhancement; SPS Distinguished Lecturer (2007–2008); chair, SPS Audio and Acoustic Signal Processing Technical Committee (2008–2010); member, IEEE SPS Overview Article Editorial Board (2010–present) and IEEE Flanagan Award Committee (2012–2014); and member-at-large, SPS BoG (2013–2015). He is a Fellow of the IEEE and received the Julius von Haast Fellowship Award from the Royal Society of New Zealand in 2011 and will receive the EURASIP Technical Achievement Group Award in 2015.

ELECTION OF DIRECTORS-AT-LARGE AND MEMBERS-AT-LARGE

Your vote is important! The election of directors-at-large for Regions 7 and 9



IMAGE LICENSED BY GRAPHIC STOCK

and Region 10 (the term is 1 January 2016–31 December 2017) and members-at-large (the term is 1 January 2016–31 December 2018) of the IEEE SPS BoG is now open. Ballots have been mailed to SPS members. The ballot includes a diverse slate of candidates for both elections, which were vetted by the SPS Nominations and Appointments Committee, as well as a space for write-in candidates.

THIS YEAR'S ELECTION OFFERS SPS MEMBERS THE OPPORTUNITY TO CAST THEIR VOTES VIA THE WEB AT [HTTPS://EBALLOT4.VOTENET.COM/IEEE](https://eballot4.votenet.com/IEEE).

This year's election offers SPS members the opportunity to cast their votes via the Web at <https://eballot4.votenet.com/IEEE> for up to one regional director-at-large for your corresponding Region: Regions 7 and 9 (Canada and Latin America) and Region 10 (Asia and Pacific Rim) and three member-at-large candidates. The IEEE Technical Activities Society Services Department at the IEEE must receive all ballots no later than 31 August 2015 to be counted. Members must meet the eligibility requirements at the time the ballot data is generated to be able to vote. To be eligible to vote in this year's Society election, you must have been an active SPS member or affil-

iate (excluding student member) prior to 1 June 2015. This is the date when the list of eligible Society voting members was compiled. The candidates for regional director-at-large are:

- *Regions 7 and 9:* Paulo S. Diniz and Fabrice Labeau
- *Region 10:* K.V.S. Hari and Hitoshi Kiya

The candidates for member-at-large are:

- Fulvio Gini
- Robert W. Heath
- Lina Karam
- Anthony Kuh
- Shoji Makino
- Nasir Memon
- Wan-Chi Siu
- Paris Smaragdis
- Min Wu.

The BoG is the governing body that oversees the activities of the SPS. The SPS BoG has the responsibility of establishing and implementing policy and receiving reports from its standing boards and committees. The BoG comprises the following 21 Society members: six officers of the Society who are elected by the BoG, nine members-at-large elected by the voting members of the Society, four regional directors-at-large elected locally by Society voting members of the corresponding region, as well as the Awards Board chair. The six officers are: the president, the president-elect, the vice president-membership, the vice president-conferences, the vice president-publications, and the vice president-technical directions. The executive director of the Society shall serve ex-officio, without vote.

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.

Regional directors-at-large are SPS members who are elected locally by Society voting members of the corresponding Region via the annual election, to serve on the Society's BoG as nonvoting members and voting members of the Society's Membership Board. **[SP]**

[from the **GUEST EDITORS**]Patrice Abry, Andrew G. Klein,
William A. Sethares, and
C. Richard Johnson, Jr.

Signal Processing for Art Investigation

Why should signal processors care about problems in cultural heritage? What can signal processing offer to the world of art scholarship and preservation?

These complementary questions have motivated recent movement in our two communities—one composed of signal processors and art experts in the other—to collaborate and try to focus the signal processing technologies on underresolved problems of art investigation.

At the start of this century, a plenary talk [1] delivered at the 2001 IEEE International Conference on Image Processing divided the subjects of a two-decade long cross-disciplinary effort between a Paris technical university (ENST) and the French national conservation labs (C2RMF) into two broad categories: 1) archiving and consulting and 2) picture processing for the fine arts. The first category incorporates image acquisition and database access. The second includes image enhancement and restoration, crack network detection, multi-source image fusion, color processing, and geometric analysis. Consider a different division of these topics into three groups: 1) image acquisition, 2) image manipulation, and 3) image feature mining. The topics of research in the 1980s and 1990s described in [1] give more attention to the first two of these groups than to the third.

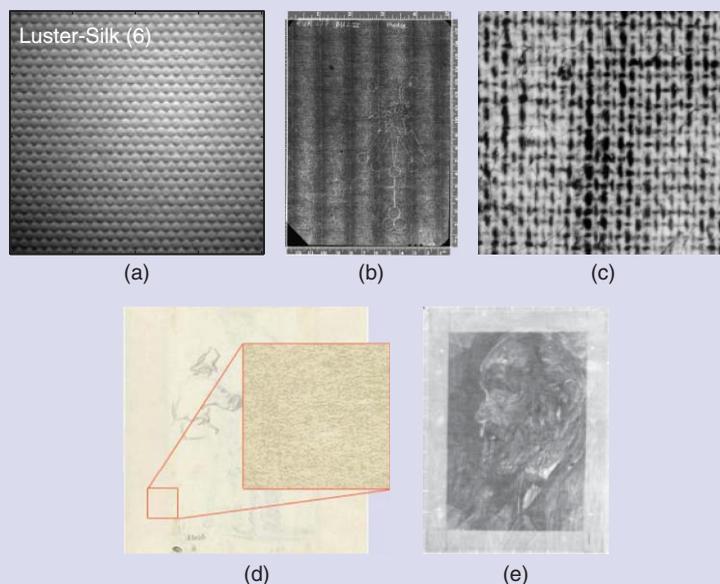
IEEE Signal Processing Magazine (SPM) has recognized previously—with a special section “Recent Advances in Applications to Visual Cultural Heritage” in its July 2008 issue [2]—the growing

interest at the start of this century in the nascent application of signal processing to art investigation. Following a vivid description of a wide range of tasks to be addressed in applying signal processing to the analysis of visual cultural heritage, six articles were grouped by the guest editors [3] into three themes: multispectral imaging, artwork analysis, and three-dimensional (3-D) digitization and modeling. By our division, this special section exhibits a balanced emphasis on

all three groupings of image acquisition, manipulation, and feature mining.

The March 2013 special issue of *Signal Processing* devoted to “Image Processing for Digital Art Work” divides [4] its seven papers into two categories: artwork restoration tools (X-ray fluorescence, crack detection) or art piece analysis procedures (canvas weave analysis, painting stylometry analysis). Here the emphasis according to our grouping has shifted to image feature mining.

SURFACE ANALYSIS FOR IDENTIFYING ARTISTIC STYLE AND ARTWORK ORIGIN



[FIGS1] A number of articles in this special issue highlight the important role of texture and surface analysis in art investigations. (a) Abry et al. employ multiscale anisotropic texture analysis to facilitate classification of historical photographic prints; (b) Johnson et al. examine chain-line features to hunt for moldmates among Rembrandt's prints, which are papers made using the same mold; (c) van der Maaten et al. focus on thread-level canvas analysis from X-ray data of the canvas supports of grand master paintings such as those by Van Gogh, and thus facilitate dating and authentication; (d) van Noord et al. learn artists' styles by applying deep neural network to detailed texture and other visual stylistic features from artworks; and (e) Yang et al. discuss the advantages of two-dimensional synchrosqueezed transforms toward quantitative analysis of canvas weave.

Digital Object Identifier 10.1109/MSP.2015.2419311

Date of publication: 15 June 2015

OBJECT RECOGNITION FOR MANAGING ART DATA



[FIGS2] (a) Anwar et al. review the classification of ancient coins by recognizing reverse motif on coins; (b) Hu et al. outline an integrated system to facilitate the study of Mayan writing, with glyph analysis and retrieval playing an important role; (c) Srinivasan et al. discuss challenges of face recognition from Renaissance portrait arts to facilitate art historians in identifying the subjects in these portraits; and (d) Picard et al. present challenges and solutions for content-based image indexing of cultural heritage collections including a wide variety of artifacts such as this tapestry which may be assigned such labels as “landscape,” “ornamentation,” “river,” “village,” “water mill,” and more.

ART ENHANCEMENT AND RESTORATION



[FIGS3] (a) Lozes et al. examine graph signal processing approaches to cultural heritage applications, such as facilitating colorization of historical art pieces; (b) Pizurica et al. discuss inpainting and other digital image processing for the restoration effort of *The Ghent Altarpiece*.

A recent plenary talk [5] delivered at the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing describes three separate tasks in computational art history utilizing signal processing algorithms to match manufactured patterns in art supports. In another plenary addressing image processing for art investigation delivered in 2014, the majority of projects described also include tasks in the category of image feature mining [6]. All of these tasks fall in the category of image feature mining.

While early research in applying signal processing to art investigation emphasized image acquisition and image manipulation, various manifestations of image feature mining have achieved more prominence in the past decade as evidenced by the preceding citations providing successive indicators of the directions being

taken in this growing field. The 11 articles in this special issue of *SPM* include studies of photographic paper classification, ancient coin classification, Mayan epigraphy analysis, 3-D color print graph signal processing, laid paper chain-line pattern matching, content-based image indexing, canvas weave analysis, crack detection for simulated in-painting, painting style characterization, and face recognition in portraits (see “Surface Analysis for Identifying Artistic Style and Artwork Origin,” “Object Recognition for Managing Art Data,” and “Art Enhancement and Restoration” for visual examples of these topics). All of these articles engage, to a substantial degree, in image feature mining. There is an enormous need in art history and conservation to locate, classify, identify, and measure features in multispectral and multidirectional images of artworks.

One barrier of entry into this field is access to images of sufficient quality and quantity of artworks. Therefore, for this special issue, the authors were requested to make the images they processed accessible to other researchers (when possible) to stimulate the exploration of new and improved solutions, and most have complied. With that added bonus, we welcome you to the modest beginnings of an emerging field that promises to offer many satisfying challenges and will require specialization of and advances in the tools and techniques of signal processing. To assist you in deciding which of the articles to use as your portal into this new domain full of ample, low-hanging fruit and many as yet undiscovered puzzles, we offer the summaries available at <http://sigport.org/189>.

Authors were encouraged to have a representative from the cultural heritage

from the **GUEST EDITORS** continued

community on their team, which is the case for most of the articles in this issue. Such cross-disciplinary collaborations are difficult. The art expert needs to acquire an appreciation of the range and limitations of the signal processor's tools so useful, viable, novel tasks can be identified and addressed. The signal processing experts must learn to describe their skills without resorting to the language of deep mathematics, and to appreciate the intellectual depth of the art expert achieved without directly resorting to computational tools. Both sides are facing new ways of thinking and conducting research. Of course, all of this makes the challenge that much more appealing to us. So it goes...

MEET THE GUEST EDITORS



Patrice Abry (patrice.abry@ens-lyon.fr) is a senior researcher with the CNRS, France. He is developing a research program on the

theoretical modeling, analysis, and applications of scale invariance.



Andrew G. Klein (andy.klein@wwu.edu) is an assistant professor at Western Washington University, Bellingham. His research interests include signal and image processing, distributed communications, and STEM education.



William A. Sethares (sethares@wisc.edu) is a professor in the Department of Electrical and Computer Engineering at the University of Wisconsin and a scientific researcher at the Rijksmuseum in Amsterdam, The Netherlands.



C. Richard Johnson, Jr. (johnson@ece.cornell.edu) is the Geoffrey S.M. Hedrick Senior Professor of engineering at Cornell University,

Ithaca, New York; a scientific researcher of the Rijksmuseum in Amsterdam, The Netherlands; and computational art history advisor to The Netherlands Institute for Art History.

REFERENCES

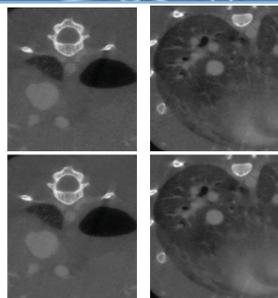
- [1] H. Maître, F. Schmitt, and C. Lahanier, "15 years of image processing and the fine arts," in *Proc. 2001 IEEE Int. Conf. Image Processing*, Oct. 2001, vol. 1, pp. 557–561.
- [2] *IEEE Signal Process. Mag.*, vol. 22, no. 4, July 2008.
- [3] M. Barni, A. Pelagotti, and A. Piva, "Image processing for the analysis and conservation of paintings: opportunities and challenges," *IEEE Signal Process. Mag.*, vol. 22, no. 5, pp. 141–144, Sept. 2008.
- [4] P. Abry, J. Coddington, I. Daubechies, E. Hendriks, S. Hughes, D. Johnson, and E. Postma, "Special issue guest editor's foreword," *Signal Processing* (Special Issue: Image Processing for Digital Art Work), vol. 93, pp. 525–526, Mar. 2013.
- [5] C. R. Johnson, Jr. (2014, May 6), "Signal processing in computational art history," in *Plenary 2014 IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Florence, Italy. [Online]. Available: <http://www.icassp2014.org/final200614/er2/0001/index.html>
- [6] I. Daubechies. (2014, July 3), "Signal analysis helping art historians and conservators," in *Plenary at IEEE 2014 International Symposium on Information Theory*, Honolulu, HI. [Online]. Available: <http://www.ee.hawaii.edu/~isit/Current/plenary-speakers.php>

SP



Spring 2015 SigPort Highlights

- ❖ Optimum decision fusion in cognitive wireless sensor Networks With Unknown Users Location (Abrardo et al.)
- ❖ Information theoretical limit of operation forensics (Chu et al.)
- ❖ A fast weighted stochastic gradient descent algorithm for image reconstruction in 3D computed tomography (Karimi et al.)
- ❖ Text-dependent speaker verification and RSR2015 speech corpus (Larcher et al.)
- ❖ Natural sound rendering for headphones (Sunder et al.)



An IEEE SPS community repository on signal and information processing, SigPort welcomes research drafts, white papers, theses, slides, posters, lecture notes, dataset descriptions, product brief, and more.

Try out SigPort Today at www.sigport.org with promotion code **SPM0915** (expires 30 September 2015). Click "About SigPort" for tutorials.



Digital Object Identifier 10.1109/MSP.2015.2429753

While the world benefits from what's new,
IEEE can focus you on what's next.

IEEE *Xplore* can power your research
and help develop new ideas faster with
access to trusted content:

- Journals and Magazines
- Conference Proceedings
- Standards
- eBooks
- eLearning
- Plus content from select partners

IEEE *Xplore*® Digital Library

Information Driving Innovation

Learn More

innovate.ieee.org

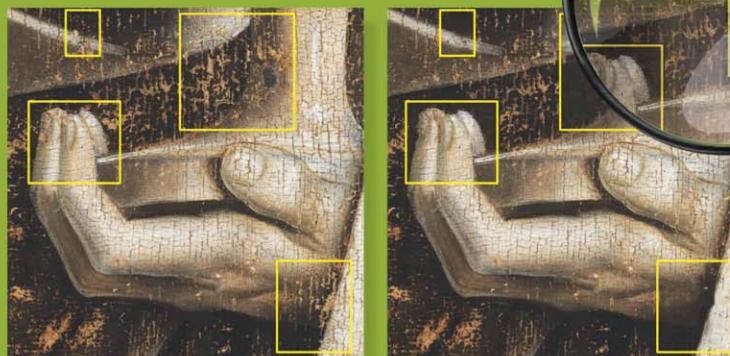
Follow IEEE *Xplore* on  

 **IEEE**
Advancing Technology
for Humanity

Patrice Abry, Stéphane G. Roux, Herwig Wendt, Paul Messier, Andrew G. Klein, Nicolas Tremblay, Pierre Borgnat, Stéphane Jaffard, Béatrice Vedel, Jim Coddington, and Lee Ann Daffner

Multiscale Anisotropic Texture Analysis and Classification of Photographic Prints

© GHENT KATHEDRALE KERKABRIEK, PHOTO COURTESY OF KIKIRPA, BRUSSELS, MANQU'YVING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

[Art scholarship meets image processing algorithms]

Texture characterization of photographic prints can provide scholars with valuable information regarding photographers' aesthetic intentions and working practices. Currently, texture assessment is strictly based on the visual acuity of a range of scholars associated with collecting institutions, such as museum curators and

conservators. Natural interindividual discrepancies, intraindividual variability, and the large size of collections present a pressing need for computerized and automated solutions for the texture characterization and classification of photographic prints. In the this article, this challenging image processing task is addressed using an anisotropic multiscale representation of texture, the hyperbolic wavelet transform (HWT), from which robust multiscale features are constructed. Cepstral distances aimed at ensuring balanced multiscale contributions are computed between pairs of images. The resulting large-size affinity matrix is then clustered

Digital Object Identifier 10.1109/MSP.2015.2402056

Date of publication: 15 June 2015

using spectral clustering, followed by a Ward linkage procedure. For proof of concept, these procedures are first applied to a reference data set of historic photographic papers that combine several levels of similarity and second to a large data set of culturally valuable photographic prints held by the Museum of Modern Art in New York. The characterization and clustering results are interpreted in collaboration with art scholars with an aim toward developing new modes of art historical research and humanities-based collaboration.

INTRODUCTION

Surface texture is a defining characteristic of a photographic print and has a significant impact on the quality and perception of the image. Manufacturers of both traditional photographic paper and new-generation inkjet materials thus carefully engineer surfaces to meet a variety of functional and expressive requirements. Smoother surfaces typically produce images with greater optical saturation and tonal range. With good lighting, smooth, featureless surfaces can visually negate the picture plane. Reducing this perceptual barrier can transform a two-dimensional (2-D) image into a more effective illusion of objective reality. Rougher surfaces cause more scattering of incident light providing the viewer with a greater tactile sense of the print as a material object. A stronger physical presence can often convey heightened intrinsic value and expressive weight. Understanding how these qualities are manipulated provides scholars insight into artistic intent and practice. More practically, as an indelible physical attribute, print texture can help categorize preferred and anomalous papers within an artist's body of work or identify anomalies (including fakes). Encyclopedic reference collections of such textures, cataloged by manufacturer and date, are currently being assembled for both traditional [1] and inkjet [2] photographic materials. Likewise, the very beginning steps are underway to catalog surface textures used by prominent photographers such as Man Ray and Lewis Hine, among others [3]. While presently useful, such surface texture collections are difficult to catalog and access, as tools for query and retrieval are only in early stages of development. At present, experts visually and manually classify an unknown texture by comparing it with identified references. This is a tedious and challenging task due to the sheer size of available reference collections exposing a significant need for (semi)automated procedures to assist in texture assessment.

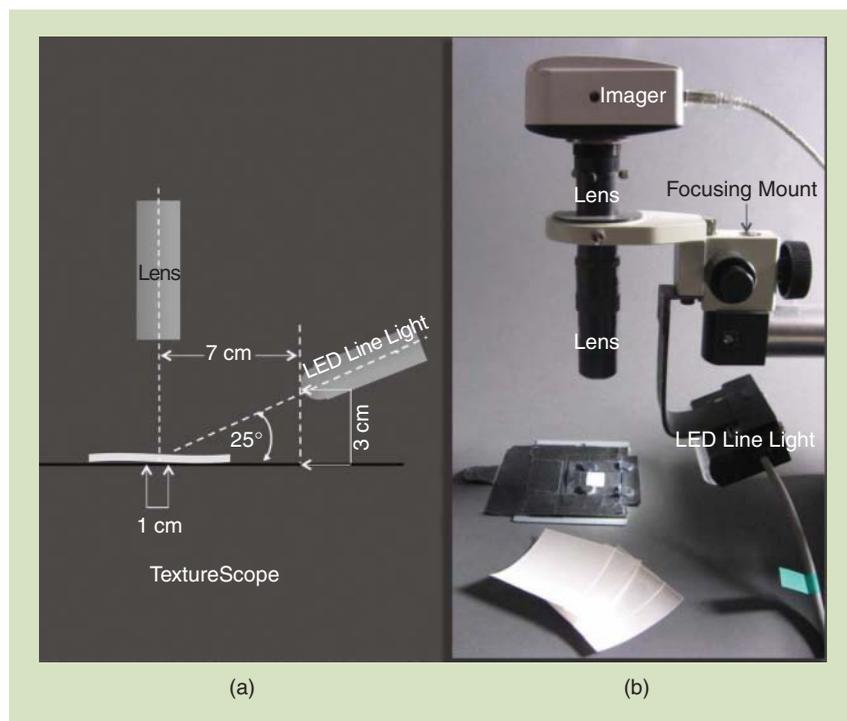
Texture analysis and characterization are long-standing topics of image processing and have been the subject of considerable research efforts over the past decades, cf., e.g., [4]–[8]. Texture characterization has relied on a variety of attributes

(from textons or primitives, i.e., gray-level statistics or geometrical features, to co-occurrence matrices or multiple spatial dependencies) and has proven effective for a wide range of different applications, e.g., in biomedical contexts [9], [10], in physics of surfaces and fractures [11], and in geophysics [12]. To a lesser extent, and only recently, texture analysis has been applied to art investigations (cf. [13]–[22] and the references therein). Among the many paradigms used for texture characterization, fractal and multiscale methods have received growing attention. Fractal analysis has further been extended to multidimensional multifractal analysis, cf., e.g., [17] and [23]. However, in most formulations, (multi)fractal or multiscale analyses do not account for the potential anisotropy of textures. Recently, however, the HWT [24] has been shown to account for anisotropy in the multiscale analysis of textures [25].

DATA SETS

TEXTURE IMAGE ACQUISITION

Presently, the simplest means to catalog surface texture is through images made using magnification and raking light. This imaging system, referred to as the *TextureScope*, has been extensively described in [19] and is shown in Figure 1. It is noncontact and nondestructive and can therefore be easily adapted for use on photographic prints of high intrinsic value. The method is relatively quick and requires minimal specialized handling so that large image sets can be produced rapidly. Created under repeatable and standardized conditions, the resulting images provide an important visual record and serve as a basis for computational analysis. The *TextureScope* depicts $1.00 \times 1.35 \text{ cm}^2$ of a paper



[FIG1] The *TextureScope*.

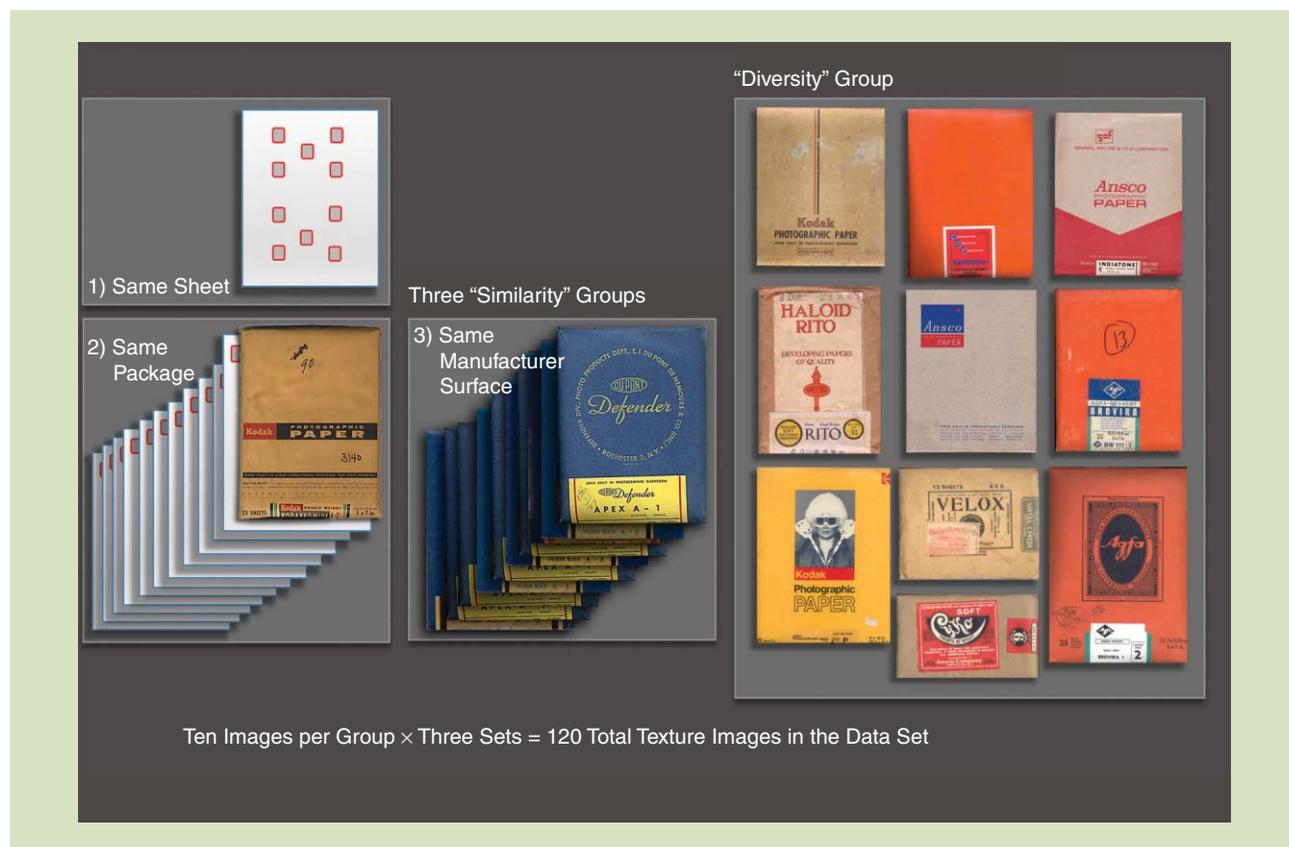
surface. This scale is selected since it reveals some microscopic features, such as matting agents occasionally used by manufacturers, but also depicts attributes recognizable to a human observer. Samples are digitized at 153.6 pixel/mm, resulting in $1,536 \times 2,080$ images, with each pixel thus corresponding to $6.51^2 \approx 42.4 \mu\text{m}^2$.

PHOTOGRAPHIC PAPER TEXTURE DATA SETS

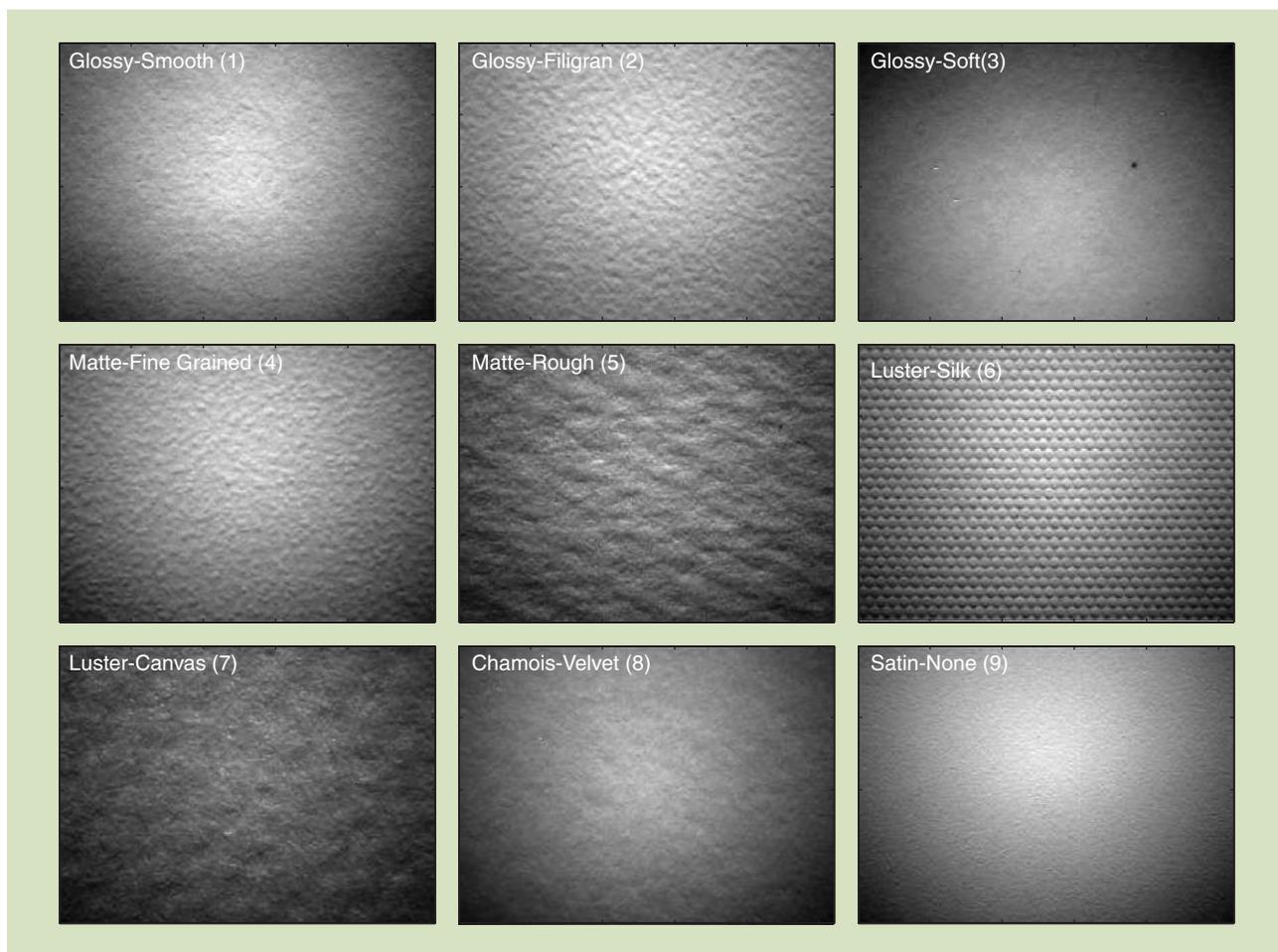
Data set 1 consists of 120 nonprinted photographic paper samples. Aside from the texture images, the data set includes manufacturer-applied semantic descriptions of the samples (manufacturer, brand, texture, and reflectance) and an approximate date of production. Three levels of similarity are built into the data set (cf. Figure 2): 1) samples from one same sheet (three subsets of ten samples each), 2) samples from sheets taken from one same package (three subsets of ten samples each), and 3) samples from papers made to the same manufacturer specifications over a period of time (three subsets of ten samples each). In addition, 30 sheets representing a fuller diversity of photographic paper textures are included. The data set and its documentation have been prepared by an expert familiar with the technical and aesthetic history of photographic paper to include both commonly used surfaces and some outliers. Data set 1, further described in [19], is publicly available within the framework of the Historic Photographic

Paper Classification Challenge (<http://papertextureid.org>) developed by Paul Messier and C.R. Johnson.

Data set 2 gathers 2,491 samples that fall into two subset categories. The first and largest subset (2,031 samples) consists of silver gelatin (traditional black and white paper) surface texture samples that were taken directly from manufacturer packages or sample books spanning the 20th century. These samples are representative of the full range of surface textures available to 20th century photographers and is carefully documented using the same manufacturer-applied metadata described for data set 1. The second subset in data set 2 contains textures from finished photographic prints. Within this group, 346 samples derive from the Thomas Walther collection held by the Museum of Modern Art in New York and contains work by leading modernist photographers primarily active in Central and Eastern Europe between World War I and World War II. This group is joined by a small but important collection of textures from 11 prints belonging to the Museum of Fine Arts, Houston. Each one of the 11 prints are by the same artist and depict the same image as 11 prints from the Walther collection in the Museum of Modern Art. Comparing the textures of these twin prints offers the possibility of determining if they are made on exactly the same, similar, or completely different papers. Discovering a shared material history between the print pairs can have significant ramifications for art historical



[FIG2] Data set 1 with nine groups of ten samples, each representing three levels of similarity (same sheet, same packet, same manufacturer) and 30 samples representing the diversity of art photographic papers.



[FIG3] Photographic paper textures. Examples of raking light photographic paper samples spanning a variety of different reflectance-texture characteristics.

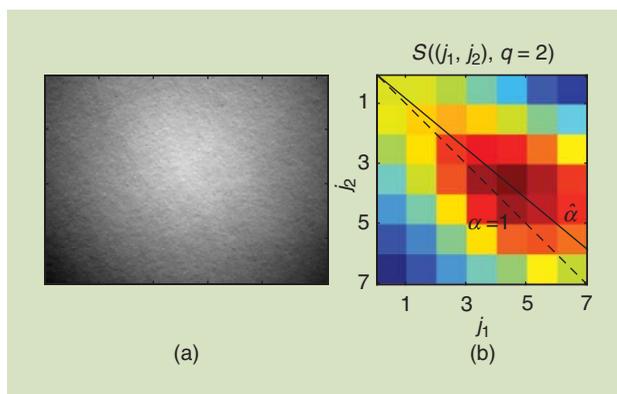
scholarship, especially if this link can be established remotely through data analysis versus often impractical side-by-side comparison. Data set 2 thus offers rich opportunities for cross-referencing discovered affinities and anomalies across time periods, manufacturers, collections, and individual makers. Typical samples from data set 2 are represented in Figure 3, spanning a variety of photographic papers.

METHODOLOGY

A texture-clustering procedure relies on the selection, design, and combination of three key ingredients: 1) features representative of the texture, 2) the distances between features providing relevant measurement of resemblance between pairs of paper surfaces, and 3) the classification procedure.

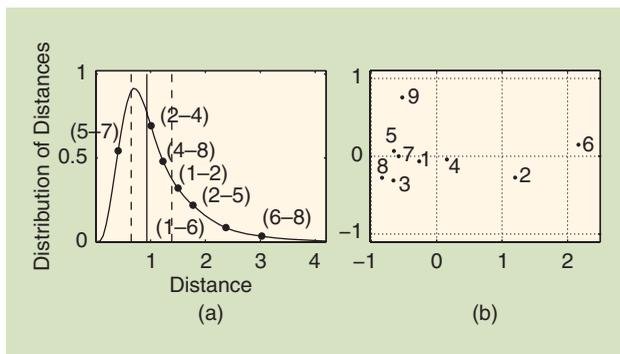
FEATURES: HYPERBOLIC WAVELET TRANSFORM

We propose to extract surface features using the HWT [24]. HWT consists of one of the many variations in image multiscale analysis, that expands on the classical 2-D-discrete wavelet transform (2-D-DWT). HWT explicitly accounts for the potential anisotropy of an image texture, as it relies on the use of two independent dilation factors $a_1 = 2^{j_1}$ and $a_2 = 2^{j_2}$

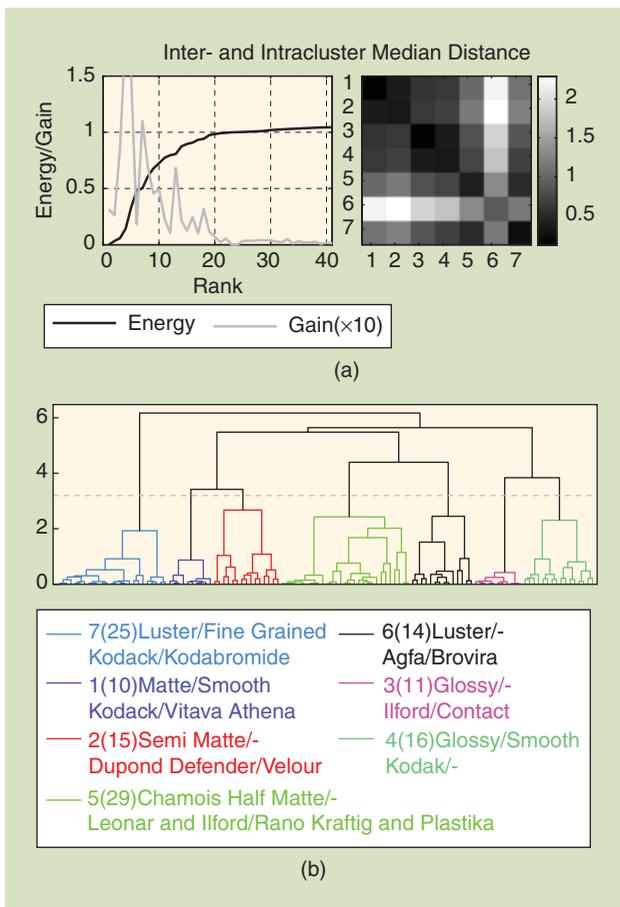


[FIG4] HWT-based features. (a) A surface sample with its (b) HWT-based anisotropic multiscale representation $S_x(j_1, j_2, q = 2)$. The estimated anisotropy angle $\hat{\alpha}_{0,q=2}$ (solid line) indicates a departure from isotropy $\alpha \equiv 1$ (dashed line).

along the horizontal and vertical axes. In [25], HWT is favorably compared against the 2-D-DWT, the former permitting to disentangle actual multiscale properties from the potential anisotropy of the analyzed texture, while the latter yields



[FIG5] The distances between (a) photographic paper textures between all pairs of data set 2, to which are reported distances between some of the samples shown in Figure 3(a), and (b) a virtual 2-D space showing proximity between samples (see the “Methodology” section for more details).



[FIG6] Data set 1: spectral clustering. The left plot in (a) shows the sorted eigenvalues (energy) and successive differences (gain), while the right plot shows the intra- and intercluster median distances. (b) A dendrogram showing seven clusters, with a posteriori interpretations of their contents from the manufacturer-applied metadata: cluster # (cluster size): reflectance/texture, manufacturer/brand.

strongly biased estimates of multiscale properties when textures are anisotropic.

The HWT coefficients of an image $X(x_1, x_2)$ are obtained by comparing it, by means of inner product, against a collection of

dilated [at scales $\mathbf{a} = (2^{j_1}, 2^{j_2})$] and translated [at locations $(2^{j_1}k_1, 2^{j_2}k_2)$] templates

$$\psi_{j_1, j_2, k_1, k_2}(x_1, x_2) = 2^{-(j_1+j_2)/2} \psi_0(2^{-j_1}x_1 - k_1, 2^{-j_2}x_2 - k_2) \quad (1)$$

of a reference mother-wavelet $\psi_0(x_1, x_2)$, $\forall (j_1, j_2) \in \mathbb{R}_+^2$:

$$T_X(j_1, j_2, k_1, k_2) = \langle X(x_1, x_2), \psi_{j_1, j_2, k_1, k_2}(x_1, x_2) \rangle. \quad (2)$$

The mother-wavelet $\psi_0(x_1, x_2)$ is classically defined as a tensor-product of one-dimensional (1-D)-multiresolution mother-wavelets (cf., e.g., [28]). A multiscale representation of $X(x_1, x_2)$ is further obtained by computing space averages (l^q -norms) of the $T_X(j_1, j_2, k_1, k_2)$ at fixed scale pairs (j_1, j_2) , $q > 0$:

$$S_X(j_1, j_2, q) = \frac{1}{n_a} \sum_{k_1, k_2} |T_X(j_1, j_2, k_1, k_2)|^q, \quad (3)$$

where n_a stands for the number of coefficients actually computed and not degraded by image border effects.

It was shown in [25] that the anisotropy of the texture can be quantified by an index $\alpha \in [0, 2]$ ($\alpha = 1$ corresponding to isotropy) and that S_X often behaves as a power law, $S_X(j\hat{\alpha}_q, j(2 - \hat{\alpha}_q), q) \simeq C2^{j\frac{q}{\hat{\alpha}_q} - \hat{H}_q}$, where $\hat{\alpha}_q = \operatorname{argmax}_\alpha \gamma(\alpha, q)$ is an estimate for α , $\hat{H}_q = -\gamma(\alpha, q)/q$ is an estimate for the (anisotropy robust) self-similarity, or Hurst, or fractal, parameter H , with $\gamma(\alpha, q) = \liminf_j \log_2(S_X(\alpha j, j(2 - \alpha j), q))/j$. An example of $S_X(j_1, j_2, q = 2)$ is shown in Figure 4.

To ensure balanced contributions from all scales despite such power law behaviors, features are computed from $\log(S_X(j_1, j_2, q))$, after a normalization across scales that ensures that the features do not depend upon a change in the intensity of the raking light and exposure variables that influence overall image brightness

$$\tilde{S}_X(j_1, j_2, q) = \log(S_X(j_1, j_2, q) / \sum_{j'_1, j'_2} S_X(j'_1, j'_2, q)). \quad (4)$$

In this article, the selected analysis scales are $1 \leq j_1, j_2 \leq 7$ and correspond to physical scales ranging from $13 \mu\text{m} \leq a_1 = 2^{j_1}, a_2 = 2^{j_2} \leq 830 \mu\text{m} \equiv 0.83 \text{ mm}$ (i.e., seven octaves) thus yielding a matrix of $7 \times 7 = 49$ multiscale features for texture characterization.

A cepstral-type distance (i.e., a log-transformed normalized L^p norm) between the multiscale representations $\tilde{S}_X(j_1, j_2, q)$ and $\tilde{S}_Y(j_1, j_2, q)$ is used to quantify proximity between textures X and Y defined as (with $p > 0$):

$$D(X, Y) = \left(\sum_{j_1, j_2} |\tilde{S}_X(j_1, j_2, q) - \tilde{S}_Y(j_1, j_2, q)|^p \right)^{1/p}.$$

In this article, $q = 2$ and $p = 1$ are used, without specific tuning to obtain optimal results. The empirical distribution of the distances computed between all $(2491 \times 2490/2)$ pairs of samples in data set 2 is shown in Figure 5(a). The distances between some pairs among the samples in Figure 3 are superimposed to that distribution, and are also mapped into a

virtual 2-D space [cf. Figure 5(b)] quantifying resemblance and disresemblance between samples (mapping obtained by a standard multidimensional scaling procedure based on a Kruskal's normalized stress criterion).

Clustering is achieved via the spectral clustering procedure (cf., e.g., [26], [27], and [29]), which can be regarded as a specific unsupervised learning technique that aims to ensure robustness of the classification by reducing the dimensionality of the space in which samples are represented. Starting from the $N \times N$ cepstral distance matrix \mathcal{D} , where N is the number of photographic paper samples, the clustering procedure used here operates as follows.

- 1) A nonlinear transformation is applied to distance matrix \mathcal{D} , $\mathcal{A} = \exp(-\mathcal{D}/\epsilon)$ (corresponding here to entry-wise exponentiation), yielding a (dis)similarity matrix, where ϵ is a constant assessing the typical closeness between images.
- 2) The eigenvalues and eigenvectors of the (random walk-type) Laplacian operator $\mathcal{L} = I - D^{-1}\mathcal{A}$ associated to \mathcal{A} are computed, where D is the diagonal matrix $D = \text{diag}(\sum_j \mathcal{A}_{ij})$.
- 3) The eigenvectors corresponding with the K smallest eigenvalues of \mathcal{L} are assembled in the $K \times N$ matrix \mathcal{S} , defining the set of robust K coordinates (hence the reduction of dimensionality, $K \ll N$) for the N samples.

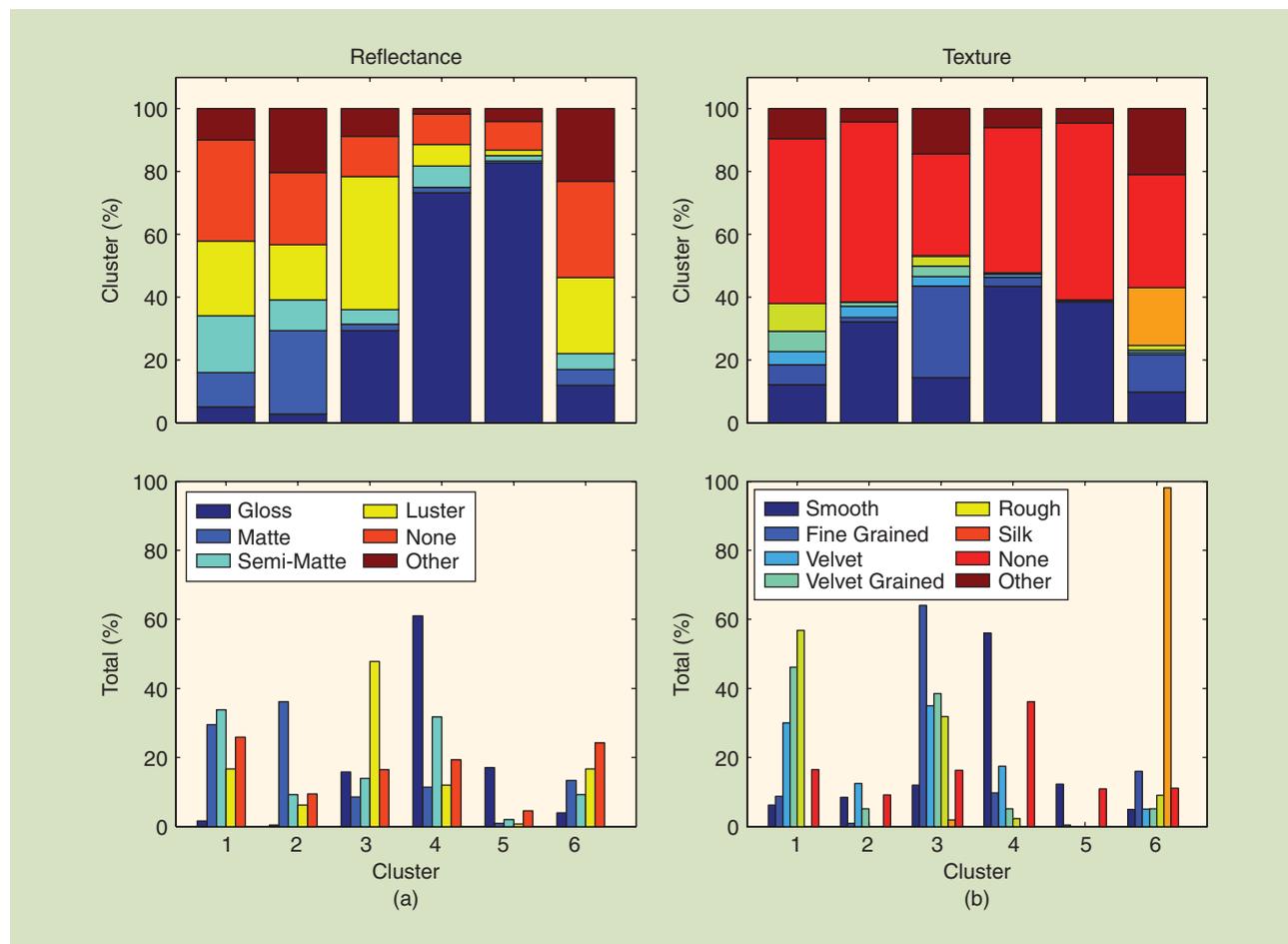
4) Hierarchical ascendant clustering (with Ward linkage) is applied to the matrix \mathcal{S} .

5) A set of thresholds is used to produce $K' \leq K$ hierarchical clusters.

RESULTS

DATA SET 1: TEST DATA SET

The analysis procedure described in the section "Methodology" yields the following results. Sorted eigenvalues of the Laplacian (and successive differences) as shown in the left plot of Figure 6(a) lead to conclude that $K = 13 \ll N = 120$ eigenvectors are sufficient to represent the distances within data set 1. The linkage procedure yields $K' = 7$ clusters that are robust to varying the linkage threshold, as shown on the dendrogram in Figure 6(b). The right plot of Figure 6(a) reports the intra-versus intercluster median distances, showing first the robustness of the achieved clustering (black diagonal squares indicate low intra-cluster median distance) and second the proximity of some clusters (e.g., 1 and 2 or 3 and 4). Inspection of the obtained clusters and comparison with the documentation available for data set 1 leads to the following striking conclusions:



[FIG7] Data set 2: gross clustering. Achieved gross clusters are mostly driven first (a) by reflectance and second (b) by texture.

■ All ten samples from a same sheet—and, similarly, all ten samples from a same package—always fall in a common cluster. The median distances between the ten samples from the same sheet and the ten samples from different sheets of the same packet are found to be of the same order of magnitude, indicating a remarkable reproducibility of the manufacturing process [19], [20].

■ Samples from one given sheet or from one given packet correctly fall within the cluster containing different samples from the same manufacturer and produced to the same specifications of texture and reflectance.

■ From the semantic descriptions applied by the manufacturer, the attribute that mostly drives the clustering is reflectance, e.g., luster, chamois, matte, semimatte, and glossy. For the same reflectance (clusters 1 and 2, 3 and 4, and 6 and 7), the clustering is further refined by the manufacturer-applied terms describing texture, e.g., smooth, grained.

■ The classification of the 30 samples representing the diversity of photo papers is found to be clearly driven by both reflectance and texture.

The contents of $K' = 7$ clusters are summarized on the dendrogram in Figure 6(b) (see also [19] and [20]).

DATA SET 2: LARGE DATA SET

Application of the procedure described in the section “Methodology” to data set 2 leads to the following comments. Inspection of the sorted eigenvalues of the Laplacian and their successive differences shows that the use of $K = 62 \ll N = 2,491$ eigenvectors yields a robust representation of distances within data set 2.

GROSS CLUSTERING

Using classical tools to assess robustness and relevance in selecting the threshold of the linkage procedure applied in this $K = 62$ dimensional space leads to an initial coarse classification into $K' = 6$ large-size clusters ($143 \leq \text{Cluster Size} \leq 702$). Compared to the semantic terms applied by manufacturers to describe gloss and texture, the analysis of these clusters leads to conclusions, reinforcing and enriching those drawn from the analysis of data set 1: clustering is mostly driven by reflectance and then texture, as illustrated in Figure 7. Clusters 4 and 5 correspond to a glossy reflectance, cluster 3 corresponds to a Luster reflectance, and clusters 1 and 2 correspond to matte and semimatte reflectances, with a rough or grained texture for the former and a velvet or smooth texture for the latter. Interestingly, cluster 6 gathers almost all of the unusually patterned



[FIG8] A cluster gathering of Willi Ruge's 1931 parachuting series. (Images courtesy of The Museum of Modern Art, New York).

silk textures (reflectance was not documented for most of these samples). See Figure 3 for representative examples of such reflectances and textures.

REFINED CLUSTERING

By decreasing the threshold of the linkage procedure, an increasing number of small-size clusters are extracted in a hierarchical manner, enabling a more detailed analysis of the data set. In some cases, these clusters have obvious interpretations for art scholars, including curators and conservators. For example, one of these clusters gathers 14 prints from a 1931 parachuting series by Willi Ruge (German, 1882–1961) from the Thomas Walther collection at the Museum of Modern Art (Figure 8). This series depicts groups of sequential exposures made to document several events. Grouping within a single cluster indicates the groups share a materials history and were likely made together. Other clusters suggest more surprising affinities and raise unexpected questions. Notably, a small cluster of three samples contains a platinum print by Alfred Stieglitz (American, 1864–1946) from 1915 and two palladium prints by Edward Weston (American, 1886–1958) from 1924 (Figure 9). Weston met the influential Stieglitz for the first time in 1922 during a short trip to New York. Weston said of this meeting [30, p. 5]: “Stieglitz has not changed my direction, only intensified it—and I am grateful.” This small cluster possibly indicates that Stieglitz’s influence was not simply a matter of artistic encouragement but perhaps also grounded in a deliberate use of the same materials. Given traditional modes of scholarships, the implied linkage between these prints would normally not receive scrutiny since these are different artists, using different imaging metals (platinum versus palladium), separated by a large geographical distance (California and New York) and a time period spanning nine years. However, clustering based on surface texture provokes new questions that might otherwise never be asked: Is the dating of the Stieglitz print secure (it is) or could it have been made later? Was Weston making a conscious effort to emulate Stieglitz even after the passage of nine years? Did manufacturers use essentially the same paper over long periods of time, even after the switch from platinum to palladium imaging metal after World War I? Are the overall warm image tones and especially the low contrast of the Stieglitz and the Weston cloud study attributes of a specific brand of paper? These and other related questions demonstrate how discovery of materials-based affinities can open the door to new modes of study.

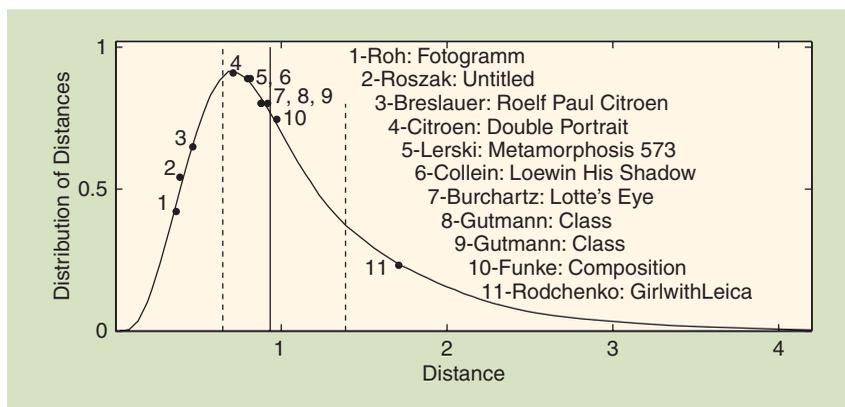
TWIN PRINTS

This promise is further illustrated by data set 2, where 11 pairs of prints, each pair attributed to the same artist and each showing the same or very similar image, are compared across two large museum collections, the Museum of Fine Arts, Houston, and the Museum of Modern Art in New York. The purpose of this comparison is to

determine whether or not the pairs in the different collections are made on the same paper and thus have a shared materials-based history. For this comparison, four fundamental attributes can be compared, print thickness, highlight color, reflectance (or gloss), and texture, of which only the two last may have relations with the quantitative features used here to characterize photo paper surfaces. Figure 10 shows the overall empirical distribution of distances between all $(2491 \times 2490/2)$ pairs in data set 2 as a reference to compare the distances between the twin prints of the 11 pairs of interest here. For examples, the pairs by Franz Roh (German, 1890–1965), Theodore Roszak (American, 1907–1981), and Marianne Breslauer (German, 1909–2001) show small distances (corresponding respectively to the 5, 6,



[FIG9] A surprising cluster gathering: one platinum print by A. Stieglitz from 1915 (left, © 2015 Georgia O’Keeffe Museum/Artists Rights Society, New York) and two palladium prints by E. Weston from 1924 (right) (images courtesy of The Museum of Modern Art, New York).



[FIG10] Distances between twin prints for the 11 pairs of interest, compared to the overall distribution of distances—the solid and dashed vertical lines denote, respectively, the 50% (or median distance) and 25% quantiles.

and 10% quantiles). Interestingly, the examination of print thickness, highlight color, reflectance, and texture had led to the conclusion that prints of these pairs were similar. For prints approaching a distance close to the distribution median distance (the solid vertical line in Figure 10), the similarities are less clear. In all respects (texture, reflectance, highlight color, and print thickness), pairs in this range by Max Burchartz (German, 1887–1961), John Gutmann (American, 1905–1998), and Jaromir Funke (Czech, 1896–1945) are classified as being on the same paper. However, pairs in this same range by Edmund Collein (German, 1906–1992), Helmar Lerski (Swiss, 1871–1956), and another Gutmann print are categorized as being on different paper when gloss, highlight color, and paper thickness are taken into account. In particular, the prints by Paul Citroen (Dutch, 1896–1983), though they have low texture differences (31% quantile), are classified as being on different paper mainly due to significant difference in gloss and print thickness. These results are not surprising given the common manufacturer practice of applying the same texture to different papers. These results suggest that an automated solution for discriminating material-based affinities across collections cannot rely on a single criterion, such as texture or reflectance, for determining results especially for distances around the median of the distribution.

Furthermore, such solutions should convey some level of confidence and context rather than returning a simple, binary determination of same or different. For distances significantly larger than the distribution median, the impact of color, thickness, and particularly gloss is still not clear. A pair by Alexander Rodchenko (Russian, 1891–1956) has a very large texture classification distance (beyond the 84% quantile) but could not be classified for this study due to missing information on gloss, highlight color, and print thickness. An interesting next step for prints such as those with large texture distances would be to determine whether additional information on, e.g., gloss or highlight color, have any impact on the determination of same or different or if instead the very large texture differences in this range are solely determinative.

CONCLUSIONS AND PERSPECTIVES

This article has quantitatively and qualitatively illustrated the potential value of basing the surface characterization and classification of photographic prints holding cultural value on anisotropic multiscale representations (HWT), combined with cepstral type distances and spectral clustering. The test data set 1 assembled in the framework of the Historic Photo Paper Classification Challenge demonstrates that the manufacturer-applied features of reflectance followed by texture are fundamental for the characterization of paper surfaces. Applied to data set 2, this methodology has promising results for art scholars, both by confirming existing conclusions and provoking new questions.

On the methodological side, this work can be expanded along several directions. There is an obvious need to compare the achieved results against those obtained with features both computed on other multiscale representations (e.g., [21]) or based on

representations of very different natures (cf. [19] for a first attempt). Devising distances that better match the perceptual, artistic, aesthetic and manufactured nature of photographic paper is a clear next step and an emergent priority. Also, at the clustering stages, tools for developing more robust assessments of both the relevance of a given cluster (compared to the benefit of further splitting it) as well as the confidence with which a given paper can be assigned to a given cluster, would significantly complement this first work. Additional strong benefits would likely result from using tools aiming at assessing the relevant levels at which data sets should be clustered, in the spirit of multiscale community clustering developed in, e.g., [29].

On the application side, texture characterization, comparison, classification, and retrieval systems have great potential within the humanities, promoting new modes of scholarship for curators, collectors, and art conservators. Dating, authentication, stylistic development, artistic intention, and spheres of artistic influence are vital scholarly questions. Networked and deployed across museum, library, and archive collections, methods for texture query and retrieval can lead to new research opportunities where a print in one location can be meaningfully compared, based on physical attributes, to others held elsewhere. Such systems will provide the means to discover material-based (not simply image-based) affinities across time and within and across artists' oeuvres. Intrinsically valuable for enhanced scholarship in the humanities, such systems also would be effective for identifying anomalies, notably including fakes.

A key future step lies in further developing tools permitting deeper and more meaningful interactions between signal and image processing experts and scholars working within a wide array of humanities-based disciplines. Besides simply making lists of clusters available to humanities-based experts, such tools must enable them to naturally apprehend the robustness of proposed clustering (or its fragility), its hierarchical nature as well as its sensitivity to methodological choices. Developed further, such systems would optimally allow for qualitative input and the modeling of results based on other bodies of nonempirical knowledge. This work would present exciting opportunities where the fields of signal and image processing can adapt to the subtleties and specificities of humanities disciplines, thus broadening applications and cross-disciplinary relevance.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of the following art conservators, who played a critical role in gathering data for this study: Hanako Murata, associate conservator of photographs for the Walther Collection research project, Museum of Modern Art; Toshiaki Koseki, Carol Crow Conservator of Photographs at the Museum of Fine Arts, Houston; and Jennifer McGlinchey, conservator of photographs with Paul Messier LLC. We also gratefully acknowledge C. Richard Johnson for founding the Historic Photographic Paper Classification Challenge and inspiring this article. This work is supported by the French ANR BLANC 2011 AMATIS BS0101102 grant and U.S. National Science Foundation CCF-1319458 grant.

AUTHORS

Patrice Abry (patrice.abry@ens-lyon.fr) is a senior researcher with the CNRS at Ecole Normale Supérieure de Lyon, France. He is developing a long-standing research program on the theoretical modeling and analysis of scale invariance and applications.

Stéphane G. Roux (stephane.roux@ens-lyon.fr) is an assistant professor at Ecole Normale Supérieure de Lyon, France, where he is a specialist in wavelet analysis and their practical application in physics, geophysics, and biology.

Herwig Wendt (herwig.wendt@irit.fr) is an associate researcher with the CNRS. His main focus is on scale invariance and multifractal analysis, particularly for images, and statistical estimation.

Paul Messier (pm@paulmessier.com) is the head of the Lens Media Lab, Yale University, New Haven, Connecticut, and consults internationally with individual collectors and institutions.

Andrew G. Klein (andy.klein@wwu.edu) is an assistant professor of engineering and design at Western Washington University, where he studies signal processing and distributed communication systems.

Nicolas Tremblay (nicolas.tremblay@ens-lyon.fr) is a Ph.D. student at Ecole Normale Supérieure de Lyon, France, and is interested in signal processing over networks and network-based classification techniques.

Pierre Borgnat (pierre.borgnat@ens-lyon.fr) is an associate researcher with CNRS at Ecole Normale Supérieure de Lyon, France, and is working on nonstationary and multiscale signal processing as well as graph signal processing, developing applications for complex systems and networks.

Stéphane Jaffard (stephane.jaffard@u-pec.fr) is a professor of mathematics at Université Paris-Est Creteil, France. He is working on wavelet and multifractal analysis. He is a former president of the French Mathematical Society.

Béatrice Vedel (beatrice.vedel@univ-ubs.fr) is an associate professor of applied mathematics at Université de Bretagne Sud, France. She is working on multifractal analysis, wavelets analysis, and function spaces theory.

Jim Coddington (jim_coddington@moma.org) is the chief conservator of the Museum of Modern Art in New York. He has published articles on imaging, structural restoration, and studies of modern painters.

Lee Ann Daffner (leeann_daffner@moma.org) is the Andrew W. Mellon Conservator of Photographs and is responsible for the preservation, conservation, and research initiatives of the Museum of Modern Art's photography collection.

REFERENCES

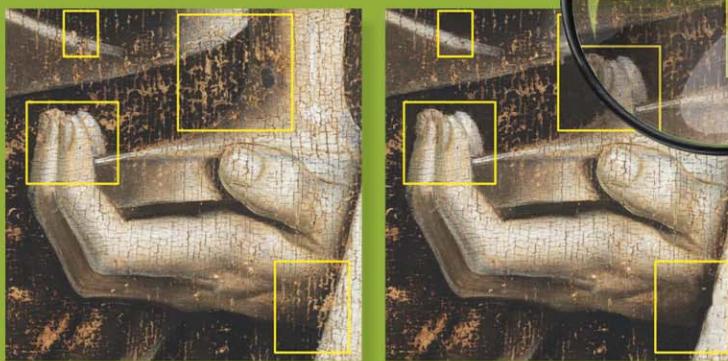
- [1] P. Messier. (2013). The Paul Messier black and white photographic papers collection. [Online]. Available: http://paulmessier.com/pm/pdf/papers/historic_photographic_papers_collection.pdf
- [2] P. Messier, C. R. Johnson, H. Wilhelm, W. A. Sethares, A. G. Klein, P. Abry, S. Jaffard, H. Wendt, S. G. Roux, N. Pustelnik, N. van Noord, L. van der Maaten, and E. Postma, "Automated surface texture classification of inkjet and photographic media," in *Proc. Int. Conf. Digital Printing Technologies*, Seattle, WA, 2013, pp. 85–91.
- [3] P. Messier, "Photographic papers in the 20th century: Methodologies for research, authentication and dating," in *Proc. FotoConservación*, Logroño, Spain, 2011.

- [4] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, 1979.
- [5] M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Vision Graph.*, vol. 4, no. 2, pp. 172–179, 1975.
- [6] K. Deguchi, "Two-dimensional auto-regressive model for analysis and synthesis of gray-level textures," in *Proc. Int. Symp. Science on Form*, 1986, pp. 441–449.
- [7] Y. Stitou, F. Turcu, Y. Berthoumieu, and M. Najim, "Three-dimensional textured image blocks model based on Wold decomposition," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3247–3261, July 2007.
- [8] A. M. Atto, Y. Berthoumieu, and P. Bolon, "2-dimensional wavelet packet spectrum for texture analysis," *IEEE Trans. Image Processing*, vol. 22, no. 22, pp. 2495–2500, June 2013.
- [9] P. Kestener, J. Lina, P. Saint-Jean, and A. Arneodo, "Wavelet-based multifractal formalism to assist in diagnosis in digitized mammograms," *Image Anal. Stereol.*, vol. 20, no. 3, pp. 169–175, 2004.
- [10] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: A review," *Med. Image Anal.*, vol. 13, no. 4, pp. 634–649, 2009.
- [11] N. Mallick, P. P. Cortet, S. Santucci, S. G. Roux, and L. Vanel, "Discrepancy between subcritical and fast rupture roughness: A cumulant analysis," *Phys. Rev. Lett.*, vol. 98, p. 255502, 2007.
- [12] S. G. Roux, A. Arneodo, and N. Decoster, "A wavelet-based method for multifractal image analysis. III. Applications to high-resolution satellite images of cloud structure," *Eur. Phys. J. B*, vol. 15, no. 4, pp. 765–786, 2000.
- [13] P. Abry, J. Coddington, I. Daubechies, E. Hendriks, S. Hughes, R. Johnson, and E. Postma, "Special issue on image processing for digital art work guest editor's forewords," *Signal Process.* (special issue), vol. 93, pp. 525–526, 2013.
- [14] M. Barni, J.-A. Beraldin, C. Lahanier, and A. Piva, "Signal processing in visual cultural heritage," *IEEE Signal Process. Mag.*, vol. 25, no. 4, pp. 10–13, 2008.
- [15] C. Benhamou, S. Poupon, E. Lespessailles, S. Loiseau, R. Jennane, V. Siroux, W. J. Ohley, and L. Pothuaud, "Fractal analysis of radiographic trabecular bone texture and bone mineral density: two complementary parameters related to osteoporotic fractures," *J. Bone Miner. Res.*, vol. 16, no. 4, pp. 697–704, 2001.
- [16] D. Rockmore, J. Coddington, J. Elton, and Y. Wang, "Multifractal analysis for Jackson Pollock," *Proc. SPIE*, pp. 6810–6813, 2008.
- [17] P. Abry, H. Wendt, and S. Jaffard, "When Van Gogh meets Mandelbrot: Multifractal classification of painting's texture," *Signal Process.* (special issue), vol. 93, no. 3, pp. 554–572, 2013.
- [18] P. Abry, S. Jaffard, and H. Wendt, "Bruegel's drawings under the multifractal microscope," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Kyoto, Japan, 2012, pp. 3909–3912.
- [19] C. R. Johnson, P. Messier, W. A. Sethares, A. G. Klein, C. Brown, P. Klausmeyer, P. Abry, S. Jaffard, H. Wendt, S. G. Roux, N. Pustelnik, N. van Noord, L. van der Maaten, E. Postma, J. Coddington, L. A. Daffner, H. Murata, H. Wilhelm, S. Wood, and M. Messier, "Pursuing automated classification of historic photographic papers from raking light photomicrographs," *J. Amer. Inst. Conserv.*, vol. 53, no. 3, pp. 159–170, 2014.
- [20] P. Abry, S. Roux, H. Wendt, and S. Jaffard, "Hyperbolic wavelet transform for photographic paper texture characterization," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Nov. 2014.
- [21] A. G. Klein, A. Do, C. A. Brown, and P. Klausmeyer, "Classification via areal-scale analysis of raking light images," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Nov. 2014.
- [22] W. A. Sethares, A. Ingle, T. Krc, and S. Wood, "Eigentextures: An SVD approach to automated paper classification," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Nov. 2014.
- [23] H. Wendt, S. G. Roux, P. Abry, and S. Jaffard, "Wavelet leaders and bootstrap for multifractal analysis of images," *Signal Process.*, vol. 89, pp. 1100–1114, 2009.
- [24] R. A. DeVore, S. V. Konyagin, and V. N. Temlyakov, "Hyperbolic wavelet approximation," *Construct. Approx.*, vol. 14, no. 1, pp. 1–26, 1998.
- [25] S. G. Roux, M. Clausel, B. Vedel, S. Jaffard, and P. Abry, "Self-similar anisotropic texture analysis: The hyperbolic wavelet transform contribution," *IEEE Trans. Image Processing*, vol. 22, no. 11, pp. 4353–4363, 2013.
- [26] U. Von Luxburg, "A tutorial on spectral clustering," *Stat. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.
- [27] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Springer Series in Statistics). New York: Springer, 2001.
- [28] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press, 1998.
- [29] N. Tremblay and P. Borgnat, "Graph wavelets for multiscale community mining," *IEEE Trans. Signal Processing*, vol. 62, no. 20, pp. 5227–5239, 2014.
- [30] E. Weston, *The Daybooks of Edward Weston*, N. Newhall, Ed. New York: Aperture, 1973.

[C. Richard Johnson, Jr., William A. Sethares, Margaret Holben Ellis, and Saira Haqqi]

Hunting for Paper Moldmates Among Rembrandt's Prints

© GHENT, KATHEDRALE KERKFABRIEK. PHOTO COURTESY OF KIK-IRPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

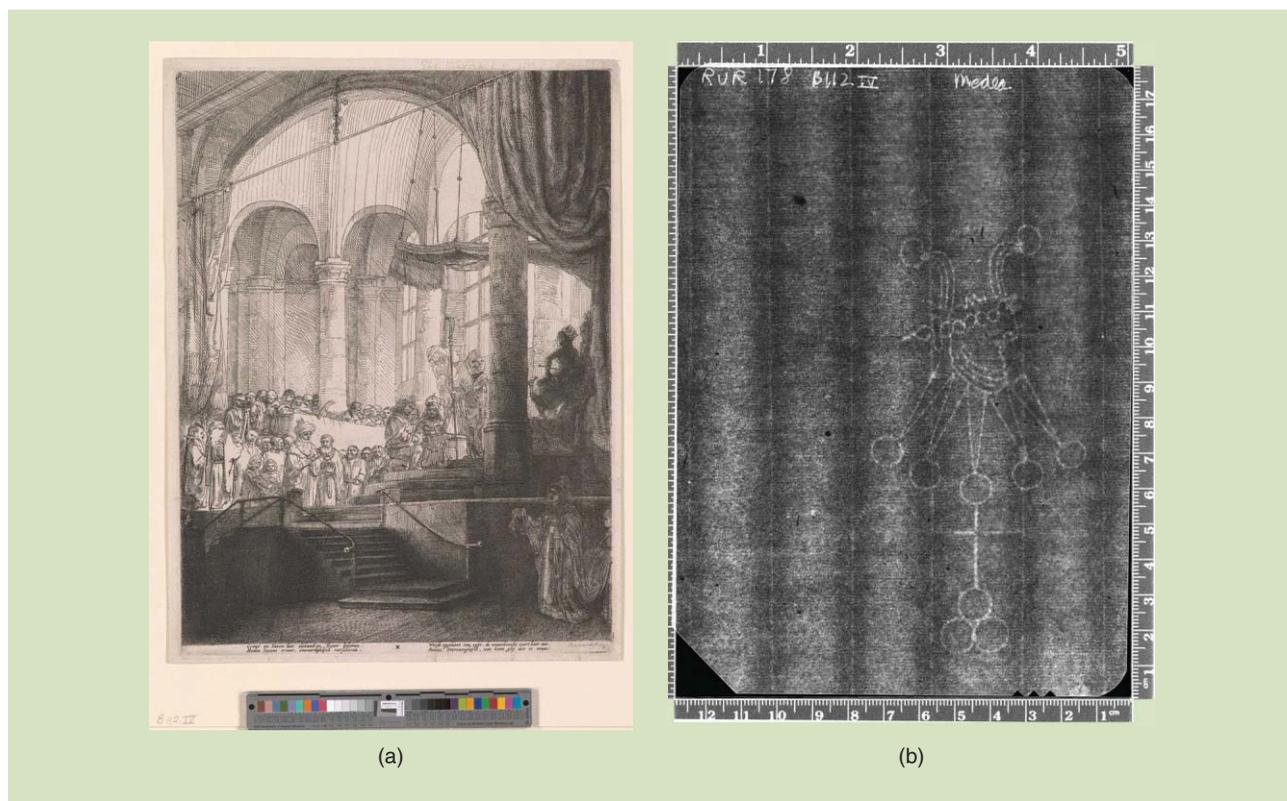
[Chain-line pattern matching]

Early paper manufacturing used sieve-like molds through which paper pulp was drained. Two pieces of paper are called *moldmates* if they were made using the same mold. When a large body of one artist's work on paper exists, the identification of moldmates can help in establishing chronology, suggest paper preferences, and indicate periods of intense activity of the artist. Rembrandt is an especially good example. With several thousand prints in existence today, the study of Rembrandt's prints has occupied scholars for over two centuries,

and the study of his printing papers occupies a prominent place within this scholarship [1]. This article examines the feasibility of moldmate identification via chain-line pattern matching and conducts a series of experiments that demonstrate how accurately the measurements can be made, how straight and parallel the lines may be, and provides a rule of thumb for the number of chain lines required for accurate moldmate identification using a simplified model. The problem of identifying moldmates among Rembrandt's prints is presented as a pair of image/signal processing tasks; our strategy is to provide basic solutions to these tasks and to then reveal the shortcomings of these solutions in the hopes of encouraging future work in the signal processing community. With the support of the Morgan Library & Museum and the Metropolitan

Digital Object Identifier 10.1109/MSP.2015.2404931

Date of publication: 15 June 2015



[FIG1] Rembrandt's etching *Medea, or the Marriage of Jason and Creusa*, (Bartsch 112iv) is shown in (a) visible light and as rendered, in a detail, by (b) beta-radiograph. The area near the watermark is shown in the beta-radiograph where the chain lines are the vertical features that are spaced a little less than 1 in apart (see the top ruler). Thanks are extended to Reba Snyder for providing the beta-radiograph of B112. (Etching photo credit: The Morgan Library & Museum, New York. RvR 178. Photography by Graham S. Haber, 2014.)

Museum of Art, both of which are in New York, we have made high-resolution data available [2] to facilitate this quest.

LAID LINES, CHAIN LINES, MOLDMATES, AND WATERMARKS

Before the introduction of the papermaking machine in the early 19th century, handmade paper was created by dipping a *mold*—a porous screen surrounded by a removable wooden frame—into a vat of mascerated and suspended pulp. Prior to 1750, the screen was fabricated from fine, densely spaced horizontal rows of laid wires held into position by thicker, widely spaced vertical chain wires. When the mold was plunged into the vat and lifted out, these wires acted as a sieve, retaining the paper pulp in thinner and thicker accumulations as the water drained away. A modern reenactment of this method of papermaking can be found online [3]. The grid-like configuration—called a *laid and chain-line pattern*—formed by the interference of the wires, is replicated in the final sheet and is easily visible when the blank paper is viewed via transmitted light or imaged via radiography. For example, Figure 1 shows one of Rembrandt's prints and a beta-radiograph of the same print in the region around the watermark. The chain lines are the whitish vertical features spaced slightly more than 2 cm apart. The laid lines are the closely spaced (and barely perceptible) horizontal lines. Each mold was made by hand and, while at first glance, two molds may appear to be identical, small

variations exist between the exact intervals of chains from one mold to the next. Papers having identical laid and chain-line patterns are commonly presumed only to occur if they have been formed on the same mold—hence, they are called *moldmates*.

A watermark designating the paper's manufacturer, size, or function was often stitched onto the heavier chain lines using a wire bent into a simple shape, for example, a star, shield, or monogram. Just like the chain and laid lines, the watermark wire affected the rate and quantity of pulp as it drained through the mold and left behind a characteristic impression on the paper. Typically, images of watermarks are captured for art historical research using beta-radiography or low-energy X-radiography [1]. For example, the watermark in Figure 1 is known as a *foolscap* and depicts the head of a jester wearing a five-pointed ruff surmounting a cross emerging from three spheres. Twenty-one variants of foolscap watermarks found on Rembrandt's prints are cataloged in [1], one of which has nine subvariants, with one of these appearing in Figure 1. This particular subvariant can be found in 16 different prints by Rembrandt according to [1]. Papers having identical watermarks may also be presumed to be moldmates.

Watermarks have received considerable attention in print, drawing, book, and manuscript connoisseurship, with scholars carefully recording watermarks in the works of Shakespeare, Beethoven, Jane Austin, Dürer, and other artists and writers. To

date, moldmates have been identified primarily by comparing watermarks. However, extracting and comparing watermarks using computer-generated algorithms is a challenging proposition, as they have complicated shapes (as shown by the foolscap) and can assume many forms and have many variations. Another drawback to the use of watermarks for moldmate identification is that not all prints under consideration have watermarks. Indeed, only approximately one-third of the extant impressions of Rembrandt's etchings include a full watermark or even a fragment [1]. Accordingly, the use of the chain-line pattern has been suggested as way to identify papers made on the same mold [4], [5]. The difficulty in manually pursuing moldmate identification via chain-line pattern matching has prompted the consideration of its automation [6]–[8], though no automated scheme has yet been universally adopted.

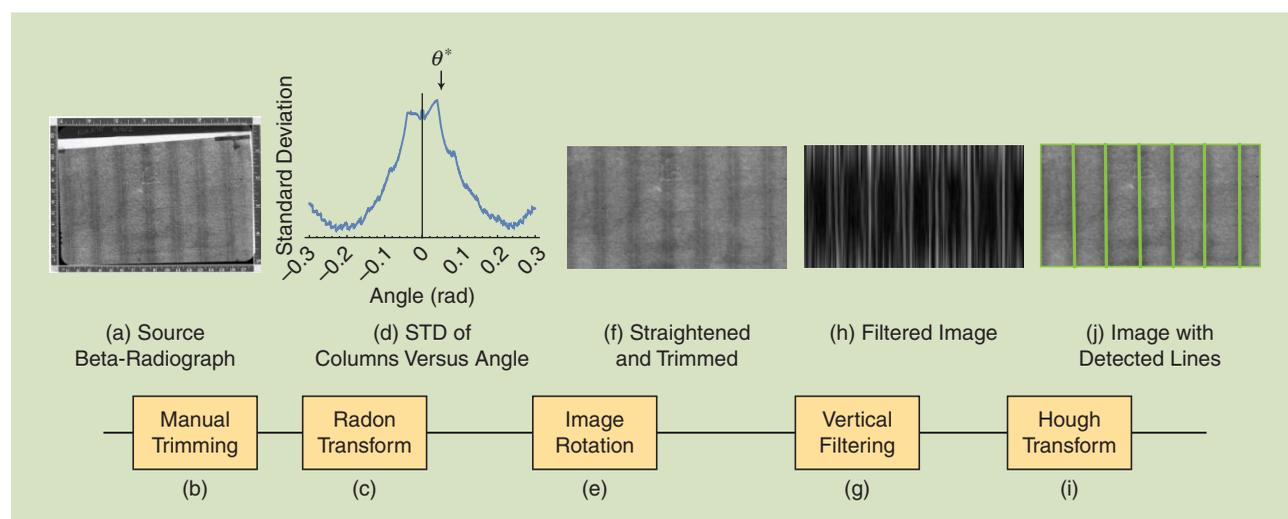
CHAIN LINE PATTERN MATCHING AND SIGNAL PROCESSING

The approach taken here is to characterize the problem of identifying moldmates among Rembrandt's prints—and, by extension, any handmade laid and chain-line patterned papers such as those found in manuscripts, printed books, archives, and collections of prints and drawings—as a basic pair of image processing tasks. (While considerable progress has been made in the application of signal processing to the technical analysis of canvas supports for paintings [9]–[13], there has been much less focus on the application of signal processing to handmade laid paper.) The first task is to locate the chain lines in images such as Figure 1; the second is to use the chain-line patterns to identify potential moldmates. Our strategy is to provide basic solutions to these two tasks that are sufficient to convince the user community to collaborate in a future imaging campaign providing full sheet images to maximize the extent of the chain-line pattern associated with each piece of paper. To help encourage members of the signal processing community to contribute to this topic of paper moldmate identification, we point out here some of the limitations of our basic

solutions and note that, with the support of the Morgan Library & Museum and the Metropolitan Museum of Art, we have made high-resolution images available [2], which will be needed for developing improvements.

Our algorithmic approach to this computational art history task is similar to the strategy in [7], which uses image processing tools (in particular Fourier and Radon transforms) to extract from a suitable image the average chain-line separation distance, chain-line orientation, number of chain lines, and the sequence of chain-line separation distances. As stated in [7]: “All these features are detected under the assumption that the lines are straight and equidistant with respect to each other.” (Here *equidistant* means *parallel*.) In [7], the average chain spacing and the chain-space sequence are combined with the laid line density in a similarity measure. Our approach in this article differs in that we drop the extraction and use of laid line density, as most of our images are of insufficient clarity to assess this feature. Plus, we abandon the assumption that the chain lines are parallel; while we retain the assumption that they are straight. The lead author of [7] modified the straightness assumption to piecewise straightness in [8], with a commensurate increase in the necessary signal processing.

In this article, we test the straightness of the chain lines in the images in our data set and find that numerically adequate near-straightness occurs in approximately 90% of our images. Thus, in our quest to establish the ability of chain-line pattern matching of a sufficient number of adjacent chain spaces by itself (to offer a reduction in a broad library to a manageable number of potential matches for further investigation), we assume chain-line straightness. We also test our data set for the occurrence of chain-line patterns with at least one nonparallel chain line with a relative angle greater than 0.12° . This is a sizable portion of our data set. We chose to include the possibility of handling images with nonparallel chain lines, for which the location of the measurement of chain spacing on the image relative to its location on the mold needs to be computed, because of the fundamental simplicity of



[FIG2] (a)–(j) A semiautomatic method of chain-line extraction.

this task. The end result is that we agree with [7] in that the chain-space sequence is “the most discriminative feature.”

LOCATING CHAIN LINES

A semiautomatic method of chain-line extraction in beta-radiographs is shown in Figure 2. This begins in (a) with a manual trimming of (b) to remove any labels, ruler markings, and blank areas from the source image. The trimmed image is input into a Radon transform (c) and the angle at which the chain lines are (closest to) vertical is given by the column with the greatest variation. The graph (d) shows the standard deviation of the elements in each column, and the angle θ^* with the maximum value gives the angle of rotation. In (e), the trimmed image is rotated by $-\theta^*$ to give the straightened image (f). This is then filtered using an aggressive vertical filter (g), which is $1/3$ the image height and three pixels wide, resulting in the filtered image (h). A Hough transform (i) then locates the most prominent lines, and the result is shown in (j), superimposed over the straightened image.

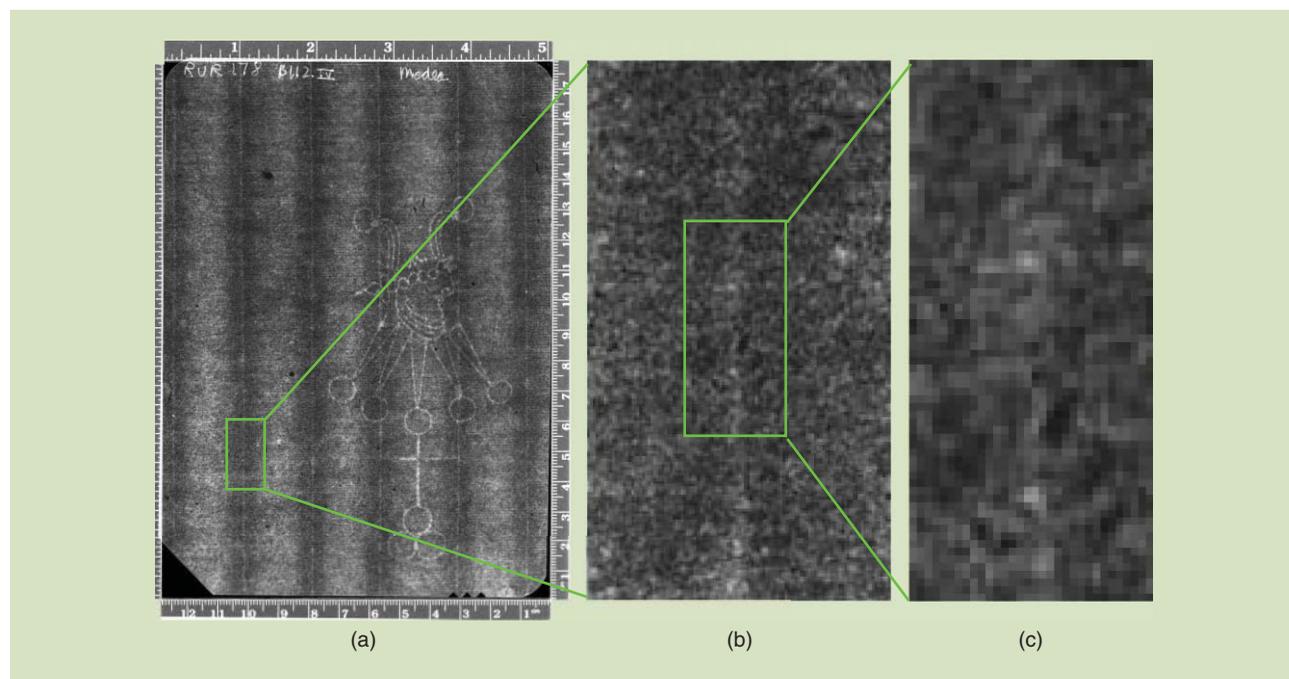
Although the positions and orientations of the lines can be determined automatically in this procedure, both the preparation of the image and the final verification are done manually. It is difficult to control the exposure of beta-radiographs, since under- and oversaturation can occur based on the properties of the paper, which may not be fully known at the time of exposure. As a result, it is common to adjust the contrast on the images and to trim unwanted portions of the raw data. The images commonly contain a ruler [as in Figures 1 and 2(a)], which is needed to verify physical dimensions, and there may be regions of all-black or all-white [such as the wedge-shaped region in Figure 2(a)] that need

to be cropped. While it may be possible to automate some of these actions, mistakes in the preprocessing almost guarantee that the chain lines will not be found properly.

Once the procedure (b)–(c)–(e)–(g)–(i) of Figure 2 has been run, it is necessary to check that the output is sensible. Typical errors in the algorithm would include failing to find some of the lines or finding too many lines. In either case, the algorithm can be rerun with different thresholds and settings within the filters and transforms. Clearly, the method suggested in Figure 2 is but one possibility. See [7] for another method that begins by exploiting the shadow around the chain lines rather than the chain line itself.

As all approaches will, our method makes several assumptions about the chain-line patterns. First, it assumes that all the lines are more or less parallel (because otherwise the Radon transform cannot locate a single preferred direction for the step (e) derotation). Second, it assumes that all the lines are genuinely straight, since curved or segmented lines cannot be reliably located by the Hough transform. The open problem of finding more automated, more general, more accurate, and/or faster algorithms is one task we are promoting here to signal processing specialists.

Rather than pursuing an improved or more general algorithm here, consider the question of accuracy. It is easy to look at the superimposed output of Figure 2(j) and to see that it “looks” correct. It has detected the correct number of lines, and they are located plausibly. But what is the “real” answer, and how close does the algorithm come to this answer? Indeed, this becomes crucial when comparing different algorithms, since only by comparing with a ground truth can the accuracy of different candidate algorithms be compared.



[FIG3] A small region containing a portion of a chain line is extracted from the beta-radiograph in Figure 1. This is then enlarged twice in (b) and (c). What appears to be a fairly clear vertical line in (a) dissolves into a blur of pixels in (c). Where exactly is the “line?”

EXPERIMENT 1: A TEST OF ACCURACY

To see why the task of locating the chain lines may not be as straightforward as it might seem, the beta-radiograph B112iv is enlarged twice; see Figure 3. While (a) and (b) clearly show the chain lines as vertical linear features, (c) shows that these are not simple geometric lines. Rather, they are composed of uneven and blurry patches, they are not of uniform width (often extending several pixels to either side), and they may be interrupted repeatedly.

To investigate, we asked a group of Cornell University engineering students taking a course on signal processing algorithms for analyzing art supports to manually identify the chain lines in several beta-radiographs. The subjects were given marking software (a simple graphical user interface built in Mathematica), which allowed placement of marks near the endpoints of the chain lines. A semitransparent line connected the two end points so that the subjects could visually verify his or her marking. Such manual markings can be used to establish the ground truth of “where the lines really are” and can also be used to assess the agreement (or disagreement) among the subjects.

Since the subjects could choose where to mark, it was not possible to compare the marked locations themselves. Rather, the lines formed by joining the two marked end points were parameterized in distance/angle form

$$\rho = x_i \cos(\theta) + y_i \sin(\theta), \tag{1}$$

where the (x_i, y_i) are the Cartesian coordinates of the two end points, ρ is the distance from the origin and θ is the angle of the line. With two marked points, (1) is a system of two equations and two unknowns, and hence can be solved for (ρ, θ) . The (ρ, θ) values for each of the lines marked by each of the subjects were tabulated. Altogether, there were 12 subjects who worked with

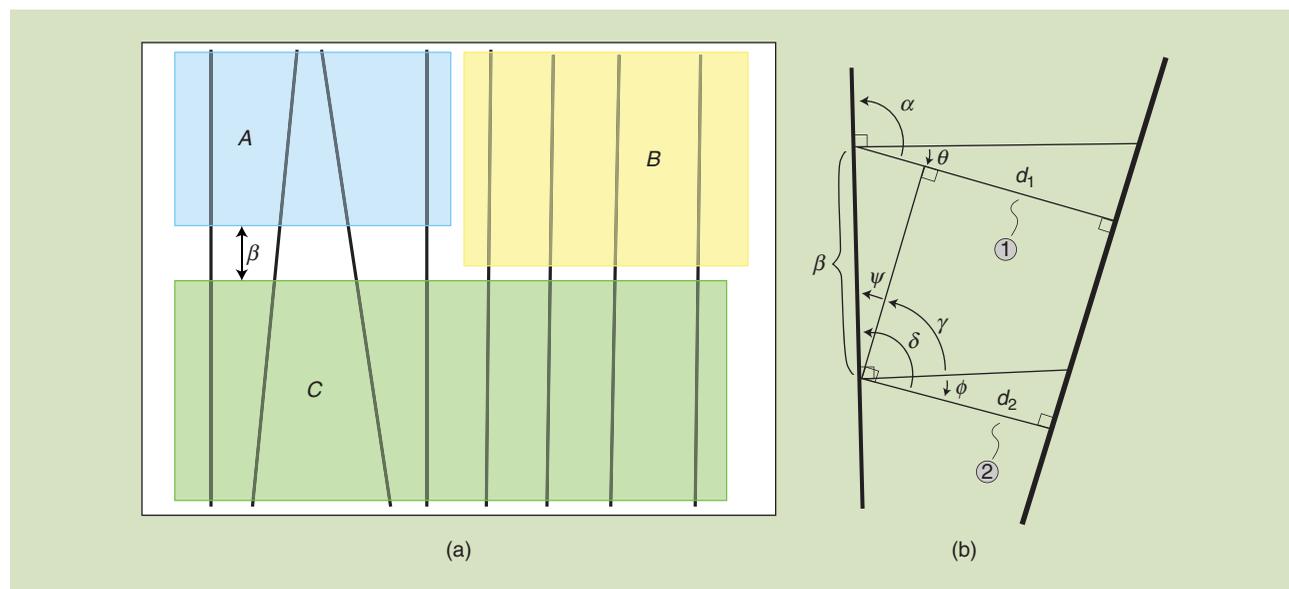
24 beta-radiographs chosen randomly from our larger set. Each subject measured four to five lines on eight different beta-radiographs and so each chain line was measured independently four times. Over this data set, the mean of the ρ values was 1.06 and the standard deviation was

$$\sigma_e = 0.009, \tag{2}$$

which corresponds to about 5.3 pixels at a nominal resolution of 600 dots per inch (dpi). This shows that despite the patchy nature of the chain lines, subjects can locate the chain lines with reasonable consistency. It also provides a way to assess the quality of an algorithm. Indeed, applying the semiautomatic procedure of Figure 2 to the same beta-radiographs gives values of ρ that are within one standard deviation of the measured values.

EXPERIMENT 2: A TEST OF STRAIGHTNESS

Both the algorithm of Figure 2 and the discussion of the geometry of moldmates in Figure 4 presume that the chain lines are straight. Staalduinen [8] observed that some chain lines may be bent and developed a method that attempted to locate the piecewise segments of such bent lines. Our observation was that the majority of chain lines did not appear to bend, and so we wanted to quantify the straightness of the chain lines. Toward this end, we asked the subjects to manually find the smallest bounding box that encloses the centered spine of each chain line. The subjects and image data were the same as in Experiment 1. The subjects now used the software to mark three points for each chain line, from which the bounding boxes can be calculated. These three points were to be placed in the horizontal center of the vertically oriented chain line with one each near either end and one at some point of extreme departure from the straight line connecting the marked end points.



[FIG4] (a) A paper mold is shown in schematic form where the nonparallel chain lines have been exaggerated. The paper has been cut into three pieces labeled A, B, and C. Given a large collection of papers, the goal of moldmate identification is to find those papers that come from the same mold. The geometry of two nonparallel chain lines is shown in (b).

Again, each bounding box is measured four times, by four different subjects. Figure 5 shows a histogram of the widths of the chain line bounding boxes. The average width is 7.0 pixels, and the median is 5.3. Both of these numbers are comparable to (2), which is the accuracy to which the measurements are made—hence, over half of the chain lines have width smaller than can be reliably measured.

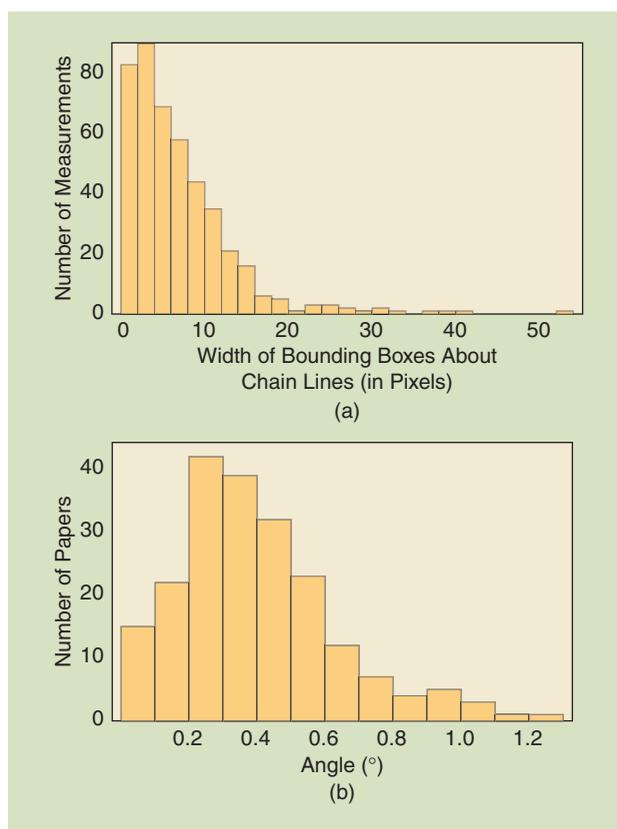
On the other hand, there are isolated chain lines that have bounding boxes with nontrivial width. In this data set [2], the largest is the left-most chain line in the image 17.37.75, which is measured as 52.7, 27.48, 29.83, and 30.97 pixels by the four subjects. While these values are not particularly consistent, they are all well above the measurement error, indicating a significantly bent chain line. The second-largest values are from the left-most chain line of image 17.37.75 (2), which were reported as 40.13, 32.08, 39.5, and 36.2. Interestingly, these two distinct images from our data set, included in the grouping of [2], are of the same of piece paper and should provide a very close match. However, they were not designated a very close match by the software we subsequently developed, which assumed chain line straightness. They were assessed as somewhat similar but not the closest match.

EXPERIMENT 3: A TEST OF PARALLELISM

A common assumption is that the chain lines in a single print are parallel [7]. To test this assumption, this third experiment uses a set of $N = 205$ beta-radiographs we received from the Morgan Library & Museum, the Metropolitan Museum of Art, and the Rijksmuseum to measure the degree to which the lines are parallel, by looking at the difference between the angles of the chain lines. Figure 5 shows a histogram of the maximum angle that occurs in each print. The mean and median values are 0.4 and 0.36°, respectively, so overall the lines tend to be fairly parallel. As will be discussed in the following sections, when the lines are not parallel, this can be quite useful as an identifying feature of the print.

THE GEOMETRY OF MOLDMATES

The dimensions of the molds used in papermaking were typically much larger than the sizes of final folded sheets as used for pages in a book or unfolded as stand-alone sheets. Figure 4 shows an exaggerated schematic of a mold with eight chain lines that has been divided into three sheets labeled *A*, *B*, and *C*. It is worth making a few observations. First, chain line matching cannot show directly that papers *A* and *B* come from the same mold since they have no chain lines in common. Second, measurements of chain spacings (interchain distances) may show a close match between *B* and *C* (because the corresponding lines are parallel), but they will not show a close match between *A* and *C* (because the lines have different angles). Third, given two pieces of paper with many chain lines, it is necessary to check all possible subsets for matches. For example, only the final four lines of *C* can match with the lines of *B*. Though not obvious from the figure, it is also worth mentioning that it is typically unknown a priori whether a paper has been flipped left-right, top-bottom (or both) when taking the beta-radiograph; thus, it is also necessary to check all four



[FIG5] (a) A histogram of the measurements of the widths of the bounding boxes about the chain lines. The majority of chain lines appear straight, though there are outliers. The methods of this paper will not be effective for these outliers. (b) A histogram of the maximum angle found between the chain lines in each print gives a measure of how parallel the chain lines are. The majority of papers have fairly parallel chain lines; those that do not may be more readily identified by the methods of this paper.

orientations when attempting to find matches. Finally, the more chain lines that overlap, the more certain one can be that two papers do (or do not) match.

Using the distance/angle form (1) for the detected chain lines allows a geometric calculation to compensate for the angle discrepancy. Essentially, this estimates the distance β in Figure 4, though it may be pictured geometrically as sliding the chain lines of paper *A* up and/or down until they best match the chain lines of paper *C*. The geometry of two nonparallel chain lines is illustrated in the right hand side of Figure 4. The distances d_1 and d_2 are two measurements of how far apart the lines are, and are related via

$$d_1 = d_2 + \beta \sin(\psi). \quad (3)$$

Because segments ① and ② are parallel, angle α equals angle δ . (Counterclockwise angles are positive while clockwise angles are negative.) Thus, because $\alpha = 90^\circ - \theta$ and $\delta = 90^\circ - \phi$, $\theta = \phi$. Furthermore, because $\gamma + \psi = 90^\circ$ and $\gamma - \phi = 90^\circ$, $\psi = -\phi = -\theta$. A key observation is that the angle off horizontal of the shortest distance line to the right line from any measurement point on the left-most line is the same.

Given a collection of N laid papers each with M chain lines (presumed to be straight but not necessarily parallel), the first step is to measure (automatically, semiautomatically, or manually) the positions of the chain lines. Using the distance/angle parameterization (1), the M chain lines in paper i can be represented by the vector

$$\{\rho^i, \theta^i\} \equiv \{\rho_1^i, \rho_2^i, \dots, \rho_M^i, \theta_1^i, \theta_2^i, \dots, \theta_M^i\} \in \mathbb{R}^{2M}, \quad (4)$$

where ρ_k^i and θ_k^i represent the distance and angle of each chain line k with respect to the origin of the i th paper. For ease of notation, these are partitioned into vectors ρ^i and θ^i , each in \mathbb{R}^M .

Consider two pieces of paper i and j that are separated by an unknown distance β on the same mold (for example, paper A and the first four chain lines of C in Figure 4). Applying the logic of (3) to each of the successive pairs of chain lines suggests estimating β using a least squares procedure

$$\min_{\beta} \left\| \begin{pmatrix} \rho_1^j - \rho_1^i \\ \rho_2^j - \rho_2^i \\ \vdots \\ \rho_M^j - \rho_M^i \end{pmatrix} - \beta \begin{pmatrix} \sin\left(\frac{\theta_1^j + \theta_1^i}{2}\right) \\ \sin\left(\frac{\theta_2^j + \theta_2^i}{2}\right) \\ \vdots \\ \sin\left(\frac{\theta_M^j + \theta_M^i}{2}\right) \end{pmatrix} \right\|. \quad (5)$$

Though the measurements θ_k^j and θ_k^i are similar, they are not identical, which is why they are averaged. Rewriting (5) using Y_{ij} for the first vector and X_{ij} for the second, the β that minimizes the least squares error $\|Y_{ij} - \beta X_{ij}\|_2 = (Y_{ij} - \beta X_{ij})^T (Y_{ij} - \beta X_{ij})$ is $\beta_{ij}^* = (X_{ij}^T X_{ij})^{-1} X_{ij}^T Y_{ij}$.

IDENTIFYING MOLDMATES

The geometric analysis of the previous section aligns the chain lines of two papers (as well as possible) by offsetting them a distance β_{ij}^* . Moreover, the value achieved at this optimum

$$d_{\beta^*}(i, j) = \|Y_{ij} - \beta_{ij}^* X_{ij}\|_2 \quad (6)$$

provides a measure of the dissimilarity between the chain lines of the papers i and j . A value of zero would mean that the shifted versions overlay exactly while a large value implies that the two sets of chain lines are very different. Moldmates should have small $d_{\beta^*}(i, j)$ while unrelated papers should have larger values.

Initial experiments showed that applying (6) directly to the problem of finding moldmates can lead to some undesirable false matches. Fortunately, many of these can be removed by considering more than just the value of $d_{\beta^*}(i, j)$. We have found four such modifications useful. First, if the difference between any of the angles is too large, any small value of $d(i, j)$ is accidental; such accidents can be removed from consideration by placing a threshold on $\max |\theta^j - \theta^i|$. Second, if all of the individual angles are small, then the computation of β^* is irrelevant, and the distance

$$d(i, j) = \|\rho^j - \rho^i\| \quad (7)$$

is more appropriate. Third, a β^* value that is larger than about 18 in (a typical size for the frame) is impossible; any such false matches can be removed by placing a threshold on β^* . Finally, if the maximum of all the $|\rho^j - \rho^i|$ is small, a distance of zero is reported to encourage consideration of this potential mold match. These can be written succinctly in two parts:

$$\begin{aligned} & \text{if } \max_j |\theta^j| < 0.005, \\ \delta(i, j) &= \begin{cases} \infty & \max |\theta^j - \theta^i| > 0.02 \\ 0 & \max |\rho^j - \rho^i| < 0.01 \\ d(i, j) & \max |\rho^j - \rho^i| > 0.01 \end{cases} \\ & \text{if } \max_j |\theta^j| \geq 0.005, \\ \delta(i, j) &= \begin{cases} \infty & \beta_{ij}^* > 18 \\ 0 & \max |\rho^j - \beta_{ij}^* \sin(X_{ij}) \rho^i| < 0.01, \\ d_{\beta^*}(i, j) & \max |\rho^j - \beta_{ij}^* \sin(X_{ij}) \rho^i| \geq 0.01 \end{cases} \quad (8) \end{aligned}$$

where X_{ij} is defined as in (5). The various thresholds (0.005, 0.01, 0.02, etc.) are representative and may require fine-tuning for specific data sets. The threshold on the differences in angle and distance are set so they are larger than the differences resulting from manual marking errors. The range of measurement point separation is set by typical mold dimensions. The threshold for declaring the chain lines parallel (so that beta need not be estimated) depends on the shift causing changes in the distance/spacing vector that exceeds the threshold that designates a match. For a modest β of 10 in and a reasonable distance threshold of 0.02 in, the sine of the relative angle should be less than 0.002. For such small values, this corresponds to the angle threshold for declaring nonparallelism of 0.002 radians or 0.12°. By this measure a large fraction of our images contain nonparallel chain lines.

HOW MANY CHAIN SPACES?

The objective in moldmate identification is not to return a single answer with the most likely fit, since this is not what the paper conservator or art historian would find most useful. Rather, the goal is to reduce a large library of chain line pattern images to a small number so that the expert user can investigate further, with the expectation that if a match exists, it will likely be in this smaller set. One issue is how many adjacent chain spaces are needed to achieve this goal. Currently, the majority of the beta-radiographs available to us are limited to the vicinity of watermarks, which typically do not cover the entire print. This is due to the earlier emphasis on watermark matching and the technical limitations of beta-radiography [14]. An assessment of the required number of adjacent chain spaces to achieve reliable matching would be helpful to persuade museums to undertake the expense in equipment and personnel costs to mount a campaign to acquire full-print images revealing its entire chain-line pattern. This section addresses that need by building a simple statistical model of the chain-line process and assessing the probability of error of the model as a function of the number of chain spaces. Specifically, the analysis calculates the probability that a sheet of interest will be confused with one or more of the existing sheets and the results provide guidelines for the number

of chain lines that need be present to reliably detect moldmates. This provides a rule of thumb for the real problem.

In the past, attempts have been made to use the average inter-chain spacing as a way of characterizing sheets of laid paper for the purpose of identifying moldmates [6]. This section begins by conducting a statistical analysis of this procedure (using data gathered as in the previously discussed experiments) to show why the mean value alone is unlikely to provide a useful characterization, which agrees with observed behavior in experiments reported in [7]. The analysis is then extended to consider the use of $M + 1$ chain lines (M interchain spaces), and a simple rule of thumb is derived that suggests how large M must be to reliably distinguish moldmate matches to a candidate sheet of paper from among a universe of $N \gg M$ sheets.

The database of images of chain-line patterns in laid papers (provided by the Morgan Library & Museum, the Metropolitan Museum of Art, and the Rijksmuseum and used in these experiments) includes $N = 205$ sheets. The mean chain spacing, calculated as $\rho_{i+1} - \rho_i$ where i ranges over all the chain lines in a given image, is $\mu = 0.977$ inches with standard deviation $\sigma_m = 0.082$. Considered as a collection of random samples, this is plausibly Gaussian, as can be seen in the histogram of Figure 6. To model the question of interest, consider a collection of N random variables $\bar{m}_i \sim \mathbf{N}(\mu, \sigma_m^2)$ each of which is measured in the presence of some small error

$$m_i = \bar{m}_i + \epsilon_i \text{ where } \epsilon_i \sim \mathbf{N}(0, \sigma_\epsilon^2). \quad (9)$$

A new candidate element $m^* \sim \mathbf{N}(\mu, \sigma_m^2)$ is measured from the same distribution as the m_i , and we wish to know how many of the m_i lie close to this candidate m^* . (These will be the potential moldmates that must be subjected to closer examination.) This also requires quantifying “close to.” For simplicity, consider two elements close if they lie within ℓ standard deviations of the measurement error, that is, if they lie within $\pm \ell \sigma_\epsilon$. This is shown schematically in Figure 7, where the shaded region S lying between the lines defined by $m^* - \ell \sigma_\epsilon$ and $m^* + \ell \sigma_\epsilon$ represents the probability that elements of the data set lie close to the measured m^* .

Let

$$f(x, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (10)$$

be the normal density function. The probability represented by the shaded region S is

$$p_S(m^*, \sigma_m, \ell \sigma_\epsilon) = \int_{m^* - \ell \sigma_\epsilon}^{m^* + \ell \sigma_\epsilon} f(x, \mu, \sigma_m) dx \\ = \frac{1}{2} [\text{Erf}(\frac{m^* + \ell \sigma_\epsilon}{\sqrt{2} \sigma_m}) - \text{Erf}(\frac{m^* - \ell \sigma_\epsilon}{\sqrt{2} \sigma_m})]. \quad (11)$$

Some observations:

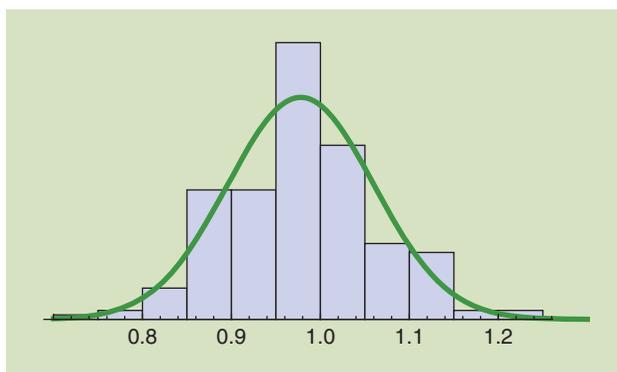
- 1) The probability p_S is independent of μ since both the data set and the new candidate element are assumed to have the same mean.
- 2) As the measurement error $\sigma_\epsilon \rightarrow 0$, the probability of being close goes to zero. In words, the more accurately the

measurements can be made, the greater the distinguishing power of the method. Conversely, larger σ_ϵ cause $p_S \rightarrow 1$.

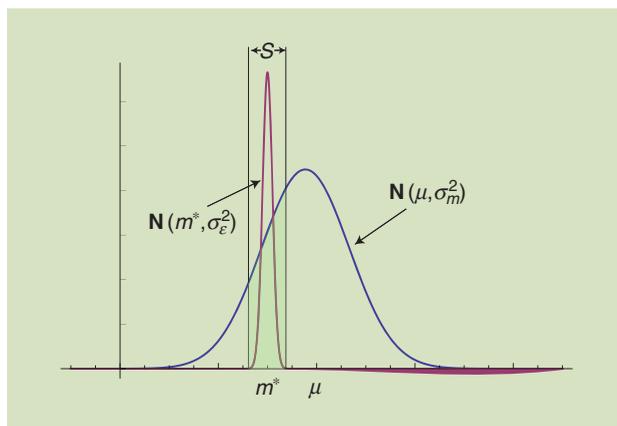
3) The probability p_S does depend on the particular m^* . For $m^* \ll \mu$ or $m^* \gg \mu$, p_S is small.

4) Conversely, $m^* = \mu$ implies that $p_S = \text{Erf}(\ell \sigma_\epsilon / \sqrt{2} \sigma_m)$. With an eight-to-one ratio of σ_m to $\ell \sigma_\epsilon$, this is 0.1.

Case 4) may be interpreted to say that roughly 10% of the elements of the database will be considered close to the candidate. Since each must be considered in four orientations, this approximately quadruples (to 40%) the percentage of sheets that will be considered potential matches. One ameliorating factor is that the measurement of the mean values are not independent of the number of chain spacings M ; larger M cause smaller effective σ_ϵ . If each independent chain line is measured with an error σ_ϵ , the average of M is effectively drawn from $\mathbf{N}(\mu, (1/M) \sigma_\epsilon^2)$ and the effective standard deviation is $\sigma_\epsilon = \sigma_\epsilon' / \sqrt{M}$. Perhaps the best use of the mean value of the chain spacings is in the cases described by case 3) where m^* deviates from the mean μ . In these cases, p_S will be small and there will be fewer possible matches to consider. From another point of view, these estimates



[FIG6] A histogram of the mean chain line spacing of the N sheets in the database is reasonably modeled as a normal probability distribution with mean $\mu = 0.977$ and standard deviation $\sigma_m = 0.082$, as shown in green.



[FIG7] The means of the chain spacings are considered as elements of $\bar{m}_i \sim \mathbf{N}(\mu, \sigma_m^2)$. The region S shows the region in which elements are “close to” the test element m^* and the probability p_S represents the percentage of elements in the database close to the candidate.

[TABLE 1] THE THREE MODELS MAKE INCREASINGLY MORE REALISTIC PREDICTIONS OF THE FRACTION OF IMAGES THAT WILL BE REGARDED AS "CLOSE," AS A FUNCTION OF THE NUMBER OF CHAIN LINES.

MODEL	FRACTION
#1 AVERAGE CHAIN-LINE SPACING	0.4
#2 INDEPENDENT SEQUENCE OF CHAIN LINES	$4(0.1)^M$
#3 DEPENDENT SEQUENCE OF CHAIN LINES	$4(0.1)(0.25)^{M-1}$

suggest that the mean value may be better used as a method of ruling out mold matches (in those cases when p_S is small) than as a method of locating mold matches in general.

Fortunately, more information is available in the chain spacings than just the mean value. The simplest way to model the chain spacings is to presume that the sequence of interchain intervals is

$$\bar{m}_i = \{m_{i,1}, m_{i,2}, \dots, m_{i,M}\} \text{ for } i = 1, 2, \dots, N, \quad (12)$$

where each component m_{ij} is independent of m_{ik} for $j \neq k$ and where \bar{m}_i is independent of \bar{m}_j for $i \neq j$. Following the logic of (9)–(11), the probability $p_S(\bar{m}^*, \sigma_m, \ell\sigma_\epsilon)$ of accidental matches from the database to a candidate vector m^* is

$$p_S(m_{i,1}^*, \sigma_m, \ell\sigma_\epsilon) p_S(m_{i,2}^*, \sigma_m, \ell\sigma_\epsilon) \cdots p_S(m_{i,M}^*, \sigma_m, \ell\sigma_\epsilon) \\ = (p_S(m^*, \sigma_m, \ell\sigma_\epsilon))^M. \quad (13)$$

Unfortunately, this exponential decay of 0.1^M represents an overly optimistic scenario in which each chain-space element is independent of the others even within a single sheet.

A more realistic model observes that each chain-space sequence has a mean value m^* and a small variance $\sigma_{\bar{m}}^2$ about that mean. Let m^* be the mean value drawn from $\mathbf{N}(\mu, \sigma_m^2)$ as before. The chain spacings are then

$$\{m^* + \hat{m}_1, m^* + \hat{m}_2, \dots, m^* + \hat{m}_M\}, \quad (14)$$

where $\hat{m}_i \sim \mathbf{N}(\mu, \sigma_{\bar{m}}^2)$ and where the standard deviation $\sigma_{\bar{m}}$ is the variation in the chain spacings within a given sheet (which is less than the variation in the complete data set). Assuming again that the m^* is independent of the \hat{m}_i and that \hat{m}_i is independent of \hat{m}_j for $i \neq j$, the desired probability $p_S(m^*)$ can be factored as

$$p_S(m^*, \sigma_m, \ell\sigma_\epsilon) p_S(\bar{m}_2, \sigma_{\bar{m}}, \ell\sigma_\epsilon) \cdots p_S(\bar{m}_M, \sigma_{\bar{m}}, \ell\sigma_\epsilon) \\ = p_S(m^*, \sigma_m, \ell\sigma_\epsilon) (p_S(\bar{m}, \sigma_{\bar{m}}, \ell\sigma_\epsilon))^{M-1}. \quad (15)$$

Since $\sigma_i < \sigma_{\bar{m}} < \sigma_m$, the probability $p_S(m^*)$ is larger than in (13), giving a more pessimistic (though also more realistic) assessment. Using the value $\sigma_{\bar{m}} = 0.034$ (the average of the standard deviations), $\ell\sigma_\epsilon \approx 0.01$ [the average measurement error from (2)], this is

$$p_S = (0.1)(0.25)^{M-1}. \quad (16)$$

By $M = 3$ (four chain lines), there is about $4p_S$ chance of false matches. The factor four occurs because of the need to inflate the number of sheets by four to consider all the possible rotations and reflections. By $M = 4$, $4p_S$ drops to 1% and continues to decrease exponentially as M increases, reducing the number of potential false matches by a factor of four with each additional chain line. This leads us to seek matching patterns of five (or more) adjacent chain lines with four (or more) chain spaces. A comparison of the predictions of the three sets of modeling assumptions is given in Table 1.

A PAIR OF MOLDMATES IDENTIFIED VIA A CHAIN-LINE PATTERN MATCH

Applying the distance measure $\delta(i, j)$ of (8) to the data set of [2], we observed that there was a potential match between Rembrandt's *Medea, or the Marriage of Jason and Creusa*, dated 1648 (B112iv, Figure 1) and a left-right flipped version of *The Artist's Mother in Widow's Dress and Black Gloves*, B344. A reproduction of the etching and its beta-radiograph is shown in [15]. What is particularly intriguing about this match is that the latter print's authorship has been questioned for many years. Most scholars after Adam Bartsch, the indefatigable 18th-century cataloguer of Rembrandt prints, agree that B344 is by a pupil of Rembrandt, who was most likely copying an earlier print, dated circa 1631, *The Artist's Mother Seated at a Table, Looking Right* (B433). This chain-line pattern match of papers in restrikes from around 1650, provides support for the conclusion that the pupil's print originated in Rembrandt's studio. Presuming creation of B344 around 1635 narrows the list of potential copyists to pupils in Rembrandt's studio at that time. The close match between the (shifted) chain lines of these two images (with the proper flips) is shown in detail in [15]. Though this match was identified solely from the close similarity of their chain-line patterns, it can be verified [1] by the match of the watermarks, though B344 has only a fragment of the watermark that is fully contained in B112iv.

CONCLUSIONS AND OPEN QUESTIONS

This article has highlighted the problem of moldmate identification of laid paper and provided a basic argument for the feasibility of the project. The potential for automated procedures to identify moldmates among collections of similar handmade laid papers is exciting to paper conservators and graphic art curators. Using our first-pass signal processing strategies and even with a modestly sized database, we were able to identify a moldmate pair of Rembrandt prints that was previously unrecognized by the owner.

This article has attempted to clearly display the moldmate problem as a set of signal processing tasks with the goal of making the problem accessible to the signal processing community, where those who are not experts in the technical analysis of paper may contribute. One key to this is a description of the various signal processing tasks that must be completed. The other major component is making the images and data sets publicly available [2] to help fuel this cross-disciplinary effort.

Some of the problems highlighted in this article are straightforward applications of signal processing techniques while others may require significant effort to realize. For example, a wonderfully complex problem is the automated extraction and comparison of watermarks (such as that in Figure 1). Such shapes are much more complex than the simple straight line we have chosen to attempt to extract and the automated comparison between such unparameterized shapes is not straightforward. On the other hand, there are many tasks that might benefit from more advanced signal processing techniques, more clever implementations of the algorithms, and/or more thoughtful metrics. For instance, the proposed algorithm for chain-line extraction (in Figure 2) can no doubt be improved, both in its accuracy and in its range of applicability, reducing the manual portions with automated extraction techniques. Similarly, the distance measure (8) can undoubtedly be expanded or improved. As shown in [15], such chain-line pattern matching software can also be used to assess concerns paper experts have regarding the possibility that differences in the wetting, pressure, and drying of intaglio printing and conservation treatments encountered separately by moldmates can distort the chain lines into distinctly different patterns.

The analytic contribution of this article is to propose a simple model where the statistics of the chain-line data set can be used to estimate the number of chain lines needed to distinguish laid papers. For example, three or four chain lines are inadequate to reliably locate moldmates, and the rule of thumb relates the certainty of any such identification with the number of chain lines, given the experimental and computational errors inherent in locating the chain lines.

AUTHORS

C. Richard Johnson, Jr. (johnson@ece.cornell.edu) received the Ph.D. degree in electrical engineering with a minor in art history from Stanford University, California. He is the Geoffrey S.M. Hedrick Senior Professor of Engineering at Cornell University, Ithaca, New York; a scientific researcher of the Rijksmuseum in Amsterdam, The Netherlands; and computational art history advisor to The Netherlands Institute for Art History. He founded the Thread Count Automation Project with the van Gogh Museum (2007), the Historic Photographic Paper Classification Challenge with the Museum of Modern Art (2010), and the Chain Line Pattern Matching Project with the Morgan Library & Museum (2012), which has since been joined by the Rijksmuseum, the Metropolitan Museum of Art, and the Dutch University Institute for Art History.

William A. Sethares (sethares@gmail.com) received the Ph.D. degree in electrical engineering from Cornell University, Ithaca, New York. He was previously with the Raytheon Company designing image processing systems and is currently a professor in the Department of Electrical and Computer Engineering at the University of Wisconsin, Madison. He has held visiting positions at the Australian National University; the Technical Institute in Gdansk, Poland; the NASA Ames Research Center in Mountainview, California; and the Rijksmuseum in

Amsterdam, The Netherlands. His research interests include adaptation and learning imaging and audio signal processing systems. He is the author of four books and holds six patents.

Margaret Holben Ellis (mhellis@themorgan.org) is the Eugene Thaw Professor, Institute of Fine Arts, New York University, where she teaches paper conservation and technical art history. She also serves as the director, Thaw Conservation Center, Morgan Library & Museum. She is a fellow of the American (AIC) and International (IIC) Institutes for Conservation of Historic and Artistic Works, a certified conservator/restorer with the Institute of Conservation, and the director, Heritage Preservation. She has received teaching and merit citations (AIC) and a Rome Prize from the American Academy in Rome. She has published and lectured on artists from Raphael to Pollock; her research on artists' materials is similarly far ranging.

Saira Haqqi (haqqis@gmail.com) is an Andrew W. Mellon Fellow in Library and Archives Conservation at the Conservation Center of the Institute of Fine Arts at New York University.

REFERENCES

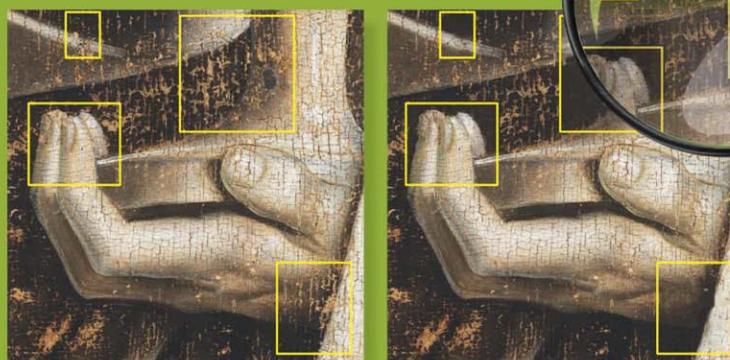
- [1] E. Hinterding, *Rembrandt as etcher: The practice of production and distribution*. Ouderkerk aan den IJssel, The Netherlands: Sound & Vision, vol. 1, chap. 1, 2006.
- [2] RKD image sharing. [Online]. Available: <http://english.rkd.nl/Services/image-sharing>.
- [3] Univ. of Iowa. Chancery papermaking. [Online]. Available: <http://now.uiowa.edu/2013/05/center-book-releases-short-film-about-papermaking>.
- [4] A. Stevenson. "Chain-indentations in paper as evidence," *Stud. Bibliograp.*, vol. 6, pp. 181–195, 1953–1954.
- [5] D. L. Vander Meulen. "The identification of paper without watermarks: The example of Pope's 'Dunciad,'" *Stud. Bibliograp.*, vol. 37, pp. 58–81, 1984.
- [6] J. C. A. van der Lubbe, E. P. van Someren, and M. J. T. Reinders, "Dating and authentication of Rembrandt's etchings with the help of computational intelligence," in *Proc. Int. Cultural Heritage Informatics Meeting (ichim 01)*, Sept. 2001, pp. 485–492.
- [7] M. van Staalduinen, J. C. A. van der Lubbe, E. Backer, and P. Paclik, "Paper retrieval based on specific paper features: Chain and laid lines," in *Multimedia Content Representation, Classification and Security* (Lecture Notes in Computer Science), vol. 4105, Springer-Verlag: Berlin Heidelberg, 2006, pp. 346–353.
- [8] M. van Staalduinen, "Content-based paper retrieval towards reconstruction of art history," Ph.D. thesis, Faculty of Elect. Eng. Math. Comput. Sci., TU Delft, 2010.
- [9] C. R. Johnson, Jr., D. H. Johnson, N. Hamashima, H. S. Yang, and E. Hendriks, "On the utility of spectral-maximum-based automated thread counting from x-rays of paintings on canvas," *Stud. Conserv.*, vol. 56, no. 2, pp. 104–114, 2011.
- [10] D. H. Johnson, C. R. Johnson, Jr., A. G. Klein, W. A. Sethares, H. Lee, and E. Hendriks, "A thread counting algorithm for art forensics," in *Proc. IEEE DSP Workshop*, Nov. 2009, pp. 679–684.
- [11] L. van Tilborgh, T. Meedendorp, E. Hendriks, D. H. Johnson, C. R. Johnson, Jr., and R. G. Erdmann, "Weave matching and dating of van Gogh's paintings: An interdisciplinary approach," *Burlington Mag.*, vol. CLIV, pp. 112–122, Feb. 2012.
- [12] W. Liedtke, C. R. Johnson, Jr., and D. H. Johnson, "Canvas matches in Vermeer: A case study in the computer analysis of fabric supports," *Metropol. Museum J.*, vol. 47, pp. 99–106, 2012.
- [13] D. H. Johnson, C. R. Johnson, Jr., and R. G. Erdmann, "Weave analysis of paintings on canvas from radiographs," *Signal Process.* (Special Issue on Image Processing for Art Investigation), vol. 93, pp. 527–540, Mar. 2013.
- [14] D. Kushel, "Radiographic methods used in the recording of structure and watermarks in historic paper," in *Fresh Woods and Pastures New: Seventeenth Century Dutch landscape Drawings from the Peck Collection*. Ackland Art Museum: Chapel Hill, 1999, pp. 116–123.
- [15] C. R. Johnson, Jr., W. A. Sethares, M. H. Ellis, S. Haqqi, R. Snyder, E. Hinterding, I. van Leeuwen, A. Wallert, D. Christoforou, J. van der Lubbe, N. Orenstein, A. Campbell, and G. Dietz. The application of automated chain line pattern (CLIP) matching to identify paper moldmate candidates in Rembrandt's prints. In *New Directions in the Study of Rembrandt and His Circle*. [Online]. Available: <http://people.ece.cornell.edu/johnson/CLIP-Herst.pdf>



[Laurens van der Maaten and Robert Erdmann]

Automatic Thread-Level Canvas Analysis

© GHEENT, KATHEDRALE KERKABRIEK, PHOTO COURTESY OF KIK-IRPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

[A machine-learning approach
to analyzing the canvas of paintings]

Canvas analysis is an important tool in art-historical studies, as it can provide information on whether two paintings were made on canvas that originated from the same bolt. Canvas analysis algorithms analyze radiographs of paintings to identify (ir)regularities in the spacings between the canvas threads. To reduce noise, current state-of-the-art algorithms do this by averaging the signal over a number of threads, which leads to information loss in the final measurements. This article presents an algorithm capable of performing thread-level canvas analysis: the

algorithm identifies each of the individual threads in the canvas radiograph and directly measures between-distances and angles of the identified threads. We present two case studies to illustrate the potential merits of our thread-level canvas analysis algorithm, viz. on a small collection of paintings ostensibly by Nicholas Poussin and on a small collection of paintings by Vincent van Gogh.

INTRODUCTION

The analysis of paintings is increasingly aided by the availability of imaging and image-processing tools, including various types of imaging to reveal underpaintings and underdrawings [1], [2], techniques for automatic brushstroke segmentation and analysis [3]–[5], and automatic face analysis techniques [6]. These

Digital Object Identifier 10.1109/MSP.2015.2407091

Date of publication: 15 June 2015

tools can help art historians, conservators, and restoration artists understand the way in which different painters worked and may provide clues about the attribution of a painting to a particular artist. One of the most commonly performed analyses is the analysis of radiographs (X-rays) of paintings, as such radiographs can reveal visible and hidden paint layers according to the radio-opacity of the paint. Radiographs do not only provide essential information on the materials that need to be used in restorations, but they may also form the basis for valuable art-historical insights.

In addition to information on the radio-opacity of the paint, radiographs also reveal the individual threads in the canvas (see Figure 1) because the ground layer, which generally contains lead, varies in thickness when it is spread over the textured surface of the bare canvas. Until a few years ago, art experts generally considered the display of the threads as a disturbance because it was obfuscating what they were truly interested in: the composition of the different paint layers. More recently, however, scholars have realized that the canvas threads visible in radiographs may carry important art-historical information [7]. This information arises from the fact that the thicknesses of the threads are irregular because of the way a loom works. Some threads are thinner than others because of natural variations in the manufacturing process: a thread with higher tension on it tends to be narrower. Such irregularities persist throughout the entire bolt of canvas. As a result, paintings made on canvas that was cut from the same bolt will likely have the same irregularities in their thread thicknesses. Thread thickness measurements may be used to identify the bolt from which the canvas originates. Specifically, if we find two paintings that have the same canvas thread thicknesses, we have obtained a strong indication that these two paintings were made in the same workshop in the same period [8].

Thread densities or thread spacings are good surrogates for the thread thicknesses that we would like to measure. Various recent studies have attempted to measure thread densities and/or spacings across the canvas, in particular, in paintings by Nicolas Poussin [9], Vincent van Gogh [10], Johannes Vermeer [11], [12], and Diego Velázquez [13]. In particular, these studies estimate the thread density in a small patch of the painting using a two-dimensional (2-D) Fourier analysis [14] or an approach based on measuring autocorrelations in small canvas patches [9]. These analyses provide valuable information, but they average information across relatively large patches of canvas (over five threads or more), which leads to low-resolution thread density maps. The averaging may hide variations in the thickness of individual threads, which makes it harder to obtain conclusive evidence that two canvases originated from the same roll.

In contrast to most prior work (the work by [12] is a notable exception), this article proposes an algorithm for thread-level analysis of the canvas. Our approach involves training a

machine-learning model to identify thread crossings in the canvas based on their visual appearance. The resulting model is used to automatically identify the millions of thread crossings inside a canvas, which, in turn, form the basis for measuring thread spacings. We show the merits of thread-level canvas analysis by using it to study a collection of three alleged Nicolas Poussin paintings as well as a small collection of paintings by Vincent van Gogh.

RADIOGRAPHS DO NOT ONLY PROVIDE ESSENTIAL INFORMATION ON THE MATERIALS THAT NEED TO BE USED IN RESTORATIONS, BUT THEY MAY ALSO FORM THE BASIS FOR VALUABLE ART-HISTORICAL INSIGHTS.

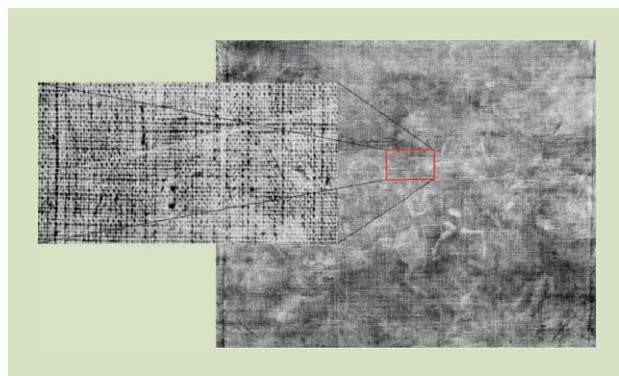
THREAD-LEVEL CANVAS ANALYSIS

Our approach to thread-level canvas analysis comprises four main parts: 1) we extract features from the radiographs that are sensitive to the sig-

nals produced by the threads, 2) we train and deploy a machine-learning model that automatically detects thread crossings based on these features, 3) we use the response of this detector to estimate the distance between neighboring threads, and 4) we automatically try to match the resulting thread-distance maps produced for different canvases to determine whether or not these canvases likely originate from the same roll. The details of these four parts of our approach are described separately below. A MATLAB implementation of our canvas analysis algorithm is publicly available from <http://lvdmaaten.github.io/canvas>.

FEATURE EXTRACTION

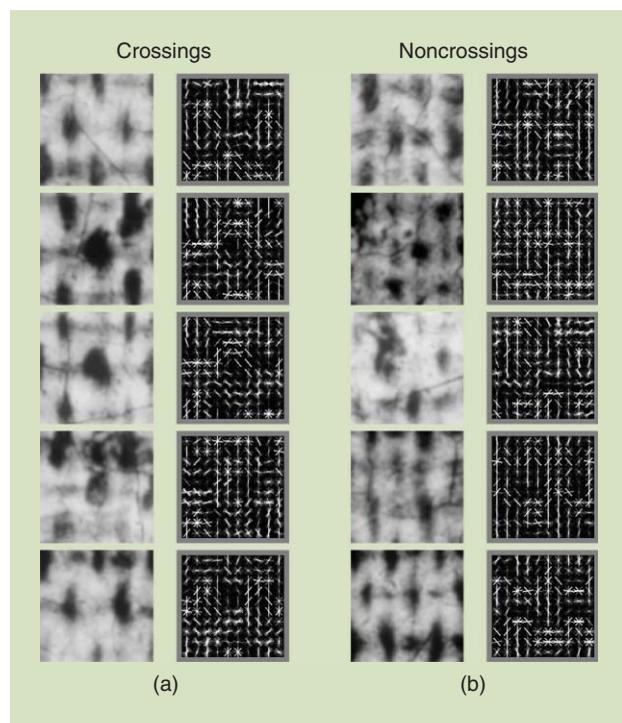
Thread-crossing detection can be performed with very high accuracy because thread crossing corresponds to visually salient locations in canvas radiographs. Our thread-crossing detector: 1) extracts histograms-of-oriented-gradient (HOG) features from the image region around the canvas location of which we want to determine whether or not it corresponds to a thread crossing and 2) uses a linear support vector machine to determine based on these features whether or not the location is a thread crossing or a “nonthread crossing.”



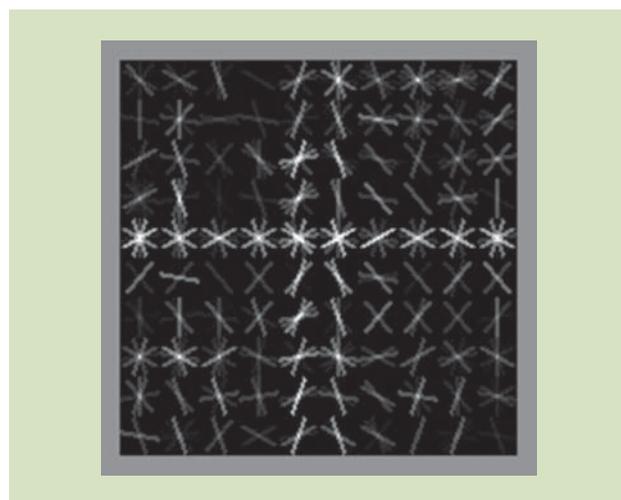
[FIG1] An example of a high-resolution radiograph of the Nicolas Poussin painting *Triumph of Bacchus*. The inset shows the individual threads in a small part of the canvas. (Radiography reproduced with permission from the Nelson-Atkins Museum of Arts in Kansas City.)

HOG features [15] describes an image location by a histogram of image gradient magnitudes for a number of quantized gradient orientations (we used eight orientations in our study). The histograms are constructed over small image patches; depending on the type of canvas and the resolution of the canvas radiographs, we used image patches of 4×4 or 8×8 pixels in this study [in

radiographs scanned at 600 dots per inch (dpi)]. Subsequently, the histograms are normalized for contrast differences by normalizing the L2-norm of all the histograms in a square, spatially connected block of four image patches. The advantage of the use of image gradients and the subsequent contrast normalization is that it produces partial invariance to larger-scale signals in the radiograph images that stem from the paint layers (in particular, from layers of white paint that contain relatively large amounts of lead, which in turn lead to strong radiograph responses). To obtain additional invariance to small variations in the gradient magnitudes, the contrast-normalized histograms are clipped at 0.2 and then renormalized according the L2-norm to produce the final HOG features. The resulting features have a particular structure near thread crossings, which is illustrated in Figure 2.



[FIG2] Examples of five canvas patches around (a) a thread crossing and (b) five randomly selected canvas patches along with the corresponding HOG feature-representation of these patches.



[FIG3] A visualization of our thread-crossing detector. The figure shows that the detector identifies crossings as a location at which prolonged horizontal and vertical edges (caused by the boundaries of the threads) cross.

THREAD-CROSSING DETECTION

To obtain a model that can automatically distinguish thread crossings from other structures in canvas radiographs, we train a logistic regression model to discriminate a set of image patches that contain manually annotated thread crossings (positive examples) from a set of image patches that are randomly sampled from the canvas (negative examples). Ideally, the set of positive examples describes the variation in the visual appearance of thread crossings, while the set of negative examples captures the visual variation in nonthread-crossings. Denoting an image patch in the training data by \mathbf{I} , the corresponding label by $z \in \{-1, +1\}$, and the HOG feature function by ϕ , the logistic regressor builds the following probabilistic model:

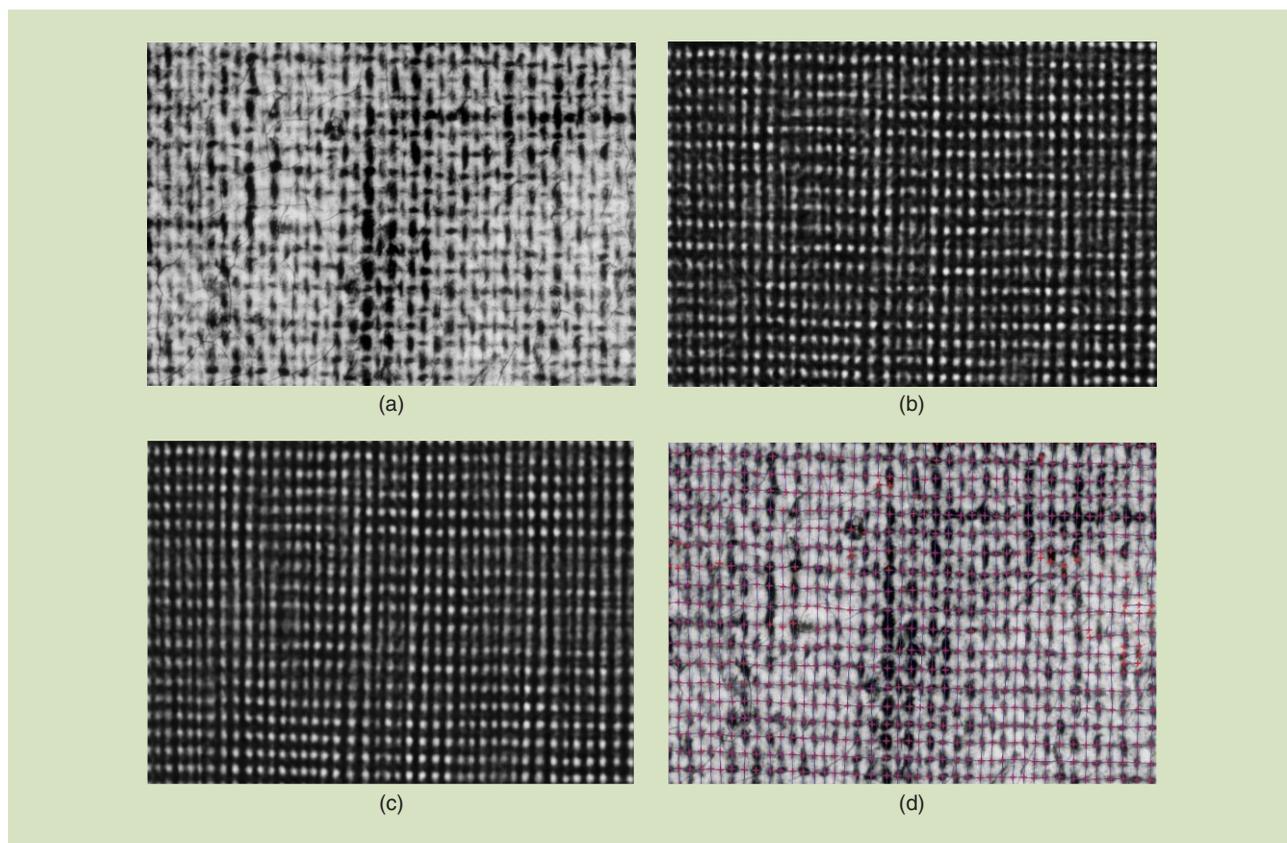
$$p(z | \mathbf{I}; \theta) = \frac{\exp[z(\mathbf{w}^\top \phi(\mathbf{I}) + b)]}{\exp[-(\mathbf{w}^\top \phi(\mathbf{I}) + b)] + \exp[\mathbf{w}^\top \phi(\mathbf{I}) + b]}.$$

Herein, the parameters $\theta = \{\mathbf{w}, b\}$ comprise a vector \mathbf{w} and a scalar bias b . After incorporating an isotropic Gaussian prior over \mathbf{w} , $p(\mathbf{w}) = \mathcal{N}(\mathbf{w} | 0, \sigma^2 \mathbf{I})$, the parameters θ are estimated via maximum a posteriori (MAP) estimation on the aforementioned data set $\mathcal{D} = \{(\mathbf{I}_1, y_1), (\mathbf{I}_2, y_2), \dots, (\mathbf{I}_N, y_N)\}$:

$$\theta^* = \operatorname{argmax}_{\theta} \sum_{n=1}^N \log p(z_n | \mathbf{I}_n; \theta) - \frac{1}{2} \sigma^2 \|\mathbf{w}\|^2.$$

Herein, the hyperparameter σ^2 is set via cross-validation. The resulting weights \mathbf{w}^* are visualized as a HOG feature in Figure 3. The figure shows that they are a template for the visual appearance of a typical thread crossing. This training procedure need to be performed only once for a particular type of canvas, assuming the imaging conditions are similar across the collection of canvas radiographs.

After training, the trained model (i.e., the template) is applied to all image patches in a canvas radiograph to predict the likelihood $p(z | \mathbf{I}; \theta)$ that a location in the canvas contains a thread crossing. An example of the resulting likelihood map is shown in Figure 4(c); brighter colors indicate a higher likelihood of the location containing a thread crossing according to the logistic-regression model. The quality of the likelihood map can be substantially improved by exploiting that the likelihood map ought to be quite regular: the likelihood of the thread-crossing presence at location (x, y) should be high when there is a high likelihood of



[FIG4] An illustration of our canvas analysis algorithm: (a) a small patch of canvas taken from Poussin's *Triumph of Bacchus*, (b) the response of our thread crossing classifier on the patch of canvas, (c) the response of our model after incorporating the pictorial-structures model that exploit canvas regularity, and (d) the final thread-crossing detections and the identified neighbor relations present in these detections. In the response images, a brighter color corresponds to a higher likelihood of a thread crossing being present in the canvas (according to our model). In (d), detected thread crossings are indicated by red crossings. The blue lines indicate the detected neighbor relations: to construct distance maps, distances that were measured between all neighboring thread crossings (i.e., over all blue lines). At locations where blue lines are absent, the distances are interpolated from neighboring thread crossings. (Figure is best viewed in color.)

thread-crossing presence near the locations $(x - d_x, y)$, $(x + d_x, y)$, $(x, y - d_y)$, and $(x, y + d_y)$, where d_x and d_y are the average distances between threads in the warp and weft directions, respectively. We employ a pictorial-structures model [16] that can exploit this information to also detect thread crossings for which little visual evidence is present (e.g., because the thread crossing is hardly visible due to the presence of lead white paint). Our pictorial-structures model computes the score s for thread-crossing presence based on the image patch $\mathbf{I}(x, y)$ extracted at location (x, y) as follows:

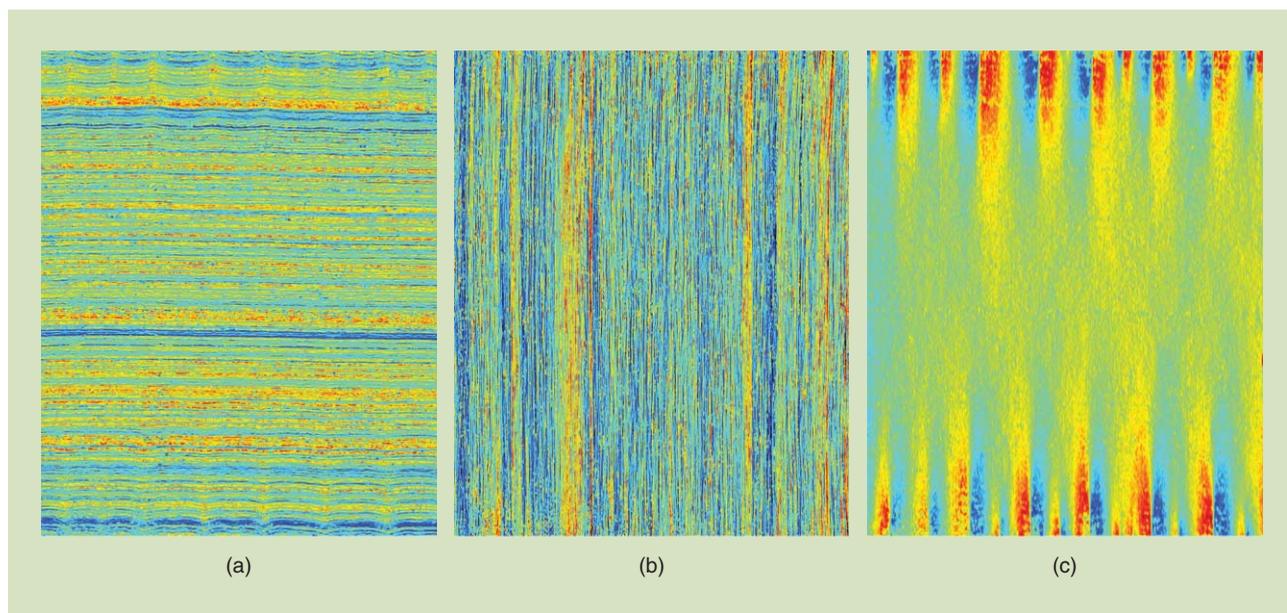
$$s(x, y; \theta) = p(z = 1 | \mathbf{I}(x, y); \theta) + \sum_{t_x} \max_{\delta_x, \delta_y} [p(z = 1 | \mathbf{I}(x + t_x d_x + \delta_x, y + \delta_y); \theta) - \alpha(\delta_x^2 + \delta_y^2)] + \sum_{t_y} \max_{\delta_x, \delta_y} [p(z = 1 | \mathbf{I}(x + \delta_x, y + t_y d_y + \delta_y); \theta) - \alpha(\delta_x^2 + \delta_y^2)],$$

where $\alpha \geq 0$ is a manually set discount factor and where $t_x \in \{-1, +1\}$ and $t_y \in \{-1, +1\}$. Intuitively, the score for a location is thus given by the sum of the likelihood for that location and the likelihood of the highest-scoring locations in a

four-lattice surrounding that location, where the score of those locations is discounted by the distance to their expected location. While the resulting score is not technically a likelihood, it may be employed in the same way. An example of a pictorial-structures score map is shown in Figure 4(c). The figure illustrates that incorporating prior knowledge on the typical structure of canvas greatly improves the performance of the thread-crossing model. The final thread-crossing detections are obtained by applying nonmaxima suppression on the score map. Nonmaxima suppression finds local maxima in the score map that are above a predefined threshold τ . An example of the resulting thread-crossing detections is shown in Figure 4(d). While some detection errors are present, the result in Figure 4(d) illustrates that the majority of thread-crossings and neighborhood relations is correctly identified.

ESTIMATING THREAD DISTANCES

After detecting the thread crossings, we need to identify which crossings are the warp and weft “neighbors” to be able to measure the distance between two weft threads or two warp threads at that



[FIG5] An illustration of the results of our canvas analysis algorithm on Nicolas Poussin's *Triumph of Bacchus*: (a) spacing between warp threads, (b) spacing between the weft threads, and (c) orientation of the warp threads. In the spacing maps, a blue color corresponds to a small distance between threads, while a red color corresponds to a large distance between threads. For the warp threads, thread spacings range between 1.2 and 1.9 mm, while for weft threads, thread spacings range between 0.8 and 1.45 mm. The warp orientation map shows strong cusping on the top and bottom, caused by the deformation of the canvas as it is placed on the stretcher. (Figure is best viewed in color.)

location. To this end, we center an anisotropic Gaussian distribution at each thread crossing location that has much more variance in either the warp or the weft direction (depending on whether we want to measure interweft or interwarp distances). The resulting density map tends to follow the direction of the threads in the canvas. The density maps may be further improved by rotating the Gaussians according to the estimated orientation of the threads at each location to become more robust to cusping, but for simplicity, we omit that in this study. To determine the neighbor of a thread crossing, we perform a large number of random walks that emanate from the thread crossing under investigation and terminate at the next thread crossing we encounter. The random walks are forced to go in a “forward” or “backward” direction, while using the density map to determine whether or not to move in the direction perpendicular to the thread. We construct a histogram over neighbor candidates that counts how often a random walk terminated in each thread crossing, and we select the crossing that has the highest count as the final neighbor candidate. The process is performed both in the “forward” and in the “backward” direction, and a neighborhood relation is only accepted if both thread crossings pick each other as neighbor to eliminate any inconsistencies (i.e., when the neighborhood relation is reciprocal). The detected thread-crossing relations are indicated by blue

lines in Figure 4(d). Finally, we estimate interthread distances on all locations where the thread identification procedure has a high confidence, while interpolating in low-confidence regions of the canvas and removing small outliers using a median filter. In a similar manner, we can measure the orientation of the thread connections to produce a thread orientation map. An example of the resulting distance and orientation maps (for both warp and weft threads) is shown in Figure 5.

MATCHING THREAD DISTANCE MAPS

To identify potential matches between different canvases based on the thread distance maps, we adopt an approach similar to that described by [17]. Specifically, we extract a small band of the distance map and take the median along this band (in the direction of the threads) to obtain an estimate of the thread distance signal. The thread distance signal is convolved with a Gaussian kernel to remove very fine-grained structure: empirically, we found matching is more accurate when performed based on features in the thread distance signal that live on a coarser scale. We match the thread-distance signals of two canvases by sliding one signal over the other (enforcing a minimum overlap), while measuring the mean absolute distance between the signals in the overlapping region. We use mean average distance as it is less sensitive to outliers than squared

**WE SHOW THE MERITS OF
THREAD-LEVEL CANVAS
ANALYSIS BY USING IT TO
STUDY A COLLECTION OF
THREE ALLEGED NICOLAS
POUSSIN PAINTINGS AS
WELL AS A SMALL COLLECTION
OF PAINTINGS BY
VINCENT VAN GOGH.**

errors. The match is repeated for a flip of one of the thread signals as one of the canvases may have been “upside down” compared to the other. A match is only accepted if one of the two minimum mean average distances is below a certain threshold.

EXPERIMENT 1: NICOLAS POUSSIN

We used our canvas-analysis algorithm to study a collection of three Nicolas Poussin paintings that were studied before by [9]: 1) *Triumph of Pan*, 2) *Triumph of Bacchus*, and 3) *Triumph of Silenus*. This set of paintings is of particular art-historical interest because the three paintings were part of single commissioning in 1636 to Cardinal de Richelieu for the Cabinet du Roi in his castle in Poitou, France. Their authenticity has been subject to strong debate: some have considered *Bacchus* to be a copy [18], [19], but most Poussin scholars now believe that *Bacchus* and *Pan* are authentic Poussin paintings. *Silenus*, however, is considered to be an early copy by its owners, the National Gallery London (Poussin's Bacchanals quickly became very popular, with the first copies being produced as early as 1665. For instance, at least seven known copies of *Bacchus* exist today). Recent canvas analysis results have challenged this belief by finding a canvas match between all three *Triumph*-paintings but were inconclusive because they were unable to perform thread-level canvas analysis. We obtained digital versions of radiographs (scanned at 600 dpi, 500 dpi, and 1,200 dpi for *Triumph of Pan*, *Triumph of Bacchus*, and *Triumph of Silenus*, respectively) and stitched them into whole-painting radiographs using algorithms described in [20]. Thereafter we manually annotated a total of 11,954 thread

crossings in these radiographs and trained our thread-crossing detector on these manually annotated positive examples (negative examples were sampled randomly from the same radiographs).

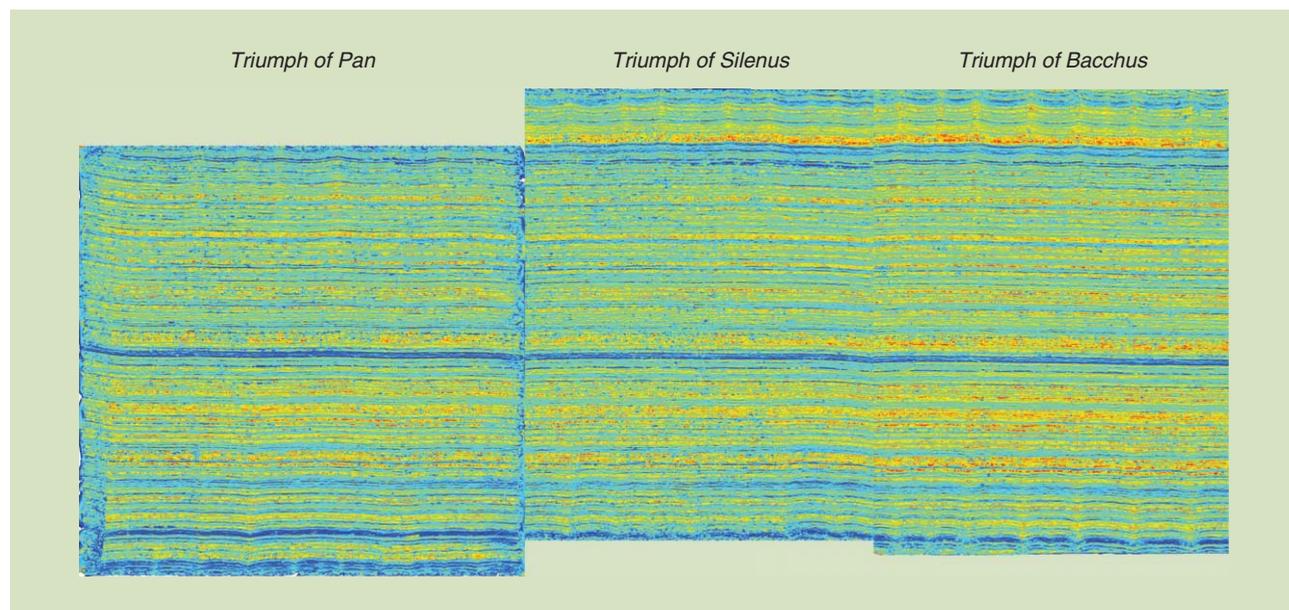
We set the value of the L2-regularization parameter in the logistic regression, σ^2 , by performing a grid search guided by the classification error on a small held-out validation set. The error of our model on the validation set was approximately 9%;

most of these errors were likely due to the set of negative examples containing some actual thread crossings by chance. The average thread distance parameters d_x and d_y were estimated by running the entire canvas-analysis procedure without the pictorial structures and taking the median interthread distance in the warp and weft direction. The nonmaxima suppression

step used a window size of 5×5 pixels, and a threshold $\tau = 0.4$ (on a scale from zero to one). The final thread distance maps were cleansed with a 7×3 or 3×7 median filter (depending on the orientations of the threads being analyzed). For weft maps, we removed distance values below 0.85 mm and above 1.45 mm from the map, while for warp maps, we removed distance values below 1.2 mm and above 1.9 mm.

Figure 5 presents the results of our analysis of *Triumph of Bacchus*. Figure 6 presents the results produced by our thread-level canvas analysis algorithm after matching the three Poussin paintings. Different colors correspond to different spacings between individual threads. The results presented in the figure provide very strong evidence that all three canvases originated from the same roll. The results are in line with earlier results from automatic and manual canvas analyses of these three

**THREAD-CROSSING DETECTION
CAN BE PERFORMED WITH
VERY HIGH ACCURACY
BECAUSE THREAD
CROSSING CORRESPONDS TO
VISUALLY SALIENT LOCATIONS
IN CANVAS RADIOGRAPHS.**



[FIG6] The results of our thread-level canvas analysis of the three *Triumph* paintings. Different colors indicate the distance between (detected) neighboring thread crossings. (The figure is best viewed in color.)

paintings [9] but provide stronger evidence because the evidence is on the level of individual threads and not on the level of multiple-thread averages. Indeed, the presented results make it highly unlikely that *Silenus* was copied 30 years later in a different location (Poussin was working in Rome, Italy, whereas a copyist likely would have worked in France), which strongly suggests that the current art-historical description of the three paintings needs to be revised. We leave such art-historical interpretations to other scholars; they are outside the scope of this work.

EXPERIMENT 2: VINCENT VAN GOGH

We also performed analyses of a small collection of paintings by Vincent van Gogh. Unlike the canvases of 17th-century Poussin, 19th-century van Gogh used canvas produced by the textile industry that has much finer threads and smaller irregularities in thread spacing [10]. Moreover, because van Gogh applies very thick paint layers, the thread structure is much harder to see in the X-rays. As a result, thread-level canvas analysis of van Gogh paintings is substantially harder than the analysis of Poussin paintings.

A large collection of roughly 180 van Gogh paintings has been studied intensively in the context of the Thread Count Automation Project [10], the goal of which is to assign all van Gogh paintings to a particular roll, as this may provide information on the order in which van Gogh made his paintings.

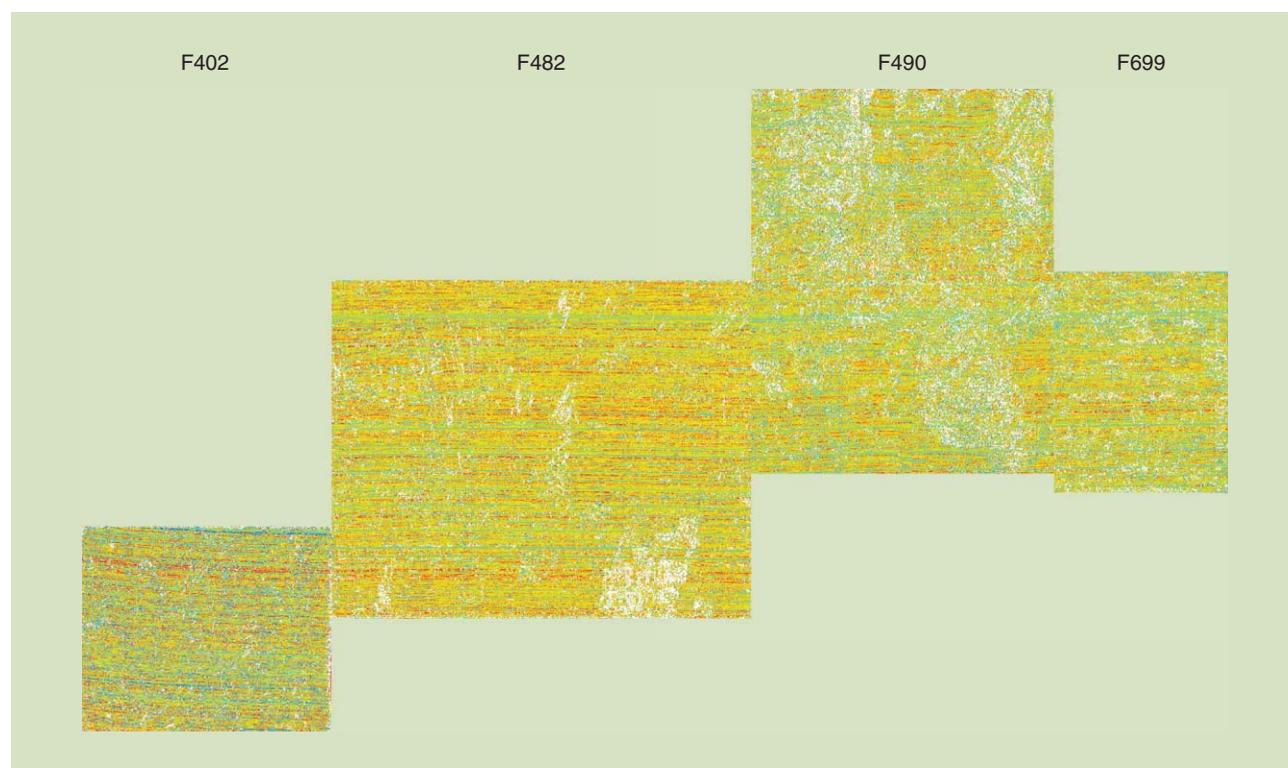
For this study, we had access to a small collection of ten radiographs van Gogh paintings that were scanned at 600 dpi:

- F402 *Two White Butterflies*
- F482 *Bedroom in Arles*
- F490 *Mother Roulin with Her Baby*
- F511 *Orchard in Blossom*
- F633 *The Good Samaritan*
- F692 *The Thresher*
- F699 *The Shepherdess*
- F720 *Enclosed Wheat Field with Rising Sun*
- F734 *The Garden of Saint-Paul Hospital*
- F822 *The Cows*.

(The F-numbers are the catalogue numbers used by the van Gogh Museum in Amsterdam, The Netherlands). Some of the ten canvases are surmised to originate from the same roll but the results of current analyses are inconclusive. Our analysis results in a group of at least four matching canvases, as illustrated in Figure 7. An extensive study on all 180 van Gogh paintings [10] is planned for a future work.

CONCLUSIONS AND OUTLOOK

We have presented a novel canvas-analysis approach that is able to perform thread-level analyses of canvas. We believe the method has two main advantages over prior work: 1) it provides more conclusive evidence on whether or not two patches of canvas have the same thread patterns and 2) it is easier for art experts to understand exactly what is being measured. We believe the second advantage is essential to get canvas-analysis



[FIG7] An illustration of canvas weave matches between four van Gogh paintings: 1) F402 *Two White Butterflies*, 2) F482 *Bedroom in Arles*, 3) F490 *Mother Roulin with Her Baby*, and 4) F699 *Shepherdess (after Millet)*. Different colors indicate the distance between (detected) neighboring thread crossings. In white regions, hardly any thread crossings were detected because the crossing signal was obfuscated by thick paint layers; these regions were ignored in the thread spacing measurements. (The figure is best viewed in color.)

technology widely used: showing art experts visualizations such as those in Figure 4 allows art experts to understand the analysis process, to identify potential errors in the measurements, and to manually correct such errors when desired, and to assess whether thread spacings are a good surrogate for thread thicknesses for the canvas a hand.

A substantial drawback of the proposed approach is that a trained thread-crossing detection model is likely only applicable to canvas of a similar type that was imaged under similar conditions: for instance, models trained on the Poussin paintings do not work well on the van Gogh paintings because van Gogh's canvases have much finer threads, which results in a different visual appearance of thread crossings. This implies that to apply our approach to a new type of canvas, it may be necessary to manually annotate a few hundred thread crossings for that canvas type. To resolve this problem, it would be very useful to establish a database with a large collection of canvas radiographs along with a crowdsourcing annotation tool. Such a database would not only facilitate systematic comparisons between canvas-analysis algorithms, but it would also allow for training thread-crossing detectors that can be applied to a wide variety of canvas types. Similar data-gathering and annotation efforts have proven instrumental in improving the state of the art in other computer-vision problems, such as object recognition [21].

ACKNOWLEDGMENTS

We thank the National Gallery London, the Nelson-Atkins Museum of Art in Kansas City, Missouri, and the van Gogh Museum in Amsterdam for generously providing the X-rays that were used in this study. In particular, we are indebted to Mary Schafer, Nicole Meyers, Teio Meedendorp, and Louis van Tilborgh. The manual annotation of thread crossings for the Poussin canvases was done by Travis Sawyer of the University of Arizona. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement 600849 (INSIDDE). We also acknowledge support from The Netherlands Institute for Advanced Study.

AUTHORS

Laurens van der Maaten (lvdmaaten@gmail.com) is an assistant professor in computer vision and machine learning at Delft University of Technology, The Netherlands. Previously, he worked as a postdoctoral researcher at the University of California, San Diego; as a Ph.D. student at Tilburg University, The Netherlands; and as a visiting Ph.D. student at the University of Toronto. His research interests include painting analysis, deformable template models, dimensionality reduction, classifier regularization, and tracking.

Robert Erdmann (erdmann@arizona.edu) is a senior scientist at the Rijksmuseum, Amsterdam, and is the Rijksmuseum Professor of Conservation Science at the University of Amsterdam, The Netherlands. He is also special professor for the visualization of art history at Radboud University in Nijmegen, The Netherlands.

He was formerly an associate professor at the University of Arizona, Department of Materials Science and Engineering and the Program in Applied Mathematics. He is a member of the Bosch Research and Conservation Project and is the director of the Thread Counting Automation Project. His interests include computational materials science, microstructural analysis, and computational art history and art conservation.

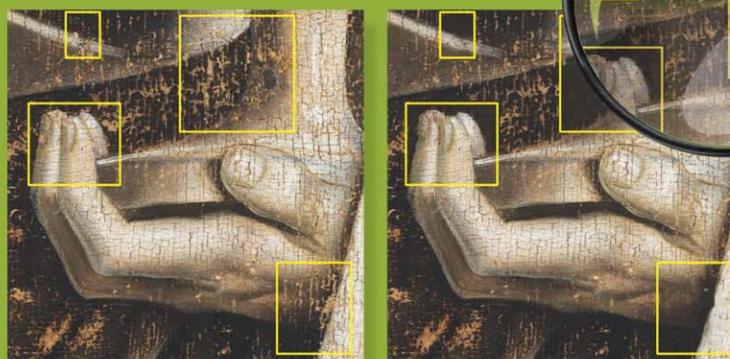
REFERENCES

- [1] A. Adam, P. Planken, S. Meloni, and J. Dik, "Terahertz imaging of hidden paint layers on canvas," *Opt. Express*, vol. 17, no. 5, pp. 3407–3416, 2009.
- [2] J. Dik, K. Janssens, G. V. D. Snickt, L. van der Loeff, K. Rickers, and M. Cotte, "Visualization of a lost painting by Vincent van Gogh using synchrotron radiation based X-ray fluorescence elemental mapping," *Anal. Chem.*, vol. 80, no. 16, pp. 6436–6442, 2008.
- [3] C.R. Johnson, Jr., E. Hendriks, I. Berezhnoy, E. Brevdo, S. Hughes, I. Daubechies, J. Li, E. Postma, and J. Wang, "Image processing for artist identification," *IEEE Signal Process. Mag.*, vol. 25, no. 4, pp. 37–48, 2008.
- [4] J. Li, L. Yao, E. Hendriks, and J. Wang, "Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1159–1176, 2012.
- [5] S. Lyu, D. Rockmore, and H. Farid, "A digital technique for art authentication," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 101, no. 49, pp. 17006–17010, 2004.
- [6] W. Smith, C. Tyler, and D. Stork, "In search of Leonardo: Computer based facial image analysis of renaissance artworks for identifying Leonardo as a portrait subject," *Hum. Vis. Electron. Imaging*, vol. 8291, no. 82911D, pp. 1–7, 2012.
- [7] E. V. der Wetering, *Rembrandt: The Painter at Work*. Oakland, CA: Univ. of California Press, 2000.
- [8] E. Hendriks, C. R. Johnson, Jr., and D. Johnson, "Interpreting canvas weave matches," *Art Matters: Int. J. Techn. Art History*, vol. 5, pp. 53–61, 2013.
- [9] R. Erdmann, C. R. Johnson, Jr., M. Schafer, J. Twilley, and T. Sawyer, "Reuniting Poussin's Bacchanals painted for Cardinal Richelieu through quantitative canvas weave analysis," *J. Amer. Inst. Conserv.*, 2013. [Online]. Available: http://erg.mse.arizona.edu/Erdmann_Reuniting_Poussins_Bacchanals.pdf
- [10] L. van Tilborgh, T. Meedendorp, E. Hendriks, D. Johnson, and C. R. Johnson, Jr., "Weave matching and dating of van Gogh's paintings: An interdisciplinary approach," *Burlington Mag.*, vol. 154, no. 1307, pp. 112–122, Feb. 2012.
- [11] W. Liedtke, C. R. Johnson, Jr., and D. Johnson, "Canvas matches in Vermeer: A case study in the computer analysis of fabric supports," *Metropolitan Museum J.*, vol. 47, pp. 99–106, 2012.
- [12] B. Tobin, "Automated canvas spacing analysis," Master's thesis, Dept. Materials Science Eng., Univ. Arizona, Tucson, AZ, 2014.
- [13] P. P. d'Ors, C. R. Johnson, Jr., and D. Johnson, "Velázquez in Fraga: A new hypothesis about the portraits of el Primo and Philip IV," *Burlington Mag.*, vol. 154, no. 1314, pp. 620–625, 2012.
- [14] D. Johnson, C. Johnson, and R. Erdmann, "Weave analysis of paintings on canvas from radiographs," *Signal Process.*, vol. 93, no. 3, pp. 527–540, 2013.
- [15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2005, pp. 886–893.
- [16] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [17] D. Johnson, C. R. Johnson, Jr., and E. Hendriks, "Matching canvas weave patterns from processing x-ray images of master paintings," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, 2010, pp. 958–961.
- [18] A. Blunt, *The Paintings of Nicolas Poussin: A Critical Catalogue*. London, U.K.: Phaidon, 1966.
- [19] A. Blunt, "French seventeenth century painting: The literature of the last ten years," *Burlington Mag.*, vol. 124, no. 956, p. 707, 1982.
- [20] D. Johnson, R. Erdmann, and C. R. Johnson, Jr., "Whole-painting canvas analysis using high- and low-level features," *Proc. 36th Int. Conf. Acoustics, Speech, Signal Processing*, 2011, pp. 969–972.
- [21] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge." [Online]. Available: <http://arxiv.org/abs/1409.0575>

[Nanne van Noord, Ella Hendriks, and Eric Postma]

Toward Discovery of the Artist's Style

© GHENT, KATHEDRALE KERKFABRIEK; PHOTO COURTESY OF KIK-IPPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

[Learning to recognize artists by their artworks]

Author attribution through the recognition of visual characteristics is a commonly used approach by art experts. By studying a vast number of artworks, art experts acquire the ability to recognize the unique characteristics of artists. In this article, we present an approach that uses the same principles to discover the characteristic features that determine an artist's touch. By training a convolutional neural network (PigeoNET) on a large collection of digitized artworks to perform the task of automatic artist attribution, the network is encouraged to discover artist-specific visual features. The trained network is shown to be capable of attributing previously

unseen artworks to the actual artists with an accuracy of more than 70%. In addition, the trained network provides fine-grained information about the artist-specific characteristics of spatial regions within the artworks. We demonstrate this ability by means of a single artwork that combines characteristics of two closely collaborating artists. PigeoNET generates a visualization that indicates for each location on the artwork who is the most likely artist to have contributed to the visual characteristics at that location. We conclude that PigeoNET represents a fruitful approach for the future of computer-supported examination of artworks.

INTRODUCTION

Identifying the artist of an artwork is a crucial step in establishing its value from a cultural, historical, and economic perspective. Typically, the attribution is performed by an experienced art expert

Digital Object Identifier 10.1109/MSP.2015.2406955

Date of publication: 15 June 2015

with a long-standing reputation and an extensive knowledge of the features characteristic of the alleged artist and contemporaries.

Art experts acquire their knowledge by studying a vast number of artworks accompanied by descriptions of the relevant characteristics (features) [1]. For instance, the characteristic features of Vincent van Gogh during his later French period include the outlines painted around objects, complementary colors [2], and rhythmic brush strokes [3]. As Van Dantzig [4] claimed in the context of his *Pictology* approach, describing works by an artist in terms of visual features enables the attribution of works to artists (see also [5]).

The advent of computers and high-resolution digital reproductions of artworks gave rise to attempts to partially automate the attribution of artworks [6]–[8]. Given the appropriate visual features, machine-learning algorithms may automatically attribute artworks to their artists. As was (and still is) common practice in traditional machine learning, feature engineering, i.e., finding or defining the appropriate features, is critical to the success of the automatic attribution task. Close cooperation with art historians and conservators facilitated the feature engineering for artist attribution, which led to promising results in the automatic attribution of artworks by van Gogh and his contemporaries [3], [6], [9], [10], highlighting the value of automatic approaches as a tool for art experts.

Despite the success of feature engineering, these early attempts were hampered by the difficulty to acquire explicit knowledge about all the features associated with the artists of artworks. Understandably, the explicit identification of characteristic features posed a challenge to art experts, because (as is true for most experts) their expertise is based on tacit knowledge that is difficult to verbalize [11]. By adopting a method capable of automatically recognizing the characteristics that are known to be important for the task at hand, the tacit knowledge of art experts may be operationalized [12].

Feature learning is an alternative to feature engineering that learns features directly from the data [12]. Feature learning is much more data intensive than feature engineering, because it requires a large number of examples to discover the characteristic features. In recent years, feature learning has shown great promise by taking advantage of deep architectures, machine learning methods inspired by biological neural networks. A typical example of a deep architecture is a convolutional neural network, which, when combined with a powerful learning algorithm, is capable of discovering (visual) features. Convolutional neural networks outperform all existing learning algorithms on a variety of very challenging image classification tasks [13]. To our knowledge, convolutional neural networks have not yet been applied for automated artist attribution. The objective of this article is to present a novel and transparent way of performing automatic artist attribution of artworks by means of convolutional neural networks.

IDENTIFYING THE ARTIST OF AN ARTWORK IS A CRUCIAL STEP IN ESTABLISHING ITS VALUE FROM A CULTURAL, HISTORICAL, AND ECONOMIC PERSPECTIVE.

When using only visual information, the following question may be raised: Is automatic artist attribution possible at all? It has been frequently argued by scholars working in the art domain that semantic or historical knowledge, as well as technical and analytical information are pivotal in the attribution of artworks. The feasibility of image-based automatic artist attribution is supported by biological studies. Pigeons [14] and honeybees [15] can be successfully trained to discriminate between artists, with pigeons correctly attributing an art work in 90% of the cases in a binary Monet–Picasso attribution task. This shows that a visual system without higher cognitive functions is capable of learning the visual characteristics present in artworks. While it is unlikely that a perfect result can be achieved without incorporating additional information, these findings do pave the way for an attribution approach that learns to recognize visual features from data rather than from prior knowledge.

In this article, we present PigeoNET, a convolutional neural network corresponding to the AlexNET architecture described in [13] to which we added a visualization component due to [16]. PigeoNET is applied to an artist attribution task by training it on artworks. As such, PigeoNET performs a task similar to the pigeons in [14], by performing artist attribution based solely on visual characteristics. This implies that, in addition to authorship, PigeoNET may also take visual characteristics into consideration that relate indirectly to the artist (e.g., the choice of materials or tools used by the artist) or that are completely unrelated to the artist (e.g., reproduction characteristics such as lighting and digitization procedure). To ensure that the visual characteristics on which the task is solved by PigeoNET make sense, human experts are needed to assess the relevance of the acquired mapping from images of artworks to artists. Our visualization method allows for the visual assessment by experts of the characteristic regions of artworks.

In our artist attribution experiments, we consider three sources of variation in the training set and assess their effects on attribution performance: 1) heterogeneity versus homogeneity of classes (types of artworks), 2) number of artists, and 3) number of artworks per artist.

After training, the performance of PigeoNET will be assessed in two ways: 1) by determining how well it attributes previously unseen artworks and 2) by generating visualizations that reveal artwork regions characteristic of the artist, or in case of artworks that are likely created by two or more artists, generating visualizations that reveal which regions belong to which artist, and could aid in answering outstanding art historical questions.

PigeoNET

A convolutional neural network can learn to recognize the visually characteristic features of an artist by adapting filters to respond to the presence of these features in an image [17]. The filters are adapted to respond to a feature by adjusting the parameters, and weights, of the filters until a suitable configuration is found. The

proper weights for this configuration are obtained by means of a learning algorithm called *back-propagation* [18], which requires no prior knowledge other than the input images and a label (e.g., the artist who created it). In the case of artist attribution, the network will learn to recognize features that are regarded as characteristic of a certain artist, allowing us to discover these characteristics. PigeoNET is a convolutional neural network designed to learn the characteristics of artists and their artworks, so as to recognize and identify their authorship.

The filters in a convolutional neural network are grouped into layers, where the first layer is directly applied to images, and subsequent layers to the responses generated by previous layers. By stacking layers to create a multilayer architecture the filters can respond to increasingly complex features with each subsequent layer. The filters in the initial layers respond to low-level visual patterns, akin to Gabor filters [19], whereas the final layers of filters respond to visual characteristic features specific to artists.

Because convolution is used to apply the filters to an image, or the response of a previous layer, the layers of filters are referred to as *convolutional layers*. The advantage of a convolutional layer, over a traditional neural network layer, is that the weights are shared, allowing the adaptive filters to respond to characteristic features irrespective of their position or location in the input [18]. To learn a mapping from the filter responses to a certain artist the convolutional layers are, typically, followed by a number of fully connected layers that translate the presence and intensity of the filter responses to a single certainty score per artist. The certainty score for an artist is high whenever the responses for filters corresponding to that artist are strong, conversely, the certainty score is low when the filter responses are weak or nonexistent. Thus, an unseen artwork can be attributed to an artist for whom the certainty score is the highest.

PIGEONET IS A CONVOLUTIONAL NEURAL NETWORK DESIGNED TO LEARN THE CHARACTERISTICS OF ARTISTS AND THEIR ARTWORKS, SO AS TO RECOGNIZE AND IDENTIFY THEIR AUTHORSHIP.

VISUALIZATION OF ARTIST-CHARACTERISTIC REGIONS

While PigeoNET's attribution of an artwork is based on the entire artwork, regions containing visual elements characteristic for an artist are assigned more weight than others to achieve a correct attribution [20]. To increase our understanding of the attribution performed by PigeoNET, we aim to visualize such artist-characteristic regions. Several methods have been proposed for visualizing trained convolutional neural networks [16], [21] and other layered algorithms [22]. We adopt the occlusion sensitivity testing method proposed by [16] for obtaining visualizations of artist-characteristic regions, which can be considered a weakly supervised

localization method. By systematically occluding a small image region of an artwork, the importance of the occluded region is determined by observing the change in the certainty score for the correct artist. When an occluded region is very important (or highly characteristic) for correctly identifying the artist, there will be a significant drop in the certainty score generated by PigeoNET. Inversely, occluding a region that is atypical for the correct artist will result in an increase in the certainty score. A region for which occlusion results in a drop of the certainty score is considered characteristic for the artist under consideration. This approach to creating visualizations allows us to show the approximate areas of an artwork which are representative of an artist.

As an illustration, Figure 1 depicts *The Feast of Acheloüs* by two artists; Peter Paul Rubens painted the persons and Jan Brueghel the Elder painted the scenery [23]. Although there is no single correct artist, the certainty score for Brueghel would decrease if the scenery were to be occluded, whereas the certainty score for Rubens would drop if the figures were occluded. Even when only part of the figures or part of the scenery were to be occluded, we see a drop in confidence scores. In a similar vein, when even smaller regions of the painting have been occluded, it becomes possible to identify important regions on a much more detailed scale.

AUTHOR ATTRIBUTION EXPERIMENT

The goal of an artist attribution task is to attribute an unseen artwork to the artist who created it. To be able to perform this task adequately, PigeoNET needs to discover features that distinguish an artist from other artists but especially to discover features that are characteristic of each artist. In the rest of this section we will discuss the data set, network architecture, training procedure, evaluation procedure, and the results.

EXPERIMENTAL SETUP

DATA SET

The characteristic features of an artist can be discovered by studying artworks that are representative of that artist. Yet,



[FIG1] Peter Paul Rubens and Jan Brueghel the Elder's work: *The Feast of Acheloüs* (45.141) in The Metropolitan Museum of Art's *Heilbrunn Timeline of Art History* (2000–present). (Image used courtesy of The Metropolitan Museum of Art; <http://www.metmuseum.org/toah/works-of-art/45.141>.)

obtaining a sufficiently large sample of such images is problematic, given the lack of (automatic) methods and criteria to determine whether an artwork is representative. A commonly utilized approach to circumvent the need for a representative sample is to take a very large sample. As such, a data set that contains a large number of images, and a large number of images per artist, is required.

The Rijksmuseum Challenge data set [24] consists of 112,039 digital photographic reproductions of artworks by 6,629 artists exhibited in the Rijksmuseum in Amsterdam, The Netherlands. All artworks were digitized under controlled settings. Within the set there are 1,824 different types of artworks and 406 annotated materials, such as paper, canvas, porcelain, iron, and wood. To our knowledge, this is the largest available image data set of artworks and the only data set that meets our requirements.

We divided the Rijksmuseum Challenge data set into a training, validation, and test set (cf. [24]). In this article, these sets are used to train PigeoNET, to optimize the hyperparameters, and to evaluate the performance of PigeoNET on unseen examples, respectively. The data set contains a number of artworks that lack a clear attribution, these are labeled as either “Anonymous” or “Unknown.” We chose to exclude these artworks because our objective is to relate visual features to specific artists.

While the Rijksmuseum Challenge data set contains a large number of images of many different types of artworks by a large number of artists, there are many artists for whom only a few artworks are available or artists who have created many different types of artworks. As stated earlier, these variations might influence the performance of PigeoNET in nonobvious ways. To this end, we consider the following three sources of variation: 1) heterogeneity versus

[TABLE 1] A LIST OF THE 34 ARTISTS WITH AT LEAST 256 ARTWORKS AND THE DISTRIBUTION OF ARTWORKS OVER MAIN TYPES (PRINTS, DRAWINGS, AND OTHER).

NUMBER	NAME	PRINTS	DRAWINGS	OTHER
1	HEINRICH ALDEGREVER	347	27	—
2	ERNST WILLEM JAN BAGELAAR	400	27	—
3	BOËTIUS ADAMSZ. BOLSWERT	592	—	—
4	SHELTE ADAMSZ. BOLSWERT	398	—	—
5	ANTHONIE VAN DEN BOS	531	3	—
6	NICOLAES DE BRUYN	515	2	—
7	JACQUES CALLOT	1,008	4	1
8	ADRIAEN COLLAERT	648	1	—
9	ALBRECHT DÜRER	480	9	2
10	SIMON FOKKE	1,177	90	—
11	JACOB FOLKEMA	437	4	3
12	SIMON FRISIUS	396	—	—
13	CORNELIS GALLE (I)	421	—	—
14	PHILIPS GALLE	838	—	—
15	JACOB DE GHEYN (II)	808	75	10
16	HENDRICK GOLTZIUS	763	43	4
17	FRANS HOGENBERG	636	—	4
18	ROMEYN DE HOOGHE	1,109	5	5
19	JACOB HOUBRAKEN	1,105	42	1
20	PIETER DE JODE (II)	409	1	—
21	JEAN LEPAUTRE	559	—	1
22	CASPAR LUYKEN	359	18	—
23	JAN LUYKEN	1,895	33	—
24	JACOB ERNST MARCUS	372	23	2
25	JACOB MATHAM	546	4	—
26	MEISSNER PORZELLAN MANUFAKTUR	—	—	1,003
27	PIETER NOLPE	344	2	—
28	CRISPIJN VAN DE PASSE (I)	841	15	—
29	JAN CASPAR PHILIPS	401	17	—
30	BERNARD PICART	1,369	132	3
31	MARCANTONIO RAIMONDI	448	2	—
32	REMBRANDT HARMENSZ. VAN RIJN	1,236	119	29
33	JOHANN SADELER (I)	578	1	—
34	REINIER VINKELES	573	50	—

homogeneity of classes (types of artworks), 2) number of artists, and 3) number of artworks per artist.

Two main types of subsets were defined to assess the effect of heterogeneity versus homogeneity of artworks: type A (for “All”)

[TABLE 2] AN OVERVIEW OF SUBSETS AND THE NUMBER OF TRAINING, VALIDATION, AND TEST IMAGES PER SUBSET.

SUBSETS	NUMBER OF EXAMPLES PER ARTIST	NUMBER OF ARTISTS (CLASSES)	NUMBER OF TRAINING IMAGES	NUMBER OF VALIDATION IMAGES	NUMBER OF TEST IMAGES
A	10	958	56,024	7,915	15,860
	64	197	37,549	5,323	10,699
	128	97	28,336	4,063	8,058
	256	34	17,029	2,489	4,838
P1	10	673	44,539	6,259	12,613
	64	165	31,655	4,484	8,983
	128	78	23,750	3,408	6,761
	256	29	14,734	2,171	4,200
P2	128	26	3,328	1,209	2,277
	128	39	4,992	1,521	2,970
	128	52	6,656	2,160	4,341
	128	78	9,984	3,408	6,761
P3	10	78	780	3,408	6,761
	64	78	4,992	3,408	6,761
	128	78	9,984	3,408	6,761

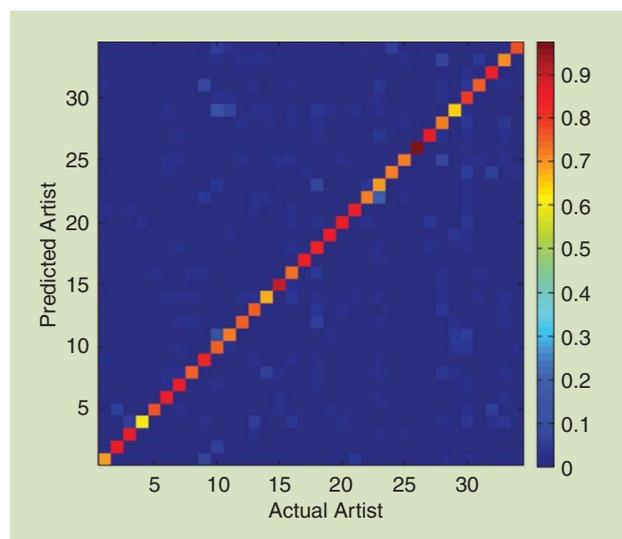
The subsets are labeled by their types. Type A (“All”) are subsets containing varying artworks, examples, and examples per artist. Type P (“Prints”) refers to subsets of prints only. P1: varying numbers of artworks, examples, and examples per artist. P2: number of examples constant (128). P3: number of artists constant (78). For A and P1, the numbers of examples per artists represent the minimum numbers, while for P2 and P3, these numbers represent the exact number of artworks per artist.

[TABLE 3] THE MCA FOR THE ARTIST ATTRIBUTION TASK ON THE 15 DATA SUBSETS.

SUBSETS	NUMBER OF EXAMPLES PER ARTIST	NUMBER OF ARTISTS (CLASSES)	MCA
A	10	958	52.5
	64	197	68.2
	128	97	74.5
	256	34	78.3
P1	10	673	60.0
	64	165	70.2
	128	78	73.3
	256	29	78.8
P2	128	26	63.9
	128	39	55.6
	128	52	52.7
	128	78	52.0
P3	10	78	13.1
	64	78	38.0
	128	78	52.0

Bold values indicate the best result per type; the overall best result is underlined.

and type P (for “Prints”), respectively. As is evident from Table 1, prints form the majority of artworks in the Rijksmuseum Challenge data set. The homogeneous type of subsets (P) has three forms: P1, P2, and P3. Subsets of type P1 have varying numbers of artists and artworks per artist (as is the case for A). Subsets of type P2 have a fixed number of artworks per artist. Finally, subsets of type P3 have a fixed number of artists. We remark that the number of examples per artist for the subsets in A and P1 are minimum values. For very productive artists these subsets may include more artworks. For subsets of types P2 and P3, the number of examples is exact and constitutes a random sample of the available works per artist. A detailed overview of the resulting 15 subsets is listed in



[FIG2] A confusion matrix for all artists with at least 256 training examples of all artwork types. The rows represent the artist estimates and the columns the actual artists. Row and column numbers (from left to right and from bottom to top) correspond to those as listed in Table 2.

Table 2. For the heterogeneous subset of at least 256 artworks of type A, Table 2 provides a more detailed listing that specifies the three most prominent categories: “Prints,” “Drawings,” and “Other.” (The largest subsets for P2 and P3 are identical, but are reported twice for clarity.) The “Other” category includes a variety of different artwork types, including 35 paintings.

All images were down-sampled to 256×256 pixels following the procedure described in [13], to adhere to the fixed input size requirement of the network architecture, and are normalized at runtime by subtracting the mean image as calculated on the training set.

ARCHITECTURE

The architecture of PigeoNET is based on the Caffe [25] implementation (see http://caffe.berkeleyvision.org/model_zoo.html) of the network described in [13], and consists of five convolutional layers and three fully connected layers. The number of output nodes of the last fully connected layer is equal to the number of artists in the data set, ranging from 958 to 26 artists.

TRAINING

An effective training procedure was used (cf. [13]), in that the learning rate, momentum, and weight decay hyperparameters were assigned the values of 10^{-2} , 0.9, and $5 \cdot 10^{-4}$. The learning rate was decreased by a factor ten whenever the error on the validation set stopped decreasing. The data augmentation procedure consisted of random crops and horizontal reflections. While orientation is an important feature, to detect authorship the horizontal reflections were used to create a larger sample size, as it effectively doubles the amount of available training data, providing PigeoNET with sufficient data from which to learn, while possibly negatively impacting PigeoNET’s ability to pick up on orientation clues to perform classification. In contrast to [13], only a single crop per image was used during training, with crops of size 227×227 pixels, and the batch size was set to 256 images per batch.

All training was performed using the Caffe framework [25] on a NVIDIA Tesla K20m card and took between several hours and several days, depending on the size of the subset.

EVALUATION

The objective of the artist attribution task is to identify the correct artist for each unseen artwork in the test set. To this end the performance is measured using the mean class accuracy (MCA), which is the average of the accuracies for all artists. This makes sure that the overall performance is not heavily biased by the performance on a single artist.

During testing the final prediction is averaged over the output of the final softmax layer of the network for ten crops per image. These crops are the four corner patches and the central patch plus their horizontal reflections.

RESULTS

The results of the artist attribution task are listed in Table 3. The results on the artist attribution task show that the three sources of

variation, [heterogeneity versus homogeneity of classes (types of artworks), number of artists, and number of artworks per artist] affect the performance in different ways. The effect of heterogeneity versus homogeneity can be assessed by comparing the results for A and P1. The results obtained with P1 are slightly better than those obtained with A (except for 128 examples per artist). However, A and P1 differ also in number of artists, which is likely to affect the performances as is evident from the results on P2 and P3.

The total number of artists (P2) and the number of examples per artist (P3) have a more prominent effect on the attribution performance of PigeoNET. Increasing the number of artists while keeping the number of examples per artist constant (as in P2) leads to a decrease in performance. With more examples per artist (P3) the performance increases tremendously.

Our results suggest that the effects of the number of artists and the number of examples per artist are closely related. This agrees with the findings reported in [13] and leads to the observation that by considering more examples per artist the number of artists to be modeled can be increased.

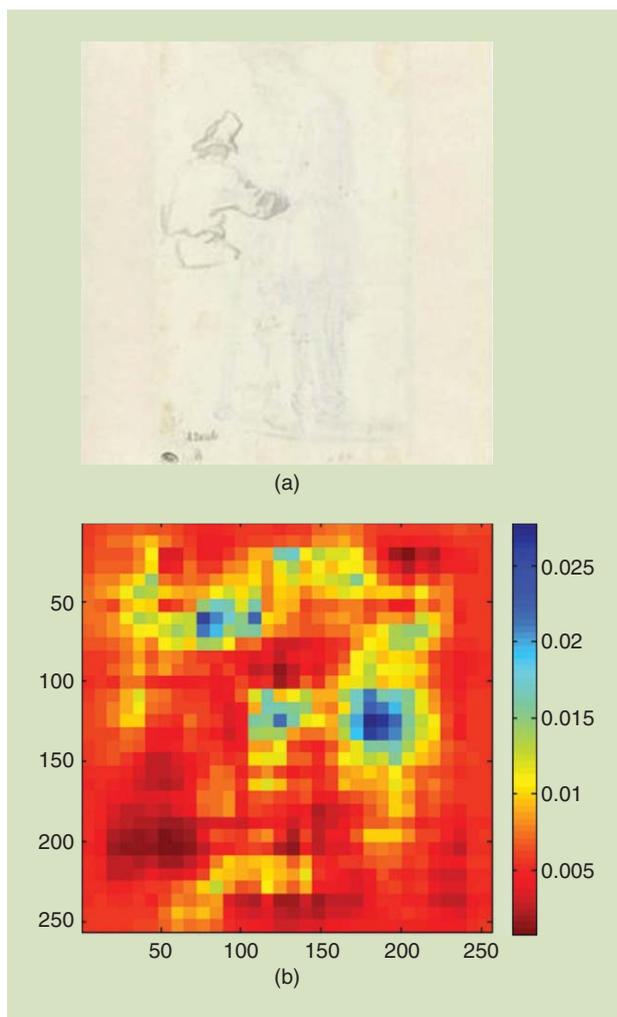
The subsets of type A are comparable to the subsets used in [24], who obtain a comparable MCA of 76.3 on a data set containing 100 artists using SIFT (which stands for “scale-invariant feature transform”) features, Fisher vectors, and one-versus-rest classification.

Figure 2 shows a visualization of the confusion matrix for the subset with at least 256 examples of all artwork types. The rows and columns correspond to the artists in Table 2. The rows represent the artist estimates by PigeoNET, the columns the actual artists. The diagonal entries represent correct attributions, which are color coded.

Upon further analysis of the results for the 256 example subset (A) of all artwork types it can be observed that the best artist-specific classification accuracy (97.5%) is obtained for Meissener Porzellan Manufaktur, a German porcelain manufacturer (class 26). Among the different types of artworks in the data set, these porcelain artworks are visually the most distinctive. Given that the visual characteristics of porcelain differ considerably from all other artworks in the data set, it is not surprising that the highest classification accuracy is achieved for this class.

The worst artist-specific classification accuracy (60.6%) is achieved for Schelte Bolswert (class 4), as indicated by the yellow square on the diagonal in the confusion matrix (fourth row from below, fourth column from left). The low accuracy may be partially explained by the confusion between Schelte Bolswert and his older brother and instructor Boëtius Bolswert (class 3). Yet, because the classification accuracy for Boëtius Bolswert (86.3) seems much less affected by the confusion, an alternative possibility is that PigeoNET is more inclined to assign visual characteristics that are present in their works to Boëtius Bolswert because his works appear more frequently in the data set.

In a similar vein, the misclassifications that occur between Fokke Simon (10) and Jan Caspar Philips (29), and between Jan Luyken (23) and Caspar Luyken (22), are notable. Fokke Simon was a student of Jan Caspar Philips, and Jan and Caspar



[FIG3] The (a) image and (b) heat map of *Studie van een op de rug geziene man* (1629–1630) by Rembrandt Harmenszoon van Rijn. The lower (red) values in the heat map correspond to greater importance in correctly identifying Rembrandt.

Luyken were father and son. Both pairs of artists have worked together on several artworks in the Rijksmuseum Challenge data set, despite the label in the data set indicating that these artworks belong to only one of these artists. We became aware of these potential dual-authorship cases after having performed our main experiment. Dual-authorship cases will be examined in more detail through visualizations in the section “Deciding Between Two Artists.”

VISUALIZATION AND ASSESSMENT

Visualizations of the importance of each region in an artwork can be generated using the regions of importance detection method described in the section “Visualization of Artist-Characteristic Regions,” where the occlusions are performed with a gray block of 8×8 pixels to indicate approximate regions that are characteristic of the artist. The regions of importance can be visualized using

heat map color coding, as shown in Figure 3(b). The value of a region in the heat map corresponds to the certainty score of PigeoNET for the artwork with that region occluded. In other words, a region with a lower value is of greater importance in correctly attributing the artwork, with (dark) red regions being highly characteristic of the artist, and (dark) blue regions being the least characteristic.

THE GOAL OF AN ARTIST ATTRIBUTION TASK IS TO ATTRIBUTE AN UNSEEN ARTWORK TO THE ARTIST WHO CREATED IT.

When comparing the artwork and heat map in Figure 3 of the drawing by Rembrandt, it is very noticeable that PigeoNET assigns much weight to seemingly empty areas. The texture of the material on which an artwork is created can be indicative of the artist who created the artwork

[26]. When taking a closer look at Figure 4, with enhanced contrast, it becomes apparent that the areas are not empty and that a distinctive visual texture is present. The visual pattern is sufficiently distinctive and artist-specific for PigeoNET to assign it a larger weight. The pattern is an example of a visual characteristic that is indirectly related to the artist. It illustrates the importance of the transparency of automatic attribution to allow human experts to interpret and evaluate the visual characteristic.

DECIDING BETWEEN TWO ARTISTS

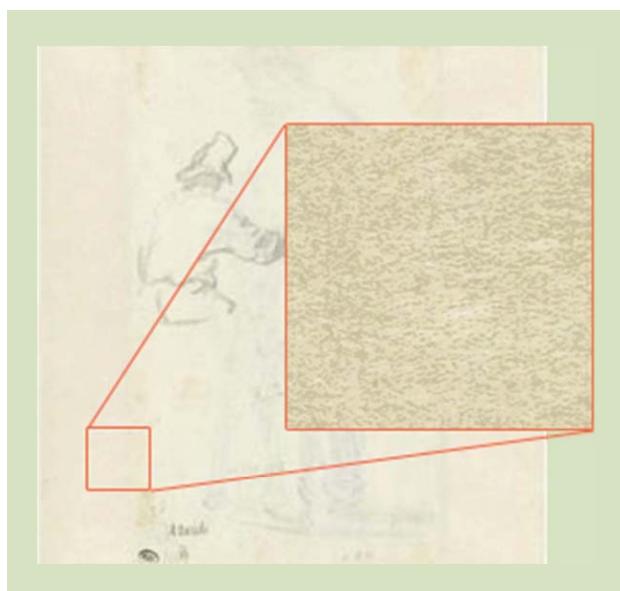
In the “Visualization and Assessment” section, we used PigeoNET to attribute an artwork to a single artist. Yet, as illustrated by the work of Peter Paul Rubens and Jan Brueghel in Figure 1, in many cases two (or more) artists have worked on the same artwork (see also [27]).

As evident from our results, PigeoNET had difficulty in correctly attributing artworks of closely collaborating artists. An intriguing explanation for PigeoNET’s failure to assign the “correct one” of two potential artists to artworks is that the artworks are created by both artists. In that case, it would not be a failure at all and indicates that PigeoNET discovered that the two artists are similar, and it recognizes the characteristic features of both artists, even if the work is attributed to only one artist. In the remainder of this section, we demonstrate the possibility of using PigeoNET to perform a fine-grained analysis of an artwork, attributing individual image regions to an artist.

DISCOVERING DUAL AUTHORSHIP

PigeoNET had difficulty in distinguishing between the works of Jan and Caspar Luyken, a father and son who worked together and created many prints. Throughout their careers, Jan Luyken chose to depict pious and biblical subjects, whereas Caspar Luyken mostly depicted worldly scenes [28]. As an example, we consider the artwork shown in Figure 5, *Over dracht der Nederlande, aan de Infante Isabella*. The work depicts the transfer of the Spanish Netherlands by Filips II to Isabella Clara Eugenia. Although arguably it is a very worldly scene, it is nevertheless attributed to Jan Luyken. Could it be possible that the artwork is incorrectly attributed to Jan Luyken? Obviously, this is a question that has to be answered by experts of their works.

Our findings may support them in their assessment. Although PigeoNET correctly attributed the artwork to Jan Luyken, the reported certainty score for Caspar Luyken is very high. Apparently, PigeoNET responds to visual features that are characteristic of Caspar Luyken. Using PigeoNET’s visualization, we are able to determine for each region how characteristic it is for each of the two artists. We created a visualization based on the certainty scores for



[FIG4] A contrast-enhanced detailed view of a highly textured region of the artwork shown in Figure 3(a).



[FIG5] An image of the *Overdracht der Nederlande, aan de Infante Isabella* (1697–1699) by Jan Luyken.

Jan Luyken and Caspar Luyken. Figure 6 shows the visualization using color coding on a yellow to blue scale. The yellow regions are characteristic for Jan Luyken, whereas the blue regions are characteristic for Caspar Luyken, the green regions are indeterminate and show characteristics of either artists in equal amounts. This example demonstrates the potential use of PigeoNET to support the study of dual authorship artworks.

DISCUSSION

Previous work on automatic artist attribution has shown that prior knowledge can be leveraged to engineer features for automatic artist attribution. In this article, we presented a novel approach that does not rely on prior knowledge and is capable of discovering characteristic features automatically enabling a successful artist attribution. Additionally, we demonstrated that PigeoNET visualizations reveal artwork regions most characteristic of the artist and that PigeoNET can aid in answering outstanding questions regarding dual-authorship.

We will now discuss the considerations regarding the data set used and address how the selection of subsets may affect the nature of visual characteristics discovered. Although the Rijksmuseum Challenge data set is the largest available data set containing digital reproductions of artworks acquired under controlled conditions [24], it does suffer from two main limitations. First, given the wide variety of artwork types, it is unclear how the “controlled conditions” were defined for different artworks. Any variation in the reproduction setting (e.g., illumination, perspective, camera type) may be picked up by PigeoNET. Presumably, our P (prints only) data sets suffer less from this problem. Still, even in these data sets subtle differences in

digitization may leave visual marks that are picked up by PigeoNET. An ideal data set for attribution would be one in which no such visual marks are present. Unfortunately, such data sets do not exist and are hard (if not impossible) to create on this scale. Therefore, transparency of the acquired features by PigeoNET and proper visualizations are essential to aid art experts in their assessment of the feasibility of classifications.

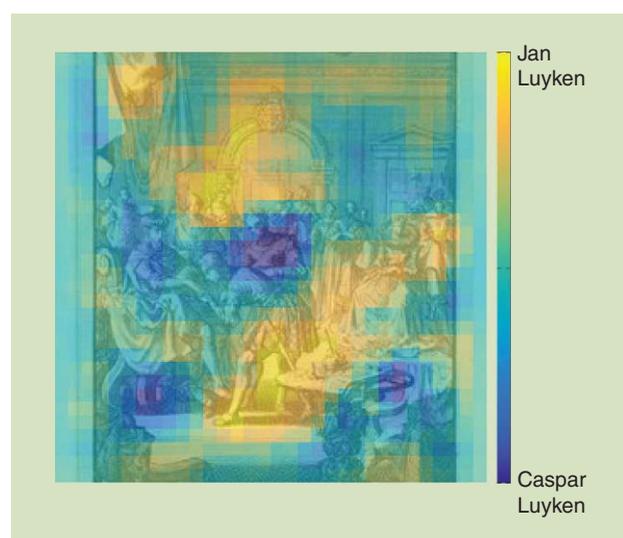
The second limitation concerns the labeling of artworks. After having performed our main experiments, we discovered that for some artworks, the Rijksmuseum catalog lists multiple contributions, whereas the Rijksmuseum Challenge data set only lists a single artist

[24]. The contributions listed in the Rijksmuseum catalog vary greatly (from inspiration to dual authorship) and do not always influence the actual attribution, but do create uncertainty about the attribution of artworks in the Rijksmuseum challenge data set. Although this significantly limits the possibility of learning stylistic features from such artworks, it does not prohibit PigeoNET from learning visual characteristics that are associated with the primary artist as such characteristics remain present in the artwork. Still, the validity and consistency of attributions is of major concern to safeguard the validity of methods such as PigeoNET. Also in the creation of such databases, the involvement of human art experts is required.

The results obtained in this work on the automatic artist attribution task show that PigeoNET is capable of accurately attributing unseen works to the correct artist. The increase of performance for the sets with a higher number of examples shows that including more examples per artist leads to a better performance. Moreover, the complete Rijksmuseum Challenge data set is a highly diverse data set with many different types of art. For some cases (e.g., the porcelain of the Meissener Porzellan Manufaktur) this results in a class that is visually very distinctive from the rest of the data set, which could make it easier to identify the correct artist. However, when comparing the performances obtained on the homogeneous P1 subsets (prints only) with those on the more heterogeneous A subsets (all artwork types), the difference in performance is quite small. This demonstrates that PigeoNET is capable of learning a rich representation of multiple artwork types without a major impact on its predictive power. Part of the types of features discovered in the A subsets are likely to distinguish between art types (e.g., a porcelain object versus a painting), rather than between author styles. In the P subsets, features will be more tuned to stylistic differences, because these subsets are confined to a single type of artwork.

Our findings indicate that the number of artists and the number of examples per artist have a very strong influence on the performance, which suggests that a further improvement of the performance is possible by expanding the data set. In future research we will determine to what extent this is the case.

**THE RESULTS OBTAINED
IN THIS WORK ON THE
AUTOMATIC ARTIST ATTRIBUTION
TASK SHOW THAT PIGEONET
IS CAPABLE OF ACCURATELY
ATTRIBUTING UNSEEN WORKS
TO THE CORRECT ARTIST.**



[FIG6] Visualization of how characteristic each image region is for the artists Jan and Caspar Luyken. The yellow regions are characteristic of Jan Luyken, whereas the blue regions are characteristic of Caspar Luyken.

CONCLUSIONS

In this article we have evaluated a feature learning system to assess to what extent it is possible to discover an artist's visually characteristic features. The results on the automatic attribution task demonstrate that the system is capable, up to a high degree of accuracy, of using visual characteristics to assign unseen artworks to the correct artist. Moreover, we demonstrated the possibility of using the visual characteristics to reveal the artist of a specific region within an artwork, which in the case of multiple artists could lead to new discoveries about the origin and creation of important works of cultural heritage. Therefore, we conclude that PigeoNET represents a fruitful approach for future computer-supported examination of artworks.

ACKNOWLEDGMENTS

We thank Teio Meedendorp and Louis van Tilborgh for their useful contributions to earlier versions of this article. Furthermore, we acknowledge the anonymous reviewers for their helpful comments. The research reported in this article is performed as part of the Reassessing Vincent van Gogh project supported by The Netherlands Organisation for Scientific Research (NWO; grant 323.54.004) in the context of the Science4Arts research program.

AUTHORS

Nanne van Noord (n.j.e.vannoord@tilburguniversity.edu) is a Ph.D. candidate working on the Reassessing Vincent van Gogh project at the Tilburg Center for Cognition and Communication, Tilburg University, The Netherlands. His research interests include machine learning, representation learning, and computer vision.

Ella Hendriks (hendriks@vangoghmuseum.nl) is a senior paintings conservator at the Van Gogh Museum, as well as an associate professor at the University of Amsterdam Training Programme for Conservation and Restoration of Cultural Heritage.

Eric Postma (e.o.postma@tilburguniversity.edu) is a full professor in artificial intelligence at the Tilburg Center for Cognition and Communication, Tilburg University, The Netherlands. His work focuses on machine learning and image processing with an emphasis on cultural heritage applications.

REFERENCES

- [1] J. Rush, "Acquiring a concept of painting style," Ph.D. dissertation, Dept. of Educational Psychology, Univ. Arizona, Tucson, AZ, 1974.
- [2] I. Berezchnoy, E. Postma, and J. van den Herik, "Computer analysis of Van Gogh's complementary colours," *Pattern Recog. Lett.*, vol. 28, no. 6, pp. 703–709, Apr. 2007.
- [3] J. Li, L. Yao, E. Hendriks, and J. Z. Wang, "Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1159–76, June 2012.
- [4] M. M. van Dantzig, *Pictology. An Analytical Method for Attribution and Evaluation of Pictures*, the van Dantzig Foundation, editor. Leiden, The Netherlands: Brill Leiden, 1973.

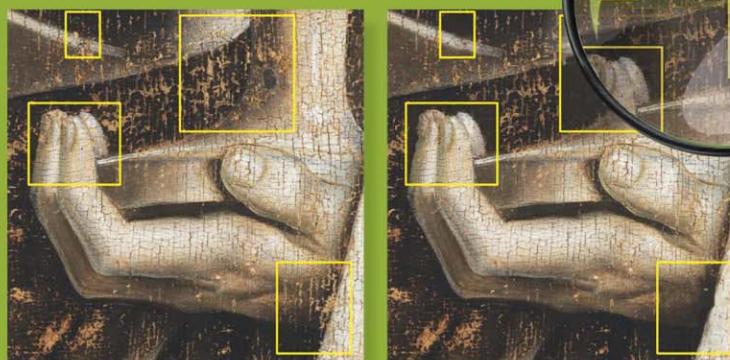
- [5] T. Lombardi, "The classification of style in fine-art painting," Ph.D. dissertation, School of Computer Science and Information Systems, Pace Univ., NY, 2005.
- [6] C. R. Johnson, Jr., E. Hendriks, I. J. Berezchnoy, E. Brevido, S. M. Hughes, I. Daubechies, J. Li, E. Postma, and J. Z. Wang, "Image processing for artist identification," *IEEE Signal Processing Mag.* vol. 25, no. 4, pp. 37–48, 2008.
- [7] J. M. Hughes, D. J. Graham, and D. N. Rockmore, "Quantification of artistic style through sparse coding analysis in the drawings of Pieter Bruegel the Elder," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 107, no. 4, pp. 1279–83, Jan. 2010.
- [8] R. Taylor, R. Guzman, and T. Martin, "Authenticating Pollock paintings using fractal geometry," *Pattern Recog. Lett.*, vol. 28, no. 6, pp. 695–702, 2007.
- [9] L. J. P. van der Maaten and E. O. Postma, "Texton-based analysis of paintings," in *SPIE Optical Imaging and Applications*, Aug. 2010, vol. 7798. [Online]. Available: <http://proceedings.spiedigitallibrary.org/volume.aspx?conferenceid=1936&volumeid=678>
- [10] P. Abry, H. Wendt, and S. Jaffard, "When Van Gogh meets Mandelbrot: Multifractal classification of painting's texture," *Signal Process.*, vol. 93, no. 3, pp. 554–572, Mar. 2013.
- [11] M. Eraut, "Non-formal learning and tacit knowledge in professional work," *Br. J. Educ. Psychol.*, vol. 70, no. 1, pp. 113–136, Mar. 2000.
- [12] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013. doi:10.1109/TPAMI.2013.50
- [13] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Adv. Neural Inform. Process. Syst.*, vol. 25, 2012, pp. 1097–1105.
- [14] S. Watanabe, J. Sakamoto, and M. Wakita, "Pigeons' discrimination of paintings by Monet and Picasso," *J. Exp. Anal. Behav.*, vol. 63, no. 2, pp. 165–174, 1995.
- [15] W. Wu, A. M. Moreno, J. M. Tangen, and J. Reinhard, "Honeybees can discriminate between Monet and Picasso paintings," *J. Comparative Physiol. A*, vol. 199, no. 1, pp. 45–55, Oct. 2012.
- [16] M. Zeiler and R. Fergus, "Visualizing and understanding convolutional neural networks," 2013, arXiv:1311.2901v2.
- [17] Y. Bengio and Y. Lecun, "Convolutional Networks for Images, Speech, and Time-Series," in *The Handbook of Brain Theory and Neural Networks*, M. A. Arbib, Ed. Cambridge, MA: MIT Press, 1995, pp. 255–258.
- [18] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computat.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.
- [19] A. K. Jain, N. K. Ratha, and S. Lakshmanan, "Object detection using Gabor filters," *Pattern Recog.*, vol. 30, no. 2, pp. 295–309, Feb. 1997.
- [20] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proc. 2014 IEEE Conf. Computer Vision Pattern Recognition*, June 2014, pp. 1717–1724.
- [21] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: visualising image classification models and saliency maps," pp. 1–8, Dec. 2013, arXiv:1312.6034.
- [22] D. Erhan and Y. Bengio, "Visualizing higher-layer features of a deep network," Tech. Rep. 1341, Univ. Montreal, 2009.
- [23] A. T. Woollett and A. van Suchtelen, *Rubens and Brueghel: A Working Friendship*. Los Angeles, CA: J. Paul Getty Museum, 2006.
- [24] T. Mensink and J. van Gemert, "The Rijksmuseum challenge: museum-centered visual recognition," in *Proc. Int. Conf. Multimedia Retrieval*, 2014, pp. 2–5.
- [25] Y. Jia. (2013). Caffe: An open source convolutional architecture for fast feature embedding. [Online]. Available: <http://caffe.berkeleyvision.org>
- [26] D. H. Johnson, C. R. Johnson, Jr., A. G. Klein, W. A. Sethares, H. Lee, and E. Hendriks, "A thread counting algorithm for art forensics," in *Proc. IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop*, 2009, pp. 679–684.
- [27] H. Bevers, L. Hendrix, W. W. Robinson, and P. Schatborn, *Drawings by Rembrandt and His Pupils: Telling the Difference*. Los Angeles, CA: J. Paul Getty Museum, 2009.
- [28] [Online]. Available: http://www.geheugenvannederland.nl/?en/collecties/jan_en_casper_luyken/vader_zoon



[Haizhao Yang, Jianfeng Lu, William P. Brown, Ingrid Daubechies, and Lexing Ying]

Quantitative Canvas Weave Analysis Using 2-D Synchrosqueezed Transforms

© GHENT, KATHEDRALE KERKFABRIEK; PHOTO COURTESY OF KIK-IPRA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

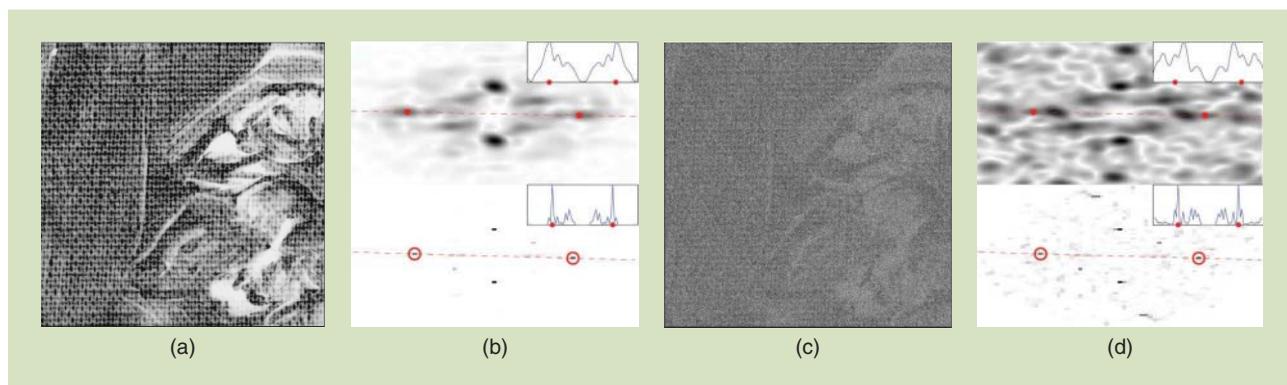
[Application of time-frequency analysis to art investigation]

Quantitative canvas weave analysis has many applications in art investigations of paintings, including dating, forensics, and canvas rollmate identification [1]–[3]. Traditionally, canvas analysis is based on X-radiographs. Prior to serving as a painting canvas, a piece of fabric is coated with a priming agent; smoothing its surface makes this layer thicker between and thinner right on top of weave threads. These

variations affect the X-ray absorption, making the weave pattern stand out in X-ray images of the finished painting. To characterize this pattern, it is customary to visually inspect small areas within the X-radiograph and count the number of horizontal and vertical weave threads; averages of these then estimate the overall canvas weave density. The tedium of this process typically limits its practice to just a few sample regions of the canvas. In addition, it does not capture more subtle information beyond weave density, such as thread angles or variations in the weave pattern. Signal processing techniques applied to art investigation are now increasingly used to develop computer-assisted canvas weave analysis tools.

Digital Object Identifier 10.1109/MSP.2015.2406882

Date of publication: 15 June 2015



[FIG1] (a) A sample swatch of an X-ray image in which canvas is clearly visible (in most places) despite the paint layers on top of the canvas. (b) The spectrum of the wFT (top) and 2DST (bottom) at one location. Local maxima (circled in red) indicate the wave vector estimates; the insets show the intensity profile on a cross section (dashed line) through two maxima. (c) The same swatch as in (a) with noise added (such that the noise level is visually comparable to the real data example in Figure 9) to test for robustness. (d) The wFT and 2DST spectra again at the same location, illustrating the more robust nature of the 2DST estimate (due to its taking into account phase information of the wFT in a neighborhood of the peaks of the absolute value of the wFT as well as the peak values). For comparison, the positions of the red circles are the same as in (b). The peaks are displaced in wFT due to noise, while the result of 2DST is not affected.

In their pioneering work [4], Johnson et al. developed an algorithm for canvas thread-counting based on windowed Fourier transforms (wFTs); further developments in [5] and [6] extract more information, such as thread angles and weave patterns. Successful applications to paintings of art historical interest include works by Vincent van Gogh [7], [8], Diego Velázquez [9], and Johannes Vermeer [10], among others [11]–[15].

A more robust and automated analysis technique was later developed by Erdmann et al. [16] based on autocorrelation and pattern recognition algorithms, requiring less human intervention (e.g., choosing proper frequency range and window size of wFTs). Unlike the Fourier-space-based approach of [4], [16] uses only the real-space representation of the canvas. Likewise, [17] also uses real-space-based features for canvas texture characterization.

In this article, we consider a new automated analysis technique for quantitative canvas analysis based on the two-dimensional (2-D) synchrosqueezed transforms (2DSTs) recently developed in [18]–[20]. This Fourier-space-based method applies the nonlinear synchrosqueezing procedure to a phase-space representation of the image obtained by wave packet or curvelet transforms. Synchrosqueezing has shown to be a useful tool in independent work by some in [18]–[22], in the general area of materials science, medical signal analysis, and seismic imaging. Using as a prior assumption that the signal of interest consists of a sparse superposition of close to but not quite periodic template functions, this mathematical tool provides sharp and robust estimates for the locally varying instantaneous frequencies of the signal components by exploiting the phase information of wFTs (i.e., not only the absolute value as in previous methods). This seemed to make it a natural candidate for canvas analysis; as illustrated by the results we obtained, reported here, this intuition proved to be correct. The method, as shown next, is very robust and offers fine-scale weave density and thread angle information for the canvas. We compare our results with those in [4]–[6] and [16].

MODEL OF THE CANVAS WEAVE PATTERN IN X-RADIOGRAPHY

We denote by f the intensity of an X-radiograph of a painting; see Figure 1(a) for a (zoomed-in) example. Because X-rays penetrate deeply, the image consists of several components: the paint layer itself, primer, canvas (if the painting is on canvas or on wood panel overlaid with canvas), possibly a wood panel (if the painting is on wood), and sometimes extra slats (stretchers for a painting on canvas, or a cradle for a painting on wood, thinned and cradled according to earlier conservation practice.) This X-ray image may be affected by noise or artifacts of the acquisition process. We model the intensity function f as an additive superposition of the canvas contribution, denoted by $c(x)$, and a remainder, denoted by $p(x)$, that incorporates all the other components. Our approach to quantitative canvas analysis relies on a simple model for the X-ray image of the weave pattern in the “ideal” situation. Since it is produced by the interleaving of horizontal and vertical threads in a periodic fashion, a natural general model is

$$f(x) = c(x) + p(x) := a(x)S(2\pi N\phi(x)) + p(x). \quad (1)$$

In this expression, S is a periodic function on the square $[0, 2\pi)^2$, the details of which reflect the basic weave pattern of the canvas, e.g., whether it is a plain weave or perhaps a twill weave. This is a generalization of more specific assumptions used in the literature—for instance, in [4] a plain weave canvas is modeled by taking for S a sum of sinusoidal functions in the x and y directions; in [6], more general weave patterns (in particular twill) are considered. The parameter N in (1) gives the averaged overall weave density of the canvas (in both directions). The function ϕ , which maps the image domain to \mathbb{R}^2 , is a smooth deformation representing the local warping of the canvas; it contains information on local thread density, local thread angles, etc. The slowly varying function $a(x)$ accounts for variations of the amplitude of the X-ray image of the canvas, e.g., due to variation in illumination conditions.

In some cases, the X-ray image fails to show canvas information in portions of the painting (e.g., when the paint layer dominates); the model (1) is then not uniformly valid. Because our analysis uses spatially localized information (analyzing the image patch by patch), this affects our results only locally: in those (small) portions of the image we have no good estimates for the canvas parameters. For simplicity, this exposition assumes that (1) is valid for the whole image.

We rewrite c by representing the weave pattern function S , periodic on $[0, 2\pi]^2$, in terms of its Fourier series,

$$c(x) = \sum_{n \in \mathbb{Z}^2} a(x) \hat{S}(n) e^{2\pi i n \cdot \phi(x)}. \quad (2)$$

This is a superposition of smoothly warped plane-waves with local wave vectors $N \nabla(n \cdot \phi(x))$. The idea of our analysis is to extract the function ϕ by exploiting that the Fourier coefficients $\{\hat{S}(n)\}$ are dominated by a few leading terms.

FOURIER-SPACE-BASED CANVAS ANALYSIS

WINDOWED FOURIER TRANSFORM

Because a and ϕ vary slowly with x , we can use Taylor expansions to approximate the function for x near x_0 as

$$c(x) \approx \sum_{n \in \mathbb{Z}^2} a(x_0) \hat{S}(n) e^{2\pi i n \cdot \phi(x_0)} e^{2\pi i N(x-x_0) \cdot \nabla_x(n \cdot \phi)(x_0)}. \quad (3)$$

The right-hand side of (3) is a superposition of complex exponentials with frequencies $w = (w_1, w_2)$, with

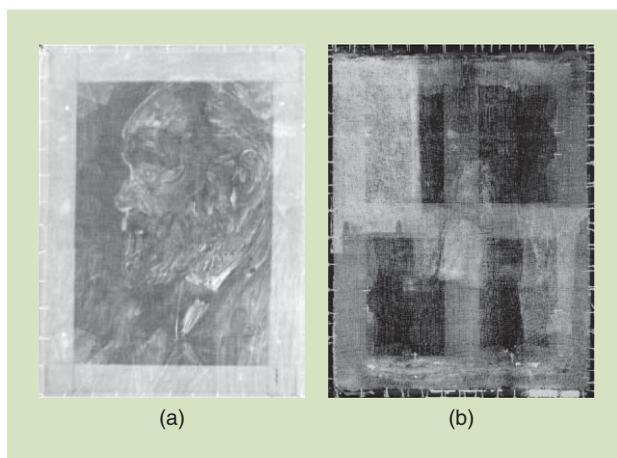
$$w_l = \sum_{l'=1}^2 n_{l'} (\partial_{l'} \phi_{l'}) (x_0);$$

these would stand out in a Fourier transform as peaks in the 2-D Fourier spectrum. Since the approximation is accurate only near x_0 , we also use a wFT with envelope given by, e.g., a Gaussian centered at x_0 with width σ . We have then

$$\begin{aligned} W(x_0, k) &:= \frac{1}{2\pi\sigma^2} \iint e^{-2\pi i k(x-x_0)} e^{-(x-x_0)^2/2\sigma^2} c(x) dx \\ &\approx \sum_{n \in \mathbb{Z}^2} a(x_0) \hat{S}(n) e^{2\pi i n \cdot \phi(x_0)} e^{-2\pi^2 \sigma^2 [k - N \nabla_x(n \cdot \phi)(x_0)]^2}. \end{aligned} \quad (4)$$

Instead of being sharply peaked, the spectrum of the wFT is thus “spread out” around the $N \nabla_x(n \cdot \phi)(x_0)$ —a manifestation of the well-known uncertainty principle in signal processing, with a tradeoff with respect to the parameter σ : a larger σ reduces the “spreading” at the price of a larger error in the approximation (3), since the Gaussian is then correspondingly wider in the real space.

The method of [4] and [6] uses the local maxima of the amplitude of the wFT to estimate the location of $\{N \nabla(n \cdot \phi)(x_0)\}$ for a selection of positions x_0 of the X-ray image (local swatches are used instead of the Gaussian envelope, but the spirit is the same). For ideal signals, (4) shows that the maxima of the amplitude $|W(x_0, \cdot)|$ identify the dominating wave vectors in Fourier space, which are then used to extract information, including weave density and thread angles. Thread density is estimated by the length of the wave vectors; the weave orientation is determined by the angles. This back-of-the-envelope calculation is fairly precise when



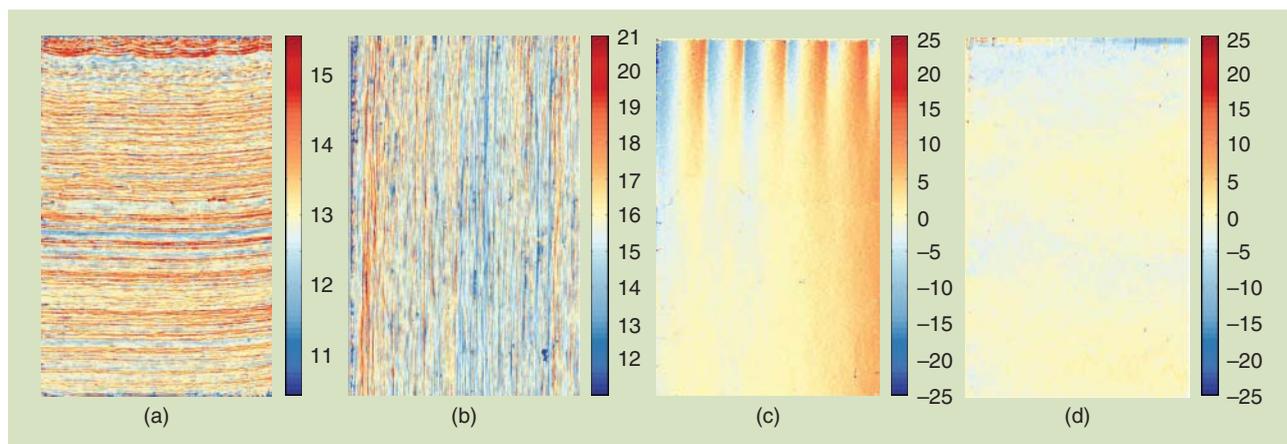
[FIG2] (a) An X-ray image of van Gogh’s painting *Portrait of an Old Man with Beard*, 1885, van Gogh Museum, Amsterdam (F205). (b) An X-ray image of Vermeer’s painting *Woman in Blue Reading a Letter*, 1663–1664, Rijksmuseum, Amsterdam, The Netherlands (L17). [X-ray images provided by Prof. C. Richard Johnson through the Rijksbureau voor Kunsthistorische Documentatie (RKD)/The Netherlands Institute for Art History data set [28].]

N is much larger than 1, resulting in a small $O(N^{-1})$ error in the Taylor expansions and stationary phase approximations. In terms of the canvas, $N \gg 1$ means that the inverse of the average thread density must be much smaller than the length scale of the variation of the canvas texture, which is typically on the scale of the size of the painting. This is essentially a high-frequency assumption, ensuring that stationary phase approximations can be applied in the time-frequency analysis. Details can be found in standard references of time-frequency analysis, e.g., [23].

In more complicated scenarios, in particular, when the X-ray signal corresponding to the canvas is heavily “contaminated” by the other parts of the painting, it is desirable to have more robust and refined analysis tools at hand than locating local maxima of the Fourier spectrum. The synchrosqueezed transforms are nonlinear time-frequency analysis tools developed for this purpose, in different [one-dimensional (1-D) and 2-D] applications that suggests they could be suitable for canvas analysis in challenging situations. A comparison of the two methods is shown in Figure 1 and will be explained next. For the sake of completeness, we note that in our implementation, we use curvelets (more or less corresponding to a nonisotropic Gaussian window, with axes-lengths adapted to the frequencies of the oscillating component) rather than wFTs with isotropic Gaussian windows, to which we have restricted ourselves in this exposition. The synchrosqueezing operation has similar effects in both cases; the curvelet implementation, while more complicated to explain in a nutshell, has the advantage of being governed by only two parameters, which set the spatial redundancy and the angular resolution. Setting these is well understood (see [24]); in addition the result is stable under small perturbations in these parameters.

SYNCHROSQUEEZED TRANSFORMS

The synchrosqueezed transforms, or more generally time-frequency reassignment techniques (see, e.g., the recent review [25]),



[FIG3] The canvas analysis results of van Gogh’s F205 using the synchrosqueezed transform: (a) and (b) thread count map of the horizontal and vertical threads and (c) and (d) the estimated horizontal and vertical thread angles. Compare with [6, Fig. 6].

were introduced to deal with the “loss of resolution” due to the uncertainty principle. Originally introduced in [26] for auditory signals, using a nonlinear squeezing of the time-frequency representation to gain sharpness of the time-frequency representation, the 1-D synchrosqueezed wavelet transform was revisited and analyzed in [27]. For the application to canvas analysis, we rely on 2-D extensions of the synchrosqueezing transforms based on wave packet and curvelet transforms [18], [19]. This 2DST has been applied to atomic-resolution crystal image analysis in [20]; the present algorithm for canvas analysis is adapted from [20], where the 2DST proved to be an excellent tool to capture and quantify deviations from a perfect lattice structure, very similar to the aims of canvas analysis. Rigorous robustness analysis of the synchrosqueezed transforms in [24] supports their application to canvas analysis where data is usually noisy and contains contaminants.

The crucial observation is that the phase of the complex function $W(x, k)$, obtained from the wFT (4) contains information on

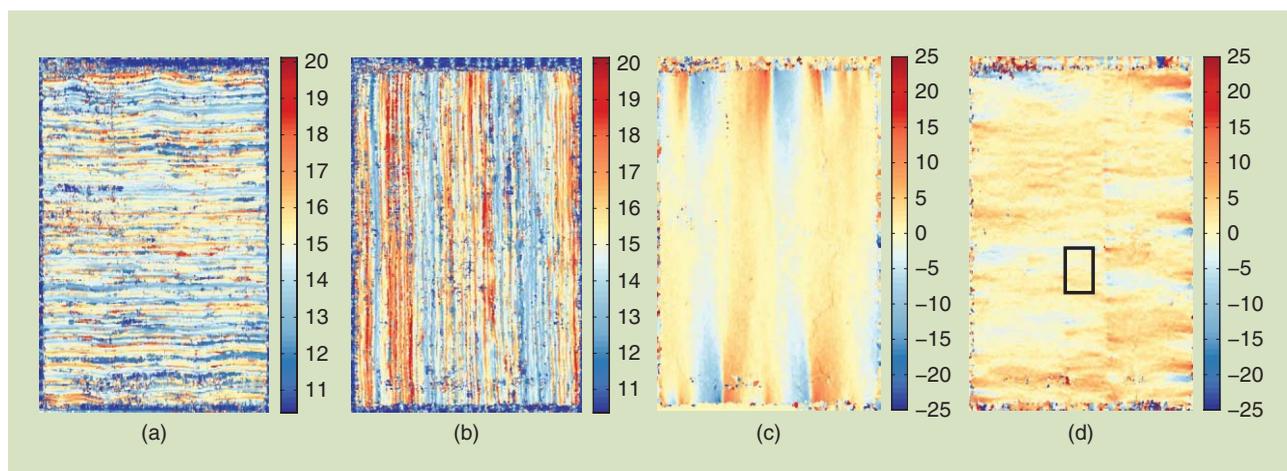
the local frequency (i.e., the instantaneous frequency) of the signal. Indeed, for (x, k) such that k is close to $N\nabla_x(n \cdot \phi)$, we have

$$w_f(x, k) := \frac{1}{2\pi} \Im(\nabla_x \ln W(x, k)) = N\nabla_x(n \cdot \phi)(x) + o(N), \quad (5)$$

where $\Im(z)$ stands for the imaginary part of the complex number z . Motivated by this heuristic, the synchrosqueezed wFT “squeezes” the time-frequency spectrum by reassigning the amplitude at (x, k) to $(x, w_f(x, k))$ as

$$T(x, \xi) := \iint |W(x, k)|^2 \delta(\xi - w_f(x, k)) d^2k. \quad (6)$$

This significantly enhances the sharpness of the time-frequency representation, leading to an estimate of the local frequency of the signal, that is more accurate as well as more robust, as we illustrate below. This gives a sharpened energy distribution on phase space:



[FIG4] Canvas analysis results of Vermeer’s L17 using the synchrosqueezed transform: (a) and (b) are a thread count map of the horizontal and vertical threads while (c) and (d) show the estimated horizontal and vertical thread angles. Average thread density is 14.407 threads/cm (horizontal) and 14.817 threads/cm (vertical). The boxed region of the (d) vertical thread angle map is shown, enlarged, in Figure 5; it is part of a striking anomaly in the vertical angle pattern in this canvas, lining up along one vertical traversing the whole canvas.

$$T(x, \xi) \approx \sum_{n \in \mathbb{Z}^2} |a(x)|^2 |\hat{S}(n)|^2 \delta(\xi - N\nabla(n \cdot \phi(x))) \quad (7)$$

in the sense of distributions. See [18]–[20] for more details, as well as an analysis of the method. The peaks of the synchrosqueezed spectrum T then provide estimates of the $N\nabla(n \cdot \phi(x))$, determining local measurement of both the thread count and the angle. Figure 1 illustrates the resulting spectrum of the 2DST, compared with the wFT for a sample X-ray image from a canvas. The reassignment carried out in (6), taking into account the local oscillation of the phase of a highly redundant wFT rather than the maximum energy of the wFT to reduce the influence of noise, results in a much more concentrated spatial frequency portrait. As illustrated by the behavior of the estimates when extra noise is added, this leads to increased robustness for the estimates of the dominating wave vectors, which determine the thread count and angle. The performance and the robustness of the 2DST are supported by rigorous mathematical analysis in [24].

APPLICATIONS TO ART INVESTIGATIONS

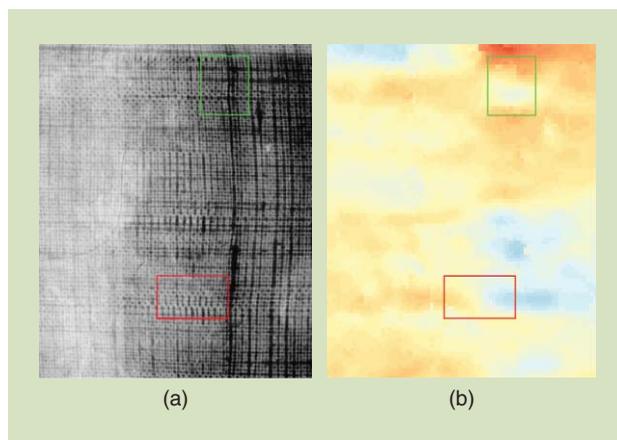
Let us now present some results of quantitative canvas analysis using 2DST. The algorithm is implemented in MATLAB. The codes are open source and available as SynLab at <https://github.com/HaizhaoYang/SynLab>.

The first example [Figure 2(a)] is the painting F205 by van Gogh, the X-ray image of which is publicly available as part of the RKD data set [28] provided by The Netherlands Institute for Art History; this was one of the first examples analyzed using the method based on the wFT; see [4, Fig. 4] and also [6, Fig. 6]. In Figure 3, the thread count and thread angle estimates are shown for horizontal and vertical threads. Comparing with the previous results in [4] and [6], we observe that the general characteristics of the canvas agree quite well. For example, [6] reports average thread counts of 13.3 threads/cm (horizontal) and 16.0 threads/cm (vertical), while our method obtains 13.24 threads/cm (horizontal) and 15.92 threads/cm (vertical). Compared to the earlier results, the current analysis gives a more detailed spatial variation of the thread counts. In particular, it captures the oscillation of the thread count on a much finer scale. We don't know whether such fine details will have applications beyond the canvas characterization already achieved by less detailed methods, but it is interesting that they can be captured by an automatic method. Note that visual inspection confirms the presence of these fine details.

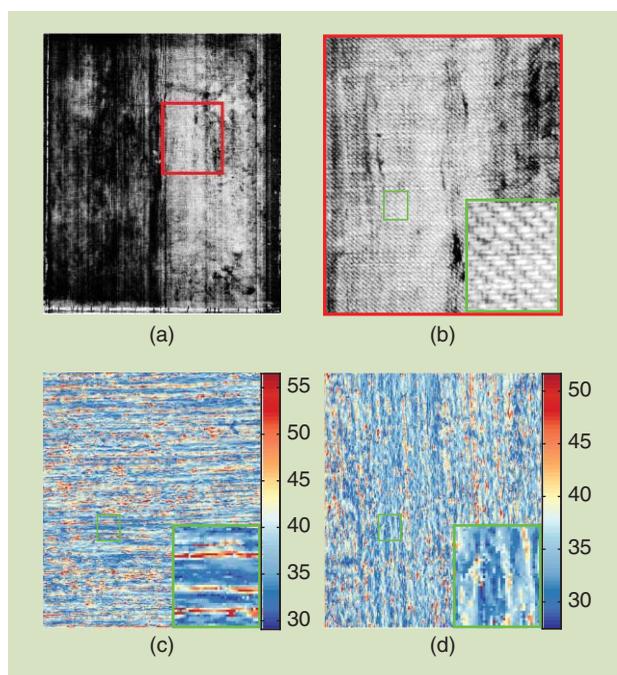
We next consider a painting of Vermeer, *Woman in Blue Reading a Letter* (L17), the X-ray image of which is also available as part of the RKD data set [28]. The canvas analysis for Vermeer's paintings is considerably more challenging than that of van Gogh's [10]. This can be understood by direct comparison of the X-ray images in Figure 2(a) and (b). The stretchers and nails significantly perturb the X-ray image for the Vermeer. The results are shown in Figures 4 and 5. Although the thread count and angle estimate are affected by artifacts in the X-ray image, they still provide a detailed characterization of the canvas weave. This is justified by the result in Figure 5, which shows a zoom-in for the X-ray image and the vertical thread angle map. It is observed that the algorithm captures (and quantifies)

detailed deviations in the vertical thread angle recognizable by visual inspection. Despite the challenges, the 2DST-based canvas analysis performs quite well on the Vermeer example.

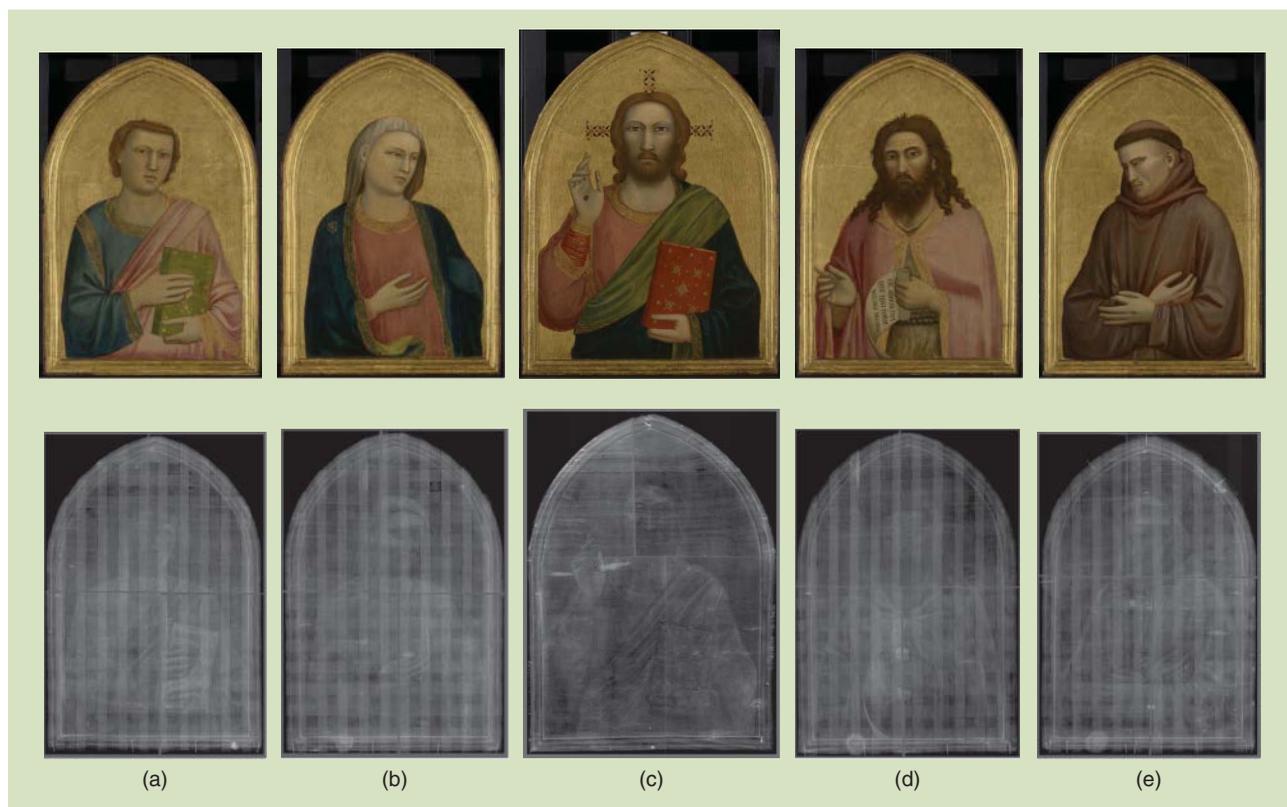
To test the algorithm on a different type of canvas weave, we applied it to the X-ray image of Albert P. Ryder's *The Pasture*, a painting on twill canvas. Figure 6 shows the result for a portion



[FIG5] Details of (a) the X-ray image and (b) the corresponding vertical thread angle map for Vermeer's L17, highlighting two examples (boxed regions) of noticeable fine scale variation of the vertical thread angle, readily recognizable also by visual inspection of the corresponding zones in the X-ray image.



[FIG6] (a) An X-ray image of Albert P. Ryder's *The Pasture*, 1880–1885, North Carolina Museum of Art, Raleigh. (b) An enlargement of the red-boxed region with clearly recognizable twill canvas weave. (c) and (d) Horizontal and vertical thread count maps corresponding to the zoomed-in region shown in (b). Note the much higher thread counts than for plain weave canvas, typical for the finer threads used in twill weave. The bottom-right insets of (b)–(d) show the further zoom-in of the green-boxed region for visual inspection. The horizontal thread count matches the changes observed in the X-ray image quite well.



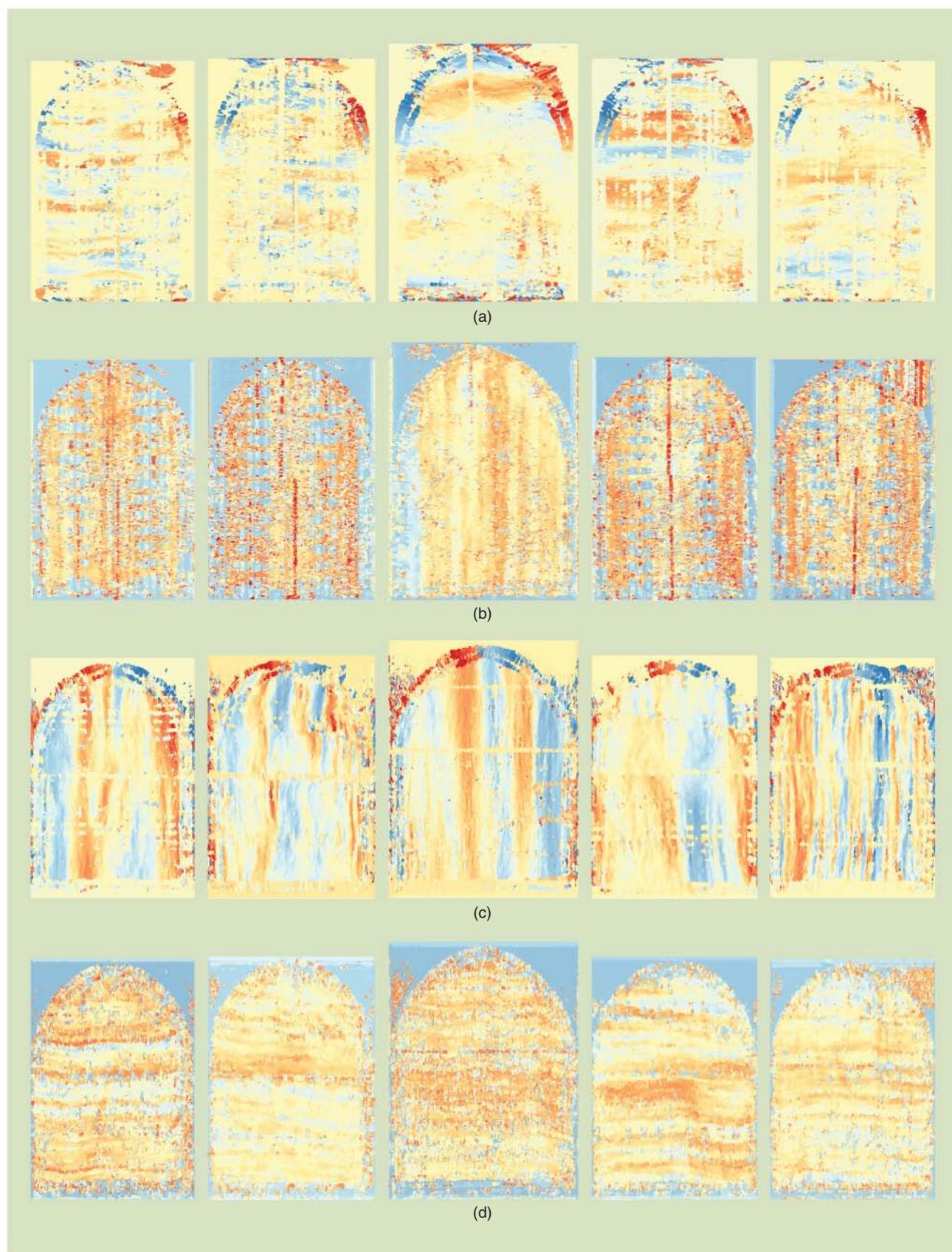
[FIG7] Giotto di Bondone and assistants, *The Peruzzi Altarpiece*, ca. 1310–1315, North Carolina Museum of Art, Raleigh. The panels are (a) John the Evangelist, (b) the Virgin Mary, (c) Christ in Majesty, (d) John the Baptist, and (e) Francis of Assisi. The resolution of the X-ray image used in the analysis is 300 dots/in. The vertical and (less obvious) horizontal stripes on the X-ray images in all panels except the central panel of Christ are caused by cradling. Each X-ray image is a mosaic of four X-ray films, leading to visible boundaries of the different pieces (thin horizontal and vertical lines) on the X-ray image.

of the canvas. The twill canvas pattern is clear on the zoomed-in X-ray image. The method is still able to capture fine-scale features of the canvas; the admittedly higher number of artifacts is due to the increased difficulty to “read” a twill versus a standard weave pattern, as well as a weaker canvas signal on the X-ray.

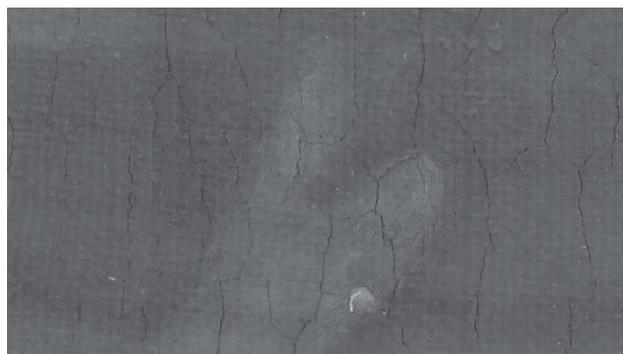
For our final example, we apply the 2DST-based canvas analysis to *The Peruzzi Altarpiece* by Giotto di Bondone and his assistants. The altarpiece is in the collection of the North Carolina Museum of Art; see Figure 7 for the altarpiece as well as the X-ray images used in the analysis. This is a painting on wood panel, but the ground of traditional white gesso was applied over a coarsely woven fabric interlayer glued to a poplar panel. We carried out a canvas analysis on the fabric interlayer, likely a handwoven linen cloth. The results of a canvas analysis based on the synchrosqueezed transform are shown in Figure 8. This example is much more challenging than the previous ones, since the X-ray intensity contributed by the canvas is much weaker because the ground does not contain lead; see, e.g., a detail of the X-ray image of the Christ panel in Figure 9. The canvas is barely visible, in sharp contrast to the X-ray images in, e.g., Figures 1(a) or 5. All panels except the central Christ panel are cradled; the wood texture of these cradles interferes with the canvas pattern on the X-ray image, introducing an additional difficulty. This difficulty is reflected in our results: e.g., the vertical thread count for the central panel has much fewer artifacts

than those of the other panels (see Figure 8). In future work, we will explore carrying out a canvas analysis after signal-processing-based virtual cradle remove (see e.g., [29]).

One interesting ongoing art investigation debate concerning this altarpiece is the relative position of the panels of John the Baptist and Francis of Assisi [Figure 7(d) and (e), respectively]. While the order shown in Figure 7 is the most commonly accepted [30], there have been alternative arguments that the Francis panel should be instead placed next to the central panel [Figure 7(c)]. As seen in X-rays, the grain of the wood typically can be used to set the relative position of panels in an altarpiece painted on a single plank of wood, but because the cradle pattern obscures an accurate reading of the X-rays of the Baptist and Francis, this proposed alternative orientation cannot be discounted. We wondered what ordering (if any) would be suggested by the canvas analysis. Under the assumption that the pieces of canvas are cut off consecutively from one larger piece of cloth, we investigated which arrangement provides the best matching. One plausible arrangement of the canvas is shown in Figure 10. Our analysis suggests that the canvas of the central panel should be rotated 90° clockwise to match with the other panels. (The larger height of the central panel, possibly exceeding the width of the cloth roll, may have necessitated this.) Moreover, a better match is achieved if the canvas of the panel of the Baptist is flipped horizontally (in other words, flipped front to back). Given our results, it seems unlikely that the Francis-panel



[FIG8] The canvas analysis result of the Giotto altarpiece. The deviation of (a) vertical thread angle, (b) vertical thread count, (c) horizontal thread angle, and (d) horizontal thread count. The panels are in the same order as in Figure 7.



[FIG9] A zoomed-in X-ray image of the central panel in *The Peruzzi Altarpiece*. The canvas texture is barely visible, even though the image is scaled such that the thread density is comparable with that of the zoomed-in X-ray in Figure 5.

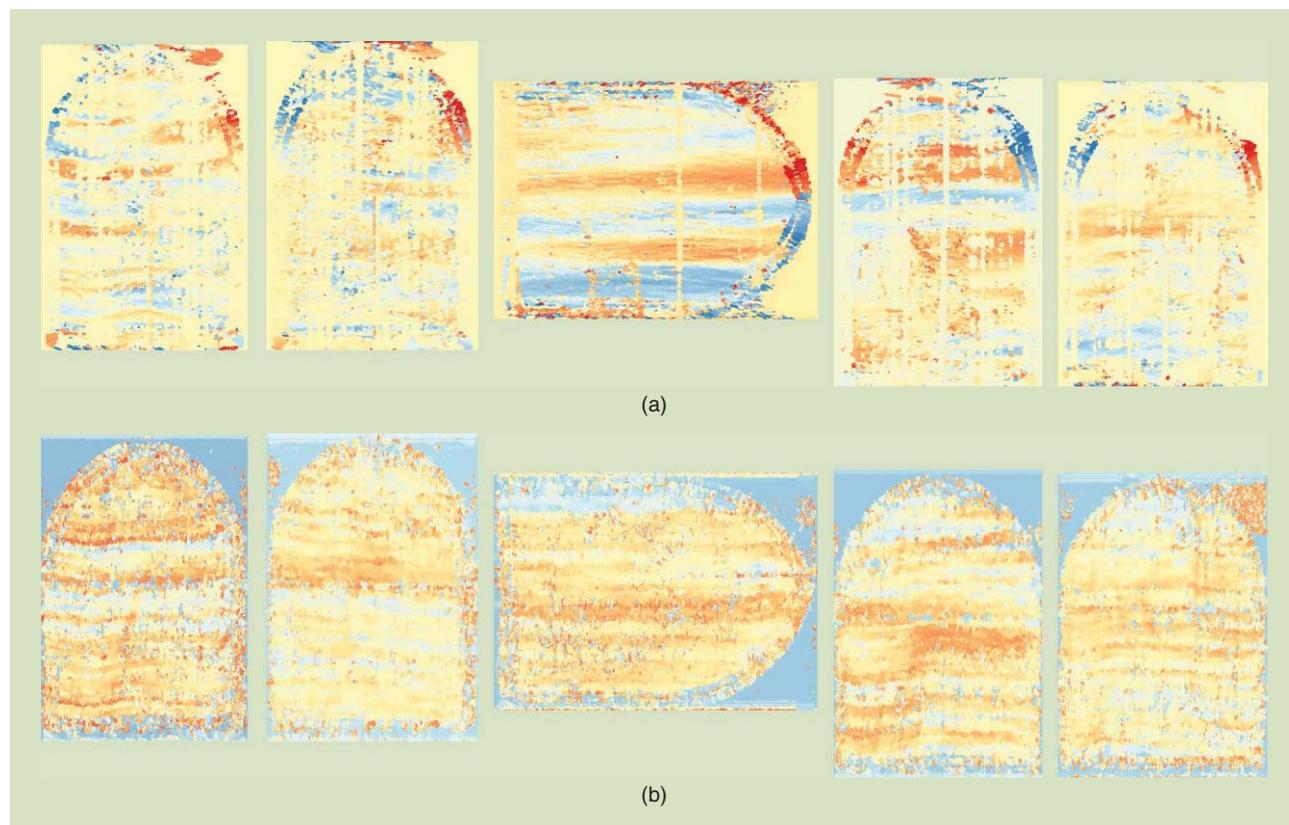
canvas would fit best to the left of the Baptist-panel canvas. A better, more precise, result will be possible after virtual cradle removal. Since pieces of the canvas could have been placed on the altarpiece plank in another order than their cutting sequence, evidence from even the most thorough study of the canvas roll arrangement would not be conclusive for the relative position of the panels themselves; nevertheless, it can play a significant role when combined with other elements in an exhaustive study.

CONCLUSIONS

We applied 2DSTs to quantitative canvas weave analysis for art investigations. The synchrosqueezed transforms offer a sharpened phase-space representation of the X-ray image of the paintings, which yields fine-scale characterization of thread count and thread angle of the canvas. We demonstrated the effectiveness of the method on art works by van Gogh, Vermeer, and Ryder. The tool is applied to *The Peruzzi Altarpiece* by Giotto and his assistants, to provide insight into the issue of panel arrangement.

ACKNOWLEDGMENTS

We are grateful to C. Richard Johnson, Jr. for providing X-ray images in the RKD data set and for his help and encouragement throughout this project. We would like to thank Bruno Cornelis and Rachel Yin for helpful discussions. We thank Noelle Ocon, conservator of paintings at the North Carolina Museum of Art, for mozaicing Giotto X-rays and also John Taormina, the director of Visual Media Center at Duke University, for assistance in scanning the X-ray film of *The Pasture* by Albert P. Ryder. The work of Ingrid Daubechies was partly supported by a grant from the Air Force Office of Scientific Research and by the National Science Foundation (NSF) award DMS-1320655. The work of Jianfeng Lu was supported in part by the Alfred P. Sloan Foundation and the NSF under award DMS-1312659. The work of Haizhao Yang and



[FIG10] A candidate canvas matching and arrangement for *The Peruzzi Altarpiece* by Giotto and his assistants. (a) The deviation of weft thread angle. (b) The deviation of warp thread count. (The weft thread count and warp thread angle are not shown as they are less helpful in inferring a possible arrangement.) The canvas pieces from left to right correspond to the panels for John the Evangelist, the Virgin Mary, Christ in Majesty, John the Baptist, and Francis of Assisi (in that order). The canvas of the central panel is rotated clockwise by 90° , and that of the Baptist is flipped horizontally.

Lexing Ying was partially supported by the NSF under award DMS-1328230 and the U.S. Department of Energy Advanced Scientific Computing Research program under award DE-FC02-13ER26134/DE-SC0009409.

AUTHORS

Haizhao Yang (haizhao@math.stanford.edu) received the B.S. degree from Shanghai Jiao Tong University, China, in 2010 and the M.S. degree from the University of Texas at Austin in 2012 both in mathematics. He is currently a Ph.D. student in the Mathematics Department at Stanford University, California.

Jianfeng Lu (jianfeng@math.duke.edu) received the B.S. degree in mathematics from Peking University, Beijing, China, in 2005 and the Ph.D. degree in applied mathematics from Princeton University, New Jersey, in 2009. He is an assistant professor of mathematics, chemistry, and physics at Duke University, Durham, North Carolina. Before joining Duke University in 2012, he was a Courant Instructor at New York University.

William P. Brown (william.p.brown@ncdcr.gov) received the B.A. degree from the University of Virginia in 1978 and the M.A. degree and certificate of advanced studies in conservation from the State University of New York College at Buffalo, New York, in 1989. He is currently the chief conservator of the Art Conservation Center of the North Carolina Museum of Art where he is responsible for the long-term preservation of the collection. In 2013, he established the Art+Science Initiative with Duke University for better understanding the application and degradation of artist materials.

Ingrid Daubechies (ingrid@math.duke.edu) received the B.S. and Ph.D. degrees in physics for the Vrije Universiteit Brussel (Belgium), where she also started her research career. In the United States, she has previously been with AT&T Bell Labs (1987–1994), Rutgers University (1991–1993), and Princeton University (1994–2010). She received the Benjamin Franklin Medal for Electrical Engineering and the James Kilby Medal of the IEEE Signal Processing Society, both in 2011. She is a James B. Duke Professor of Mathematics at Duke University and a Fellow of the IEEE.

Lexing Ying (lexing@stanford.edu) received the B.S. degree from Shanghai Jiao Tong University in 1998 and the Ph.D. degree from New York University in 2004. He was a postdoctoral scholar at the California Institute of Technology from 2004 to 2006. He was a faculty member in the Department of Mathematics at the University of Texas at Austin between 2006 and 2012. Currently, he is a professor of mathematics at Stanford University.

REFERENCES

- [1] K. Lister, C. Peres, and I. Fiedler, "Appendix: Tracing an interaction: Supporting evidence, experimental grounds," in *Van Gogh and Gauguin: The Studio of the South*, D. Druick and P. Zegers, Eds. London: Thames & Hudson, 2001, pp. 354–369.
- [2] E. van de Wetering, *Rembrandt: The Painter at Work*. Amsterdam, The Netherlands: Amsterdam Univ. Press, 1997.
- [3] A. Kirsh and R. Levenson, *Seeing Through Paintings: Physical Examination in Art Historical Studies*. New Haven, CT: Yale Univ. Press, 2000.
- [4] D. Johnson, C. Johnson, Jr., A. Klein, W. Sethares, H. Lee, and E. Hendriks, "A thread counting algorithm for art forensics," in *Proc. IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop (DSP/SPE 2009)*, Jan. 2009, pp. 679–684.
- [5] D. H. Johnson, R. G. Erdmann, and C. R. Johnson, Jr., "Whole-painting canvas analysis using high- and low-level features," in *Proc. 36th Int. Conf. Acoustics, Speech, Signal Processing*, 2011, pp. 969–972.

- [6] D. H. Johnson, C. R. Johnson, Jr., and R. G. Erdmann, "Weave analysis of paintings on canvas from radiographs," *Signal Process.*, vol. 93, no. 3, pp. 527–540, 2013.

- [7] E. Hendriks, L. Jansen, J. Salvant, E. Ravaud, M. Eveno, M. Menu, I. Fielder, M. Geldof, L. Megens, M. van Bommel, C. R. Johnson, Jr., and D. H. Johnson, "A comparative study of Vincent van Gogh's Bedroom Series," in *Studying Old Master Paintings: Technology and Practice—The National Gallery Technical Bulletin 30th Anniversary Conf. Postprints*, M. Spring, Ed. London: Archetype Publications, 2011, pp. 237–243.

- [8] L. van Tilborgh, T. Meedendorp, E. Hendriks, D. H. Johnson, C. R. Johnson, Jr., and R. G. Erdmann, "Weave matching and dating of van Gogh's paintings: An interdisciplinary approach," *Burlington Mag.*, vol. CLIV, pp. 112–122, Feb. 2012.

- [9] P. P. D'Ors, C. R. Johnson, Jr., and D. H. Johnson, "Velazquez in Fraga: A new hypothesis about the portraits of El Primo and Philip IV," *Burlington Mag.*, vol. CLIV, pp. 620–625, Sept. 2012.

- [10] W. Liedtke, C. R. Johnson, Jr., and D. H. Johnson, "Canvas matches in Vermeer: A case study in the computer analysis of fabric supports," *Metropol. Museum J.*, vol. 47, pp. 99–106, 2012.

- [11] D. H. Johnson, E. Hendriks, and C. R. Johnson, Jr., "Interpreting canvas weave matches," *Art Matters*, vol. 5, pp. 53–61, 2013.

- [12] C. R. Johnson, Jr., D. H. Johnson, I. Verslype, R. Lugtigheid, and R. G. Erdmann, "Detecting weft snakes," *Art Matters*, vol. 5, pp. 48–52, 2013.

- [13] C. R. Johnson Jr., E. Hendriks, P. Noble, and M. Franken, "Advances in computer-assisted canvas examination: Thread counting algorithms," in *Proc. 37th Annu. Meeting American Institute Conservation Historic and Artistic Works*, Los Angeles, CA, May 2009, pp. 2129.

- [14] D. H. Johnson, E. Hendriks, M. Geldof, and C. R. Johnson, Jr., "Do weave matches imply canvas roll matches?" in *Proc. 38th Annu. Meeting American Institute Conservation Historic and Artistic Works*, Milwaukee, WI, May 2010, pp. 97–102.

- [15] C. R. Johnson, Jr., D. H. Johnson, N. Hamashima, H. S. Yang, and E. Hendriks, "On the utility of spectral-maximum-based automated thread counting from X-rays of paintings on canvas," *Stud. Conserv.*, vol. 56, no. 2, pp. 104–114, 2011.

- [16] R. Erdmann, C. R. Johnson, Jr., M. Schafer, and J. Twilley, "Reuniting Poussin's Bacchanals painted for Cardinal Richelieu through quantitative canvas weave analysis," in *Proc. 41st Annu. Meeting American Institute Conservation Historic and Artistic Works*, Indianapolis, IN, May 2013.

- [17] B. Cornelis, A. Doms, A. Munteanu, J. Cornelis, and P. Schelkens, "Experimental study of canvas characterization for paintings," *Proc. SPIE*, vol. 7531, p. 753103, Feb. 2010.

- [18] H. Yang and L. Ying, "Synchrosqueezed wave packet transform for 2D mode decomposition," *SIAM J. Imaging Sci.*, vol. 6, no. 4, pp. 1979–2009, 2013.

- [19] H. Yang and L. Ying, "Synchrosqueezed curvelet transform for two-dimensional mode decomposition," *SIAM J. Math. Anal.*, vol. 46, no. 3, pp. 2052–2083, 2014.

- [20] H. Yang, J. Lu, and L. Ying. (2014). Crystal image analysis using 2d synchrosqueezing transforms. Preprint. [Online]. Available: <http://arxiv.org/abs/1402.1262>

- [21] H.-T. Wu, S.-S. Hseu, M.-Y. Bien, Y. R. Kou, and I. Daubechies, "Evaluating the physiological dynamics via synchrosqueezing: Prediction of the ventilator weaning," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 3, pp. 736–744, 2014.

- [22] Y.-T. Lin, H.-T. Wu, J. Tsao, H.-W. Yen, and S.-S. Hseu, "Time-varying spectral analysis revealing differential effects of sevoflurane anesthesia: nonrhythmic to rhythmic ratio," *Acta Anaesthesiol. Scand.*, vol. 58, no. 2, pp. 157–167, 2014.

- [23] P. Flandrin, *Time-Frequency/Time-Scale Analysis*. London: Academic Press, 1998.

- [24] H. Yang and L. Ying. (2014). Robustness analysis of synchrosqueezed transforms. Preprint. [Online]. Available: <http://arxiv.org/abs/1410.5939>

- [25] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Processing Mag.*, vol. 30, no. 6, pp. 32–41, 2013.

- [26] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds. Boca Raton, FL: CRC Press, 1996, pp. 527–546.

- [27] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool," *Appl. Comp. Harmonic Anal.*, vol. 30, no. 2, pp. 243–261, 2011.

- [28] RKD (Netherlands Institute for Art History). (14 July, 2014.) Image sharing. [Online]. Available: <http://english.rkd.nl/Services/image-sharing>

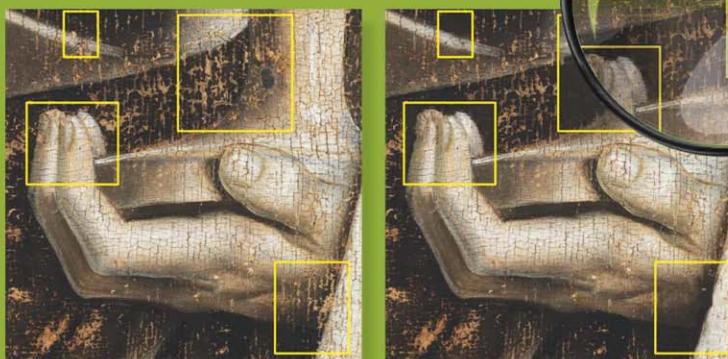
- [29] R. Yin, D. Dunson, B. Cornelis, W. Brown, N. Ocon, and I. Daubechies, "Digital cradle removal in x-ray images of art paintings," in *Proc. IEEE Int. Conf. Image Processing*, Paris, Oct. 2014, pp. 4299–4303.

- [30] D. Steel, "Catalogue entry: 1. Giotto and his workshop, 'Christ blessing with Saint John the Evangelist, the Virgin Mary, Saint John the Baptist, and Saint Francis (Peruzzi Altarpiece)'," in *Florence at the Dawn of the Renaissance, Painting and Illumination: 1300-1350*, C. Sciacca, Ed. Los Angeles, CA: The J. Paul Getty Museum, 2012, pp. 24–28.

[Hafeez Anwar, Sebastian Zambanini, Martin Kampel, and Klaus Vondrovec]

Ancient Coin Classification Using Reverse Motif Recognition

© GHENT, KATHEDRALE KERKFABRIEK; PHOTO COURTESY OF KIK-IRPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

[Image-based classification of Roman Republican coins]

We propose a holistic system to classify ancient Roman Republican coins based on their reverse-side motifs. The bag-of-visual-words (BoW) model is enriched with spatial information to increase the discriminative power of the coin image representation. This is achieved by combining a spatial pooling scheme with co-occurrence encoding of visual words. We specifically address the required geometric invariance properties of image-based ancient coin classification, as coins from different collections can be located at differing image locations, have various scales in the images

and can undergo various in-plane rotations. We evaluate our method on a data set of 2,224 coin images from three different sources. The experimental results show that our proposed image representation is more discriminative than the traditional bag-of-visual-words model while still being invariant to the mentioned geometric transformations. For 29 motifs, the system achieves a classification rate of 82%. It is considered to act as a helpful tool for numismatists in the near future, which facilitates and supports the traditional coin classification process by a faster presorting of coins.

INTRODUCTION

In ancient times, coins had become the central embodiment of money, starting from around the seventh century B.C. in Greece and spreading over other civilizations like the Roman

Digital Object Identifier 10.1109/MSP.2015.2409331

Date of publication: 15 June 2015

Empire, Byzantium, India, or China [1]. While ancient coins were everyday objects in the past, today they are also considered as pieces of art that reflect the individualism of the engravers who manually cut the dies used for minting the coins [2]. Roman coins, for instance, often depict portraits of gods, influential persons, or historical events, similar to sculptures or paintings from this era. [2]. The classification of coins according to standard reference books is the fundamental work of coin experts [2] since this provides additional information such as accurate dating, political background, or minting place. However, classifying ancient coins is a highly complex task that requires years of experience in the entire field of numismatics [3]. Additionally, efficiency is a vital issue, both among humans and machines, especially when it comes to the processing of large coin collections or maybe even larger hoard finds. As a substantial part of numismatic coin analysis, coin classification can be supported and facilitated by an automatic system.

The basic elements of Roman Republican coin design are depicted in Figure 1. The obverse coin sides frequently show a portrait of the goddess Roma as the central motif or, at a later stage, that of a historic person, whereas the reverse shows some kind of significant objects or scenes. Usually those main motifs are combined with a legend. In addition to that, there are minor images, such as a tiny club, or numerals, which are referred to as *symbols* in classical numismatics. Symbols split up otherwise large and uniform coin issues of a specific year.

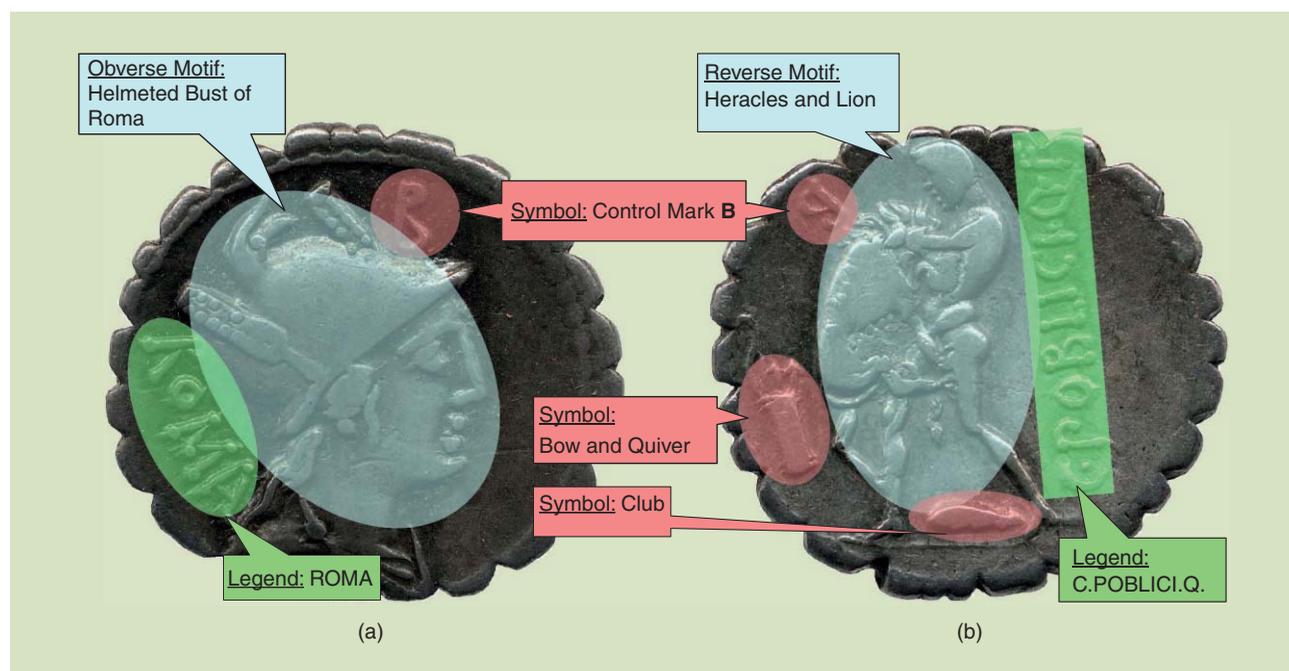
WHILE ANCIENT COINS WERE EVERYDAY OBJECTS IN THE PAST, TODAY THEY ARE ALSO CONSIDERED AS PIECES OF ART THAT REFLECT THE INDIVIDUALISM OF THE ENGRAVERS WHO MANUALLY CUT THE DIES USED FOR MINTING THE COINS.

In contrast to existing works where the classification is based either on obverse portraits [4], [5], legends [6], [7], or dense image-to-image matching [8], [9], we exploit the reverse-side motifs for classification in this articles. Accordingly, the presented method closely mirrors the human approach, as the reverse images usually hold a critical

amount of information that is required for successful numismatic classification. This is also true for periods other than the Roman Republican time. Naturally the central (reverse) motif is most prominent in all descriptions and, compared to legends or minor symbols, larger and more resistant against the inevitable degree of wear that is generally found on ancient coins. If manual numismatic classification is not possible for exactly that reason, it is commonly agreed upon to give a description of the visible parts (usually the central motif) and, if possible, the resulting probable reference numbers from the reference books [2].

PRIOR WORK AND CONTRIBUTIONS

In contrast to methods dedicated to present-day coins [10], [11], image-based classification of ancient coins has become a recent research interest, owing to the higher complexity of the problem due to the challenging conditions of ancient coins. Ancient coins do not have a rigid shape and are worn to a certain degree because of their age. Consequently, it was experimentally shown by [3] that the success of classification methods for present-day coins cannot be transferred to the



[FIG1] An example of the basic coin elements on the (a) obverse and (b) reverse side of a Roman Republican coin.

domain of ancient coins. The first method exclusively dedicated to ancient coins was proposed by Kampel and Zaharieva [12]. The method compares two coin images by establishing correspondences between them and taking their total number as similarity value.

Given such a similarity function, classification is achieved in an

CLASSIFYING ANCIENT COINS IS A HIGHLY COMPLEX TASK THAT REQUIRES YEARS OF EXPERIENCE IN THE ENTIRE FIELD OF NUMISMATICS.

comparing the image signals, one can leverage a priori known

exemplar-based manner where a query image is compared to class exemplar images in the data set. This idea was later extended to more sophisticated similarity functions that also take geometric constraints into account [8], [9].

Alternatively, or in addition to



[FIG2] Reverse motifs used for classification.

semantic background information for classification. Arandjelović [6] and Kavelar et al. [7] proposed to exploit the legend shown on ancient coins. The background information here is a lexicon of known legend words linked with coin classes, which are detected by means of graphical models that integrate likelihoods of letter appearances to words. However, the two methods are designed for two different Roman coin periods: the method of Kavelar et al. [7] is designed for Republican coins and has to consider more degrees of freedom in the word search than the method of Arandjelović [6] for Imperial coins. On Imperial coins, the legend is arranged along the coin border and the image can be normalized for orientation by means of a log-polar transformation.

Apart from the legend, for Roman Imperial coins the person portrayed on the obverse coin side has been used for classification [4], [5], as this is typically the issuer defining the coin class [13]. The method of Arandjelović [4] relies on detected scale-invariant feature transform (SIFT) features [14] that are quantized into a fixed vocabulary of visual words. The spatial configuration of visual words is then encoded by locally biased directional histograms (LBDHs), which are again subject to vocabulary creation and the histogram of LBDH words serves as final image feature. However, a recently proposed method by Kim and Pavlovic [5] outperformed this method. Their approaches are more adapted to the field of face recognition and achieve recognition rate of 82–86% on a 15-class-problem by using the deformable parts model [5], [15].

Our proposed method is different from prior work both from a contextual and methodological perspective. We aim at exploiting a different semantic information for classification instead of the legend or the person on the obverse, particularly the motifs on the reverse of Roman Republican coins as shown in Figure 2. This is an essential part of coin descriptions and thus highly practical for classification. On the methodological side, we aim at the rotation-invariant spatial encoding of local features, as coin rotations cannot be handled by the best-performing methods for image-to-image coin matching [9] and obverse side recognition [5].

ROTATION-INVARIANT SPATIAL EXTENSIONS TO THE BoW IMAGE REPRESENTATION FOR ANCIENT COIN CLASSIFICATION

Our proposed method is based on the BoW model, which has become a standard paradigm for image classification in the last decade [16]. Local features sampled from a set of images are quantized using a clustering technique such as the k -means. These cluster centers are called the *visual words* and the collection of all visual words is called the *visual vocabulary*. This visual vocabulary is then used to represent novel images. From a given image, local features are extracted and

visual words from the vocabulary are assigned to these features using a similarity measure such as the Euclidean distance. Finally, the image is represented as a histogram where each bin represents the count of the respective visual word in the image. This histogram is called the *histogram of visual words* or the *BoW*, and the size of this histogram is equal to the size of the visual vocabulary.

A basic problem in the BoW image representation is the lack of spatial information of the quantized local features [17]–[19]. The histogram counts the number of visual words without considering their spatial positions in the two-dimensional (2-D) image space. However, the spatial information provides discriminating details in problems such as object recognition-based image classification [17], [18] because objects have specific geometric structures. There are two main principles used by various methods to add spatial information to the BoW image representation. The first principle deals with

the splitting of the image space into subregions of various shapes [17], [18]. From each subregion, statistics of visual words are collected to form the final histogram of visual words. Since these subregions are either rectangular [17], [20] or log-polar [18], the methods based on such partitionings are not invariant to translations and image rotations. The second principle of incorporating spatial information to the BoW image representation is based on modeling various relationships of the visual words such as their co-occurrence [19] or their geometric relationships [21]. However, modeling the co-occurrences of visual words in the case of large vocabularies proves to be computationally expensive [19]. The methods based on geometric relationships of visual words such as the one proposed by Khan et al. [21] are also unable to achieve invariance to image rotations.

In this article, the recognition of the reverse motifs of ancient coins is carried out by using the BoW image representation. Due to the specific geometric structures of the reverse motifs, spatial information incorporated to the BoW image representation results in improved performance. However, such spatial information should be added by considering certain issues that are specific to ancient coins. Due to flat and circular nature of the coins, they can exhibit rotations in the image plane. In addition to rotations, the images can come from different data sets and thus the coin region can be in different scales and at different image positions. Consequently, the method should be invariant to changes of the coin position, rotation, and scale.

Our proposed method for adding spatial information to BoW is a combination of both the main principles, i.e., splitting the image space into subregions and modeling the relationships of visual words. To this end, we propose a three-step strategy to achieve an image representation for ancient coin classification

IMAGE-BASED CLASSIFICATION OF ANCIENT COINS HAS BECOME A RECENT RESEARCH INTEREST, OWING TO THE HIGHER COMPLEXITY OF THE PROBLEM DUE TO THE CHALLENGING CONDITIONS OF ANCIENT COINS.

that is invariant to translation, scale changes, and image rotations. We first use the automatic coin image segmentation [22] to segment the coin from the background, which helps to achieve invariance to scale changes and translation. Second, we apply the circular tiling scheme [20] over the segmented coin image to increase the discriminative power of the model while maintaining rotation invariance. Finally, the rotation-invariant geometric relationships of visual words [23], [24] in each subregion of the circular tiling scheme are modeled to obtain the final image representation. In the following, we will give a brief description of each step of the proposed strategy.

A BASIC PROBLEM IN THE BOW IMAGE REPRESENTATION IS THE LACK OF SPATIAL INFORMATION OF THE QUANTIZED LOCAL FEATURES.

for the automatic segmentation of ancient coins. This method is based on the following two assumptions:

- The area of the image depicting the coin contains more information contents than the rest of the image.
- The coin is the most circular object present in the image.

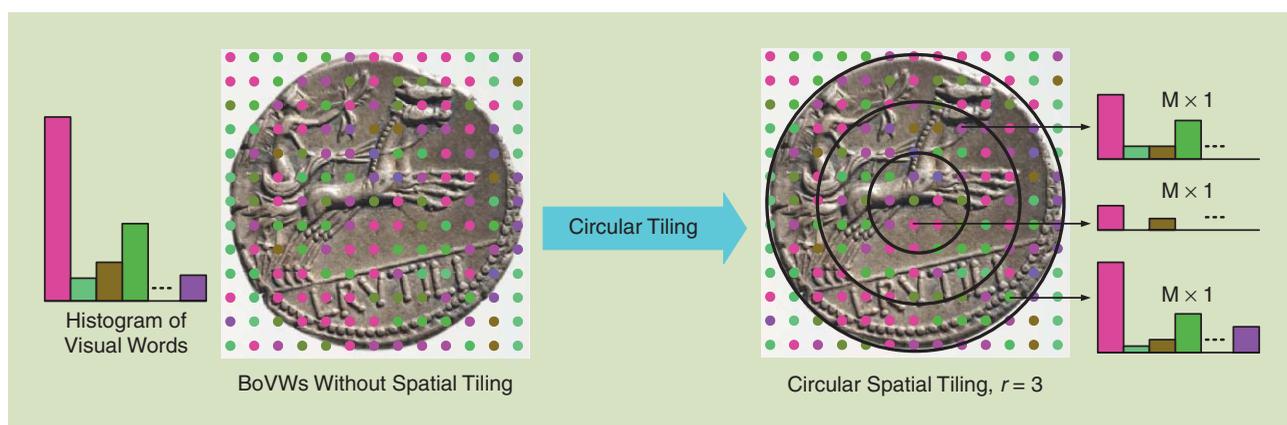
Consequently, the automatic coin image segmentation method is a two step process. In the first step, two filters are applied to the image to measure the local information content. The first filter measures the local entropy while the second one measures the range of gray values. The outputs of both filters are summed up to obtain the intensity image which is then normalized to the range zero and one. To achieve the final segmentation, seven empirically defined threshold values are applied to the normalized intensity image and a confidence score is calculated for each achieved segmentation. This confidence score is based on the form factor [25] calculated from

AUTOMATIC COIN IMAGE SEGMENTATION

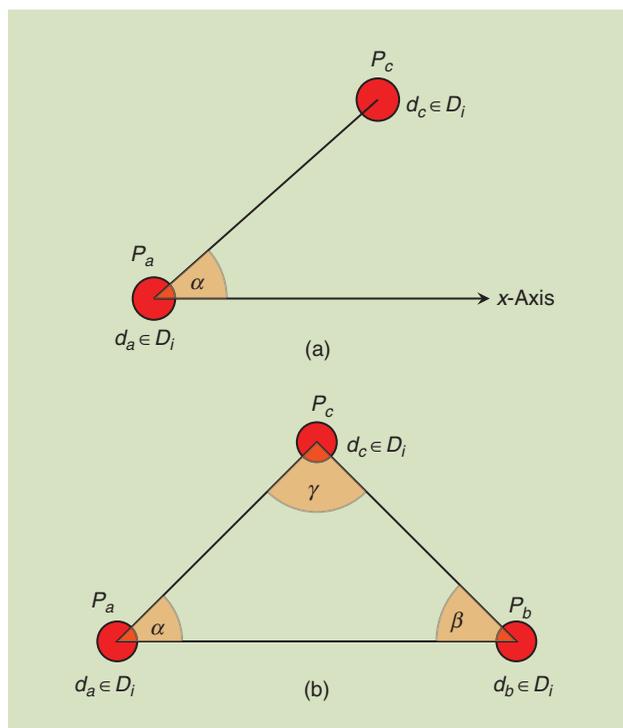
Coin image segmentation is used as a preprocessing step to achieve invariance to translation and scale changes. The method proposed in [22] is used as it is specifically proposed



[FIG3] Images of ancient coins at various scales and positions along with their segmentation masks.



[FIG4] The BoW image representation with and without circular tiling.



[FIG5] (a) PIWs and (b) TIWs.

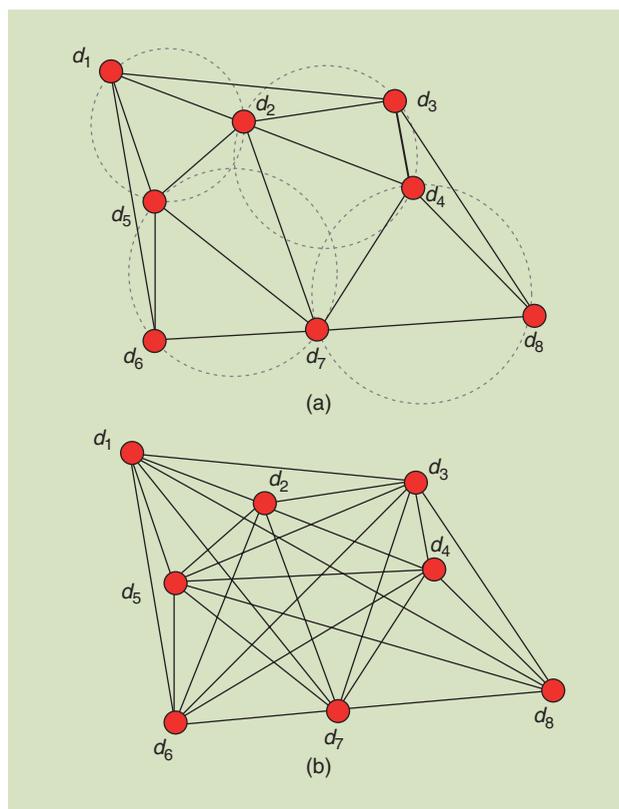
the area and perimeter of the binary mask. The form factor is sensitive to both the elongation and jaggedness of the border. It is equal to one for a circle and less than one for any other shape. Since the shape of the coins is supposed to be circular or nearly circular, the form factor is a reasonable choice for the confidence measurement. Finally, a segmentation is only accepted if the area of the segmented region is less than 90% and more than 5% of the image area. Figure 3 depicts two ancient coins and their respective segmentation masks demonstrating the effectiveness of the segmentation method for the images of ancient coins at arbitrary scales and positions. With the known location of the coin, the image is cropped and normalized to a specified standard size of 480×480 to achieve the required invariance to object location and scale.

CIRCULAR TILING

Once the coin image is normalized with respect to location and scale, circular tiling is applied. In the circular tiling scheme, concentric circular subregions are imposed over the BoW image representation as shown in Figure 4. For a vocabulary size of M , a histogram of M visual words is generated for each circular tiling. These histograms are then concatenated in a single feature vector of length $M \cdot r$.

GEOMETRIC RELATIONSHIP OF IDENTICAL VISUAL WORDS

The angles of a triangle are invariant to scale changes and rotations. Tao and Grosky [26] use the triangular relationships of identical color patches of an image to construct the so-called anglograms for spatial color indexing. The

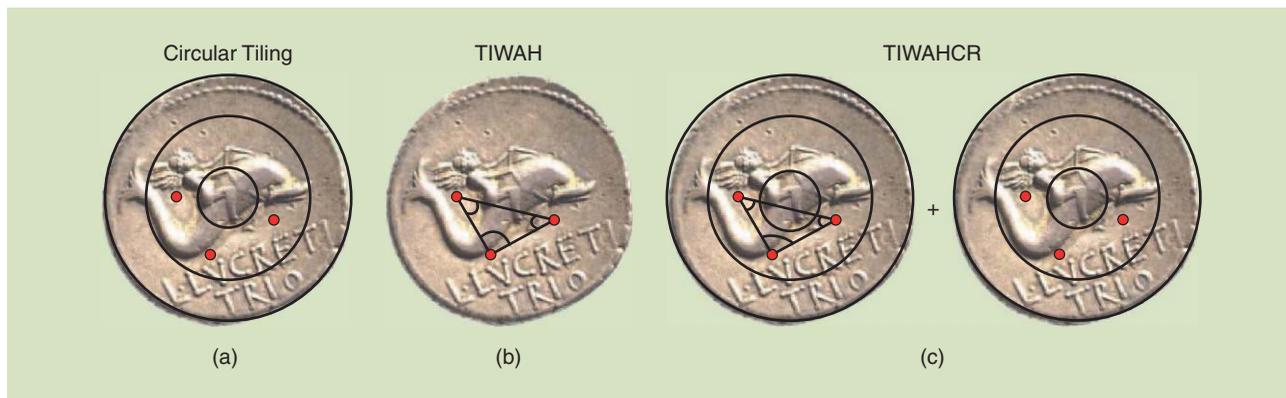


[FIG6] Triangulation methods. (a) Delaunay triangulation. (b) Combinatorial triangulation.

anglograms-based image representation is used for scale- and rotation-invariant image retrieval. To achieve the triangulation among identical image patches, they use the Delaunay triangulation, which is an efficient and well-known triangulation method from computational geometry [27]. In the BoW model, identical visual words represent similar image patches. To add spatial information to the BoW model, Khan et al. [21] proposed the use of identical visual words. They use the angles made by pairwise identical visual words (PIWs) for the construction of the pairwise identical visual words angles histogram (PIWAH) to represent the image. PIWAH is not invariant to image rotations because it is made from angles that are calculated with respect to the x -axis. Inspired by the idea of Tao and Grosky [26], we consider three identical words to calculate the angles and call them *triplets of identical visual words (TIWs)*, as shown in Figure 5. This will achieve a rotation-invariant triangular relationship among the words of a given triplet. Based on the angles calculated among members of each TIW, we produce the angles histogram in a similar manner as proposed by Khan et al. [21] and call them *triplets of identical visual words angles histogram (TIWAH)*.

In the BoW model, a visual vocabulary $\text{voc} = \{v_1, v_2, v_3, \dots, v_M\}$ consists of M visual words. A given image is first represented as a set of descriptors

$$I = \{d_1, d_2, d_3, \dots, d_N\}, \tag{1}$$



[FIG7] (a) Simple circular tiling, (b) TIWAH, and (c) TIWAHCR.

where N is the total number of descriptors. A given descriptor d_k is then mapped to a visual word v_i using a similarity measures like the Euclidean distance as follows:

$$v(d_k) = \underset{v \in \text{voc}}{\operatorname{argmin}} \operatorname{Dist}(v, d_k), \quad (2)$$

where d_k is the k th descriptor in the image and $v(d_k)$ is the visual word assigned to this descriptor based on the distance $\operatorname{Dist}(v, d_k)$. The value of the bin b_i of this histogram gives the number of occurrences of a visual word v_i in an image. Let D_i be the set of all descriptors mapped to a visual word v_i , then the i th bin of the histogram of visual words b_i , is the cardinality of the set D_i .

$$b_i = \operatorname{Card}(D_i) \text{ where } D_i = \{d_k, k \in [1, \dots, N] \mid v(d_k) = v_i\}. \quad (3)$$

We previously proposed [23] to use all the distinct pairs of three descriptors from set D_i to calculate angles between the spatial positions of the descriptors as shown in Figure 6. We call that method *combinatorial triangulation*, as the triangulation is done for all the three distinct pairs of descriptors belonging to a given visual word. The spatial position of a descriptor is given by its position on the dense sampling grid. The set of all TIWs related to a visual word v_i is defined as

$$\operatorname{TIW}_i = \{(P_a, P_b, P_c) \mid (d_a, d_b, d_c) \in D_i^3, d_a \neq d_b \neq d_c\}, \quad (4)$$

where P_a, P_b , and P_c are the spatial positions of the descriptors d_a, d_b , and d_c , respectively. The value of the i th bin of the histogram shows the frequency of the visual word v_i . Therefore, in case of *combinatorial triangulation*, the cardinality of TIW_i is ${}^{b_i}C_3$, which is the number of all possible pairs of three distinct elements among the elements of D_i . The positions of the elements of each pair make a triangle. Calculating angles for such a huge number of triangles is time-consuming. For instance, if the cardinality b_i of the set D_i is 80 then the number of unique three combinations is 82,160. Consequently, for the computation of the angles among TIWs of a given visual word, we propose to use the Delaunay triangulation where the number of triangles is much

smaller. In Delaunay triangulation, the three points should not be collinear and the circumscribed circle defined by the three points should not contain any other point. The principles of the Delaunay triangulation significantly reduce the number of triangles for angle computation among TIWs. Figure 6 shows both the Delaunay and the combinatorial triangulations. It can be observed that for eight descriptors belonging to a visual word, combinatorial triangulation results in 56 triangles while the Delaunay triangulation results in nine triangles. The angles of all the triangles are calculated using the law of cosines. The angles histogram is built from these angles for which the bins are empirically chosen between 0 and 180°. The angles histogram for a specific word v_i is named as TIWAH_i . The i th bin of the histogram of visual words associated with visual word v_i is replaced with TIWAH_i in such a way that the spatial information is added without altering the frequency information of v_i . Finally TIWAH_i of all the visual words are combined to represent a given image

$$\operatorname{TIWAH} = (\psi_1 \operatorname{TIWAH}_1, \psi_2 \operatorname{TIWAH}_2, \dots, \psi_M \operatorname{TIWAH}_M) \quad (5)$$

$$\text{where } \psi_i = \frac{b_i}{\|\operatorname{TIWAH}_i\|},$$

where ψ_i is the normalization coefficient. For a visual vocabulary of size M , if the number of bins in angles histogram is θ , then the size of the TIWAH is $M\theta$.

In each partition or tiling of the circular tiling scheme, the geometric relationships of identical visual words are modeled using TIWAH as shown in Figure 7. The TIWs are now limited to the circular tilings only. Doing so not only extracts more information from the circular tilings but also reduces the number of unique combinations of TIWs. Finally, the histogram of visual words and TIWAH of all the circular tilings are concatenated as described in (6). We call the final histogram TIWAHCR, which contains the information of TIWs and visual words other than TIWs for a given circular tiling.

$$\operatorname{TIWAHCR} = [\operatorname{TIWAH}^1, \operatorname{TIWAH}^2, \dots, \operatorname{TIWAH}^r, \operatorname{CR}^1, \operatorname{CR}^2, \dots, \operatorname{CR}^r], \quad (6)$$



[FIG8] The variations in reverse motifs of coins obtained from different sources: (a) the Vienna Museum of Fine Arts; (b) the British Museum of London; and (c) acsearch, an online auction Web site for ancient coins.

where for r number of circular tilings, $TIWAH^j$ represents the angles histogram and CR^j represents the histogram of visual words for the j th circular tiling. Therefore for a vocabulary size of M , number of circular tilings r , and number of bins in angles histogram θ , the length of TIWAHCR is $(r \cdot M \cdot \theta) + (r \cdot M)$.

DATA SET

Our data set belongs to the largest and the most diverse among the works that deal with image-based analysis of

ancient coins [3], [4], [6]–[8], [12], [22]–[24]. Our data set consists of 2,224 images that belong to 29 different reverse motifs. Only the recent work by Kim and Pavlovic [5] uses more images but with only 15 various obverse motifs. Our images have been collected from the following three sources:

1) *The Vienna Museum of Fine Arts* (Kunsthistorisches Museum Wien): The museum has 6,000 Roman Republican coins from which the images of 4,200 coins were collected for the Image-Based Classification of Ancient Coins (ILAC) project

[TABLE 1] AVERAGE TIME IN SECONDS TAKEN BY EACH TRIANGULATION SCHEME FOR 29 IMAGES.

COMBINATORIAL TRIANGULATION	2,860
COMBINATORIALF TRIANGULATION	41.6
DELAUNAY TRIANGULATION	6.4

The bold number signifies the least time taken by the Delaunay triangulation.

[28]. In our data set, 1,014 coin images are from this museum. 2) *The British Museum London*: The Department of Coins and Medals at the British Museum is home to one of the largest collections of ancient coins in the world. 857 of the images in our data set are from the British Museum. 3) *acsearch*: acsearch (<http://www.acsearch.info>) is an online auction Web site for ancient coins. It is also one of the largest online sources for images of the ancient coins. Our data set contains 353 coin images from this Web site.

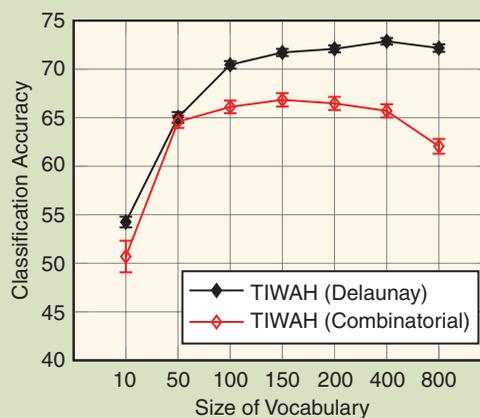
Sample coin images from all three sources are shown in Figure 8 where the differences among the images can be observed. The coin images acquired from the Vienna Museum of Fine Arts differ from those of the British Museum and acsearch due to the background and illumination conditions. The images from the British museum vary from one another due to strong illumination changes. Similarly, the dirt on the coins in the images obtained from acsearch also cause variations among coin images of the same class. Finally, a significant variation can be observed among the coin images from the Vienna Museum due to rotations.

EXPERIMENTS AND RESULTS

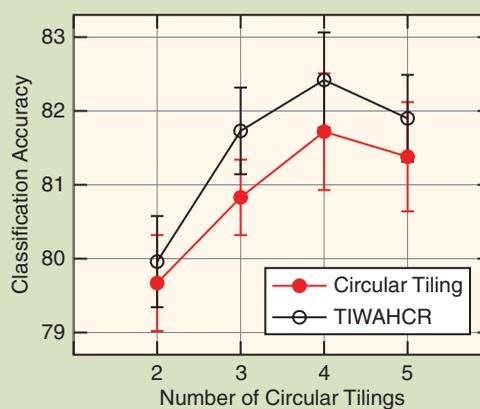
The data set is divided into two disjoint training and test sets. The size of the training set is 1,426 while the test set consists of the remaining 798 images. The histogram representations (TIWAHCR) of these images are obtained using the method mentioned in the section “Rotation-Invariant Spatial Extensions to the BoW Image Representation for Ancient Coin Classification.” These histograms are then used as feature vectors for classification with a support vector machine. Experiments are performed for the following parameters.

COMPUTATIONAL COMPLEXITY OF THE TRIANGULATION METHODS

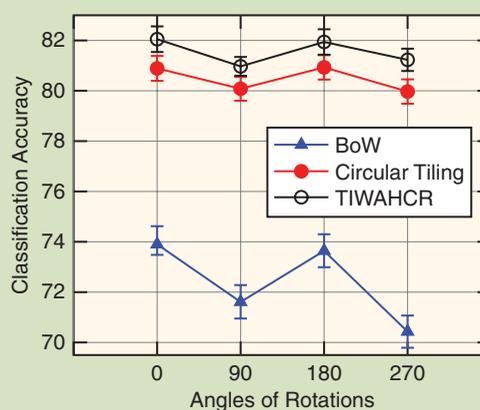
Since our proposed method relies upon the triangulation of identical visual words, in this section we compare the efficiency of both the combinatorial and Delaunay triangulations in terms of computation time. We collect 29 images from the whole data set by randomly selecting one image from each class. Images are of standard size, which is 480×480 . The TIWAH representation of these images is constructed using both the triangulation methods implemented in MATLAB. We also use a faster “C” implementation of the combinatorial triangulation and denote it by *combinatorialF* triangulation. The experiments are repeated ten times on a single core and the mean calculation time taken by each method is reported in Table 1. It can be concluded from these results that the Delaunay triangulation is even faster than the combinatorialF triangulation.



(a)



(b)



(c)

[FIG9] (a) The performances of various triangulation methods with respect to the size of vocabulary. (b) The performances of TIWAHCR and circular tiling with respect to the number of tilings. (c) The performances of various methods on rotated images.

SIZE OF VOCABULARY

The size of vocabulary is an important parameter in the BoW model [29]. We optimize for the size of vocabulary on our current data set by empirically selecting the values from the set {10, 50, 100, 150, 200, 400, 800}. For each value of the vocabulary size, classification runs are performed ten times and the

mean performances of the TIWAH constructed with both the types of triangulation methods are shown in Figure 9(a). TIWAH constructed with Delaunay triangulation is not only more efficient but also performs better than the one constructed with combinatorial triangulation. It can also be observed that no significant performance improvement is gained from the vocabulary sizes larger than 150.

NUMBER OF TILINGS

We also optimize for the number of tilings in both the simple circular tiling and TIWAHCR. The values for the number of tilings are empirically selected from {2, 3, 4, 5}. Experiments are performed ten times and the mean performances achieved by both the settings on all the empirically selected values are shown in Figure 9(b). The performance of TIWAHCR is superior to that of the simple circular tiling because it contains additional spatial information of the visual words in each tiling. On four tilings, TIWAHCR achieves the maximum performance. However with the increase in the number of tilings, its performance tends to converge towards that of the simple circular tiling. This is due to the fact that the width of the tilings decrease with the increase in their number thus decreasing the occurrences of identical words in any given tiling.

EVALUATION OF ROTATION-INVARIANCE

We generate synthetically rotated coin images to evaluate the rotation-invariance of the proposed method. The rotated coin images are produced by making one of them as reference and then rotating this reference image in 90° steps. The synthetically rotated images allow us to test stronger rotation differences as those already present in our data set (generally lower than 90°; see Figure 8). Experiments are repeated 20 times and the mean performances of the BoW, simple circular tiling and TIWAHCR are shown in Figure 9(c). The number of tilings in the circular tiling and TIWAHCR is three. The size of the vocabulary is 150 and at each iteration a new vocabulary is constructed. Local rotation-invariance is achieved by using rotation-invariant local features such as SIFT. The experimental results reflect the theoretical foundations of our rotation-invariant image representation and show that rotations have no significant influence on the classification performance. Our proposed method clearly outperforms the BoW model on rotated images while marginally performs better than the circular tiling.

THE ART EXPERT'S OPINION ON THE ACHIEVED RESULTS

Given the fundamental task of the already closed project [28], a rate of 80% accurate results proves that it is possible to gain a real benefit from the developed algorithms. Of course, to

push this rate to 98 or 99%, a multiple amount of reference images and probably more “fine-tuning” (e.g., their proper selection based on the degree of preservation) would be required.

So ultimately, this approach is of the highest interest, both in respect of future automated coin classification and by considering all other uses of coin matching, be that monitoring of (illegal) trade, retrieving scientific data from otherwise unspecified archives of images, or counting known specimens per type. The same procedure can potentially be extended to count coin-dies, which would lead to answering the question of how many coins were produced, which is an often-discussed issue of historic economies.

CONCLUSIONS

We presented a system for image-based classification of ancient coins using their reverse motifs. Due to the lack of spatial information in the BoW model and the requirements of the image-based classification of ancient coins, we proposed a method for adding spatial information to the BoW model, which is invariant to scale changes, image rotations, and translation. This information is added using a three-step strategy that involves the automatic coin segmentation, application of a circular tiling scheme over the segmented image, and modeling the triangular geometric relationship of identical visual

words in each tiling. It is shown that such a representation not only outperforms the BoW model but also was invariant to image rotations, scale changes, and translations. However, it was observed that the proposed method was more discriminating on smaller vocabulary sizes, e.g., by modeling spatial relationships of nonidentical words. In the future, we plan to address this issue so that it can also work with larger vocabularies. Furthermore, we also plan to apply our proposed method on other art-related problems such as sculpture recognition [30].

ACKNOWLEDGMENTS

This research has been supported by the Austrian Science Fund (FWF) under grant TRP140-N23-2010 (ILAC) and the Vienna Ph.D. School of Informatics, Vienna University of Technology, Austria (<http://www.informatik.tuwien.ac.at/teaching/phdschool>).

AUTHORS

Hafeez Anwar (hafeez@caa.tuwien.ac.at) received his B.S. degree in computer systems engineering from the University of Engineering and Technology Peshawar, Pakistan, and his M.S. degree in computer engineering from Myongji University, South Korea, where he also acted as a research assistant on a funded project related to the design and development of an intra-oral three-dimensional scanner. He is currently pursuing his Ph.D. degree at the Computer Vision Lab of the Vienna University of Technology, Austria, where he is funded by the scholarship program of the

THE SAME PROCEDURE CAN POTENTIALLY BE EXTENDED TO COUNT COIN-DIES, WHICH WOULD LEAD TO ANSWERING THE QUESTION OF HOW MANY COINS WERE PRODUCED, WHICH IS AN OFTEN-DISCUSSED ISSUE OF HISTORIC ECONOMIES.

Vienna Ph.D. School of Informatics. His research interest is in object category recognition-based image classification with application to cultural heritage.

Sebastian Zambanini (zamba@caa.tuwien.ac.at) received his M.S. degree in computer graphics and digital image processing in 2007 and his Ph.D. degree in 2014 from the Vienna University of Technology, Austria. Currently, he is employed as a university assistant at the Computer Vision Lab of the Vienna University of Technology. He has collaborated in various computer vision research projects addressing the areas of medical image processing, surveillance, ambient assisted living, and cultural heritage. His research interests comprise object recognition, visual feature extraction, video analysis, and gesture recognition, and he has coauthored more than 40 publications on these topics.

Martin Kampel (martin.kampel@tuwien.ac.at) received the Ph.D. degree in computer science (computer graphics, pattern recognition, and image processing) in 2003 and the “*venia docendi*” (habilitation) in applied computer science in 2009, both from the Vienna University of Technology, Austria. He is an associate professor (Privatdozent) of computer vision at the Institute for Computer Aided Automation, Vienna University of Technology, Austria, engaged in research, project leading, industry consulting, and teaching. His research interests are three-dimensional vision and cultural heritage applications, visual surveillance, and ambient assisted living. He has been working for more than ten years on the acquisition and reconstruction of small archaeological artefacts like pottery fragments or coins.

Klaus Vondrovec (klaus.vondrovec@khm.at) is the curator of ancient coins in the Coin Cabinet in the Museum of Fine Arts, Vienna, Austria. He has been working extensively in the field of classification, interpretation, and statistic analysis of Roman coin-finds in Austria. This has involved creating a scientific database-scheme that is currently being used by the Coin-Cabinet and the Numismatic Commission of the Austrian Academy of Sciences. Another result of his work is the online database dFMRÖ (since 2005) of the Numismatic Commission. Another field of activity was the digital photography of coins and their postprocessing up to publication that also involved the development of an automated plate-generating tool.

REFERENCES

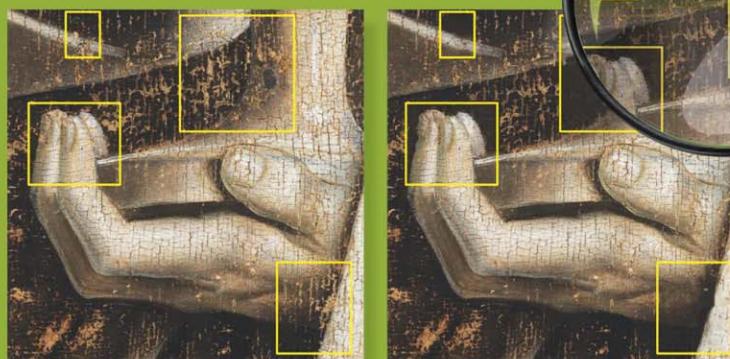
- [1] P. Grierson, *Numismatics*. Oxford, U.K.: Oxford Univ. Press, 1975.
- [2] M. H. Crawford, *Roman Republican Coinage*, 2 vols. Cambridge, U.K.: Cambridge Univ. Press, 1974.
- [3] M. Zaharieva, M. Kampel, and S. Zambanini, “Image-based recognition of ancient coins,” in *Proc. Int. Conf. Computer Analysis of Images and Patterns (CAIP)*, 2007, pp. 547–554.
- [4] O. Arandjelović, “Automatic attribution of ancient Roman imperial coins,” in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 1728–1734.
- [5] J. Kim and V. Pavlovic, “Improving ancient roman coin recognition with alignment and spatial encoding,” in *Proc. European Conf. Computer Vision (ECCV) Workshop on Computer Vision for Art Analysis*, 2014.
- [6] O. Arandjelović, “Reading ancient coins: Automatically identifying denarii using obverse legend seeded retrieval,” in *Proc. European Conf. Computer Vision (ECCV)*, 2012, pp. 317–330.
- [7] A. Kavelar, S. Zambanini, and M. Kampel, “Reading the legends of roman republican coins,” *J. Comput. Cult. Herit.*, vol. 7, no. 1, pp. 5:1–5:20, 2014.
- [8] S. Zambanini and M. Kampel, “Coarse-to-fine correspondence search for classifying ancient coins,” in *Proc. eHeritage Workshop, Asian Conf. Computer Vision (ACCV)*, 2012, pp. 25–36.
- [9] S. Zambanini, A. Kavelar, and M. Kampel, “Classifying ancient coins by local feature matching and pairwise geometric consistency evaluation,” in *Proc. IEEE Int. Conf. Pattern Recognition (ICPR)*, 2014, pp. 3032–3037.
- [10] M. Nölle, H. Penz, M. Rubik, K. Mayer, I. Holländer, and R. Granec, “Dagobert—A new coin recognition and sorting system,” in *Proc. Int. Conf. Digital Image Computing: Techniques and Applications (DICTA)*, 2003, pp. 329–338.
- [11] M. Reisert, O. Ronneberger, and H. Burkhardt, “An efficient gradient based registration technique for coin recognition,” in *Proc. MUSCLE CIS Coin Competition Workshop*, 2006, pp. 19–31.
- [12] M. Kampel and M. Zaharieva, “Recognizing ancient coins based on local features,” in *Proc. Int. Symp. Visual Computing (ISVC)*, 2008, pp. 11–22.
- [13] D. Sear, *Roman Coins and Their Values*, 5th ed. London: Spink & Son, 2003.
- [14] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [15] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [16] K. Grauman and B. Leibe, *Visual Object Recognition*. San Rafael, CA: Morgan and Claypool, 2010.
- [17] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR)*, 2006, pp. 2169–2178.
- [18] E. Zhang and M. Mayo, “Enhanced spatial pyramid matching using log-polar-based image subdivision and representation,” in *Proc. Int. Conf. Digital Image Computing: Techniques and Applications (DICTA)*, 2010, pp. 208–213.
- [19] S. Zhang, Q. Tian, G. Hua, Q. Huang, and W. Gao, “Generating descriptive visual words and visual phrases for large-scale image applications,” *Trans. Imag. Processing*, vol. 20, no. 9, pp. 2664–2677, 2011.
- [20] H. Anwar, S. Zambanini, and M. Kampel, “Supporting ancient coin classification by image-based reverse side symbol recognition,” in *Proc. Int. Conf. Computer Analysis of Images and Patterns (CAIP)*, 2013, pp. 17–25.
- [21] R. Khan, C. Barat, D. Muselet, and C. Ducottet, “Spatial orientation of visual word pairs to improve bag-of-visual-words model,” in *Proc. British Machine Vision Conf. (BMVC)*, 2012, pp. 1–11.
- [22] S. Zambanini and M. Kampel, “Robust automatic segmentation of ancient coins,” in *Proc. Int. Conf. Computer Vision Theory and Applications (VISAPP)*, 2009, pp. 273–276.
- [23] H. Anwar, S. Zambanini, and M. Kampel, “Encoding spatial arrangements of visual words for rotation-invariant image classification,” in *Proc. German Conf. Pattern Recognition (GCPR)*, 2014, pp. 407–416.
- [24] H. Anwar, S. Zambanini, and M. Kampel, “A rotation-invariant bag of visual words model for symbol-based ancient coin classification,” in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, 2014, pp. 5257–5261.
- [25] J. C. Russ, *The Image Processing Handbook*, 5th ed. Boca Raton, FL: CRC Press, 2006.
- [26] Y. Tao and W. I. Grosky, “Spatial color indexing using rotation, translation, and scale invariant anglograms,” *Multimedia Tools Appl.*, vol. 15, no. 3, pp. 247–268, 2001.
- [27] M. d. Berg, O. Cheong, M. v. Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 3rd ed. New York: Springer-Verlag TELOS, 2008.
- [28] A. Kavelar, S. Zambanini, M. Kampel, K. Vondrovec, and K. Siegl, “The ILAC-project: Supporting ancient coin classification by means of image analysis,” in *Proc. ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2013, vol. XL-5/W2, pp. 373–378.
- [29] E. Nowak, F. Jurie, and B. Triggs, “Sampling strategies for bag-of-features image classification,” in *Proc. European Conf. Computer Vision (ECCV)*, 2006, pp. 490–503.
- [30] R. Arandjelović and A. Zisserman, “Smooth object retrieval using a bag of boundaries,” in *Proc. Int. Conf. Computer Vision (ICCV)*, 2011, pp. 375–382.



Rui Hu, Gulcan Can, Carlos Pallán Gayol, Guido Krempel, Jakub Spotak,
Gabrielle Vail, Stephane Marchand-Maillet, Jean-Marc Odobez, and Daniel Gatica-Perez

Multimedia Analysis and Access of Ancient Maya Epigraphy

© GHEENT, KATHEDRALE KERKFABRIEK; PHOTO COURTESY OF KIKI-IPA,
BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

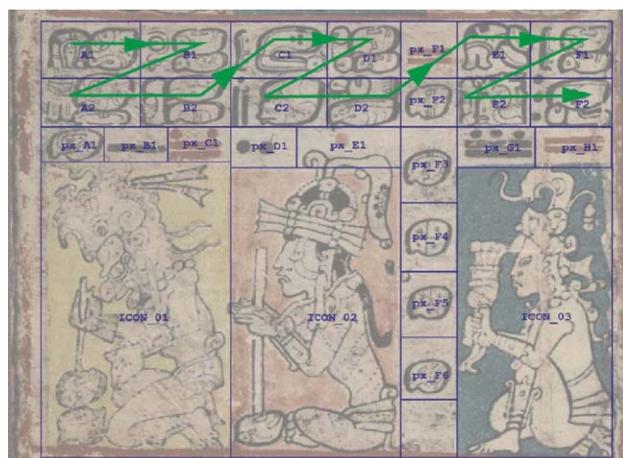
[Tools to support scholars on Maya hieroglyphics]

This article presents an integrated framework for multimedia access and analysis of ancient Maya epigraphic resources, which is developed as an interdisciplinary effort involving epigraphers (someone who deciphers ancient inscriptions) and computer scientists. Our work includes several contributions: a definition of consistent conventions to generate high-quality representations of Maya

hieroglyphs from the three most valuable ancient codices, which currently reside in European museums and institutions; a digital repository system for glyph annotation and management; as well as automatic glyph retrieval and classification methods. We study the combination of statistical Maya language models and shape representation within a hieroglyph retrieval system, the impact of applying language models extracted from different hieroglyphic resources on various data types, and the effect of shape representation choices for glyph classification. A novel Maya hieroglyph data set is given, which can be used for shape analysis benchmarks, and also to study the ancient Maya writing system.

Digital Object Identifier 10.1109/MSP.2015.2411291

Date of publication: 15 June 2015



[FIG1] A detailed template (page 6b) of the Dresden codex, showing individual constituting elements (glyph blocks, captions, calendric signs, and icons) framed by blue rectangles. Green arrows indicate the reading order of the blocks. (Image mapping by Carlos Pallán based on E. Förstemann 1880 and 1892 facsimiles.) High-resolution images of the whole codex are available in [4].

INTRODUCTION

The Maya culture is one of the major pre-Columbian civilizations that developed in ancient Mesoamerica. It began to flourish during the Preclassic period (2000 B.C. to A.D. 250), reached its peak during the Classic period (A.D. 250–900), and continued throughout the Post-Classic (A.D. 1000–1519) period. Ancient Mayan languages were recorded by means of a highly sophisticated system of hieroglyphic writing, comprising several thousand hieroglyphic signs, which has left us with an exceptionally rich artistic legacy.

The vast majority of ancient Maya texts were produced during the Classic Period, throughout which hieroglyphic texts were carved or painted on various media types, including stone monuments, architectural elements such as columns, lintels, capstones, and mural paintings, as well as personal items such as precious stones, ceramic vessels, bones, etc., to record diverse dedicatory, historical, astronomical, and mythological events. A rare type of Maya textual source is the so-called codex, which are screenfold books made of bark paper, coated with a film of lime plaster upon which textual and icon information was painted using a brush. Only three of these books are known to have survived the Spanish Conquest. Although their exact provenience and dating is not entirely known, they were, in all likelihood, produced within the greater Peninsula of Yucatan at some point during the Post-Classic period; see Figure 1 for an example.

Maya texts are typically composed of glyph blocks arranged in double columns. The most common reading order of glyph blocks is from left to right and from top to bottom within these double columns (see green arrows in Figure 1). One glyph block (also referred to as *block* in the rest of the article) could contain single or multiple glyphs. Due to the large variety of resources at their disposal, Maya scribes could choose among several writing conventions to render a particular Maya term. As a result, graphic conventions within a single block can vary greatly, it

could correlate with a phoneme (syllabic sign), an individual word (logograph), or even an entire sentence. Individual glyph recognition is a key and challenging step of Maya script decipherment.

Maya decipherment has undergone nearly 200 years of scholarly research [20]. While Maya archaeologists have discovered and documented a vast number of hieroglyphic monuments and their related context, epigraphers have achieved significant progress in deciphering the hieroglyphic script, and historical linguists have determined the languages recorded. Over 1,000 signs have been classified thus far by scholars in several catalogs. It has been estimated that approximately 80% of signs occurring in the known hieroglyphic corpus can be read phonetically [18].

Maya hieroglyphic analysis requires epigraphers to spend a significant amount of time browsing existing catalogs to identify individual glyphs from each block, as a necessary step for generating transcriptions, transliterations, and translations of Maya texts. Technological advances in automatic analysis of digital images and information management are allowing the possibility of analyzing, organizing, and visualizing hieroglyphic data in ways that could facilitate research aimed at advancing hieroglyphic analysis. However, there are several challenges for automatic Maya hieroglyphic data analysis. First, the available digitized Maya glyph data is limited. Second, the surviving Maya scripts have often lost their visual quality over time. Third, glyphs segmented from blocks are often partially missing due to occlusion. Finally, due to the artistic flexibility, glyphs of a same sign category may vary with time, location and styles; at the same time, glyphs of different categories may share similar visual features. In this article, we address automatic glyph recognition as image retrieval and classification problems.

Our work is a multidisciplinary effort, where computer scientists work closely with Maya epigraphers to design, develop, and evaluate computational tools that can robustly and effectively support the work of Maya hieroglyphic researchers. Our contributions include:

- novel Maya glyph data sets with unique historical and artistic value, which could potentially be used as shape benchmarks
- shape-based glyph retrieval and classification methods
- an in-depth study of the statistical Maya language model for automatic glyph retrieval
- a multimedia repository for data parsing and annotation.

RELATED WORK

Computer vision algorithms have shown potential to provide new insights into the realm of digital humanities. Various systems have been proposed to aid the analysis of cultural, historical, and artistic materials, which can significantly facilitate the daily work of scholars in the field.

The automatic analysis of historical manuscripts is the domain most related to our work. A large body of literature in this field examines the digitization and automatic analysis of cultural heritage data, produced from medieval times to the early 20th century [6], [13]. The methodologies include applying machine vision algorithms for page layout analysis, text line extraction, character recognition, and information retrieval. However, the

application of computer vision technologies for ancient manuscript analysis, such as Maya writing, is still a novel field. Previous work by our team contributed one of the first studies of visual Maya hieroglyph analysis and addressed glyph retrieval as a shape-matching problem [22], [23].

Shape-based visual information retrieval has been used for searching natural image data sets [10], [16], trademark images [19], technique drawings [14], three-dimensional objects [11], hand drawn images, or clip art [26]. Traditional shape-based retrieval systems include curve fitting [9], point-to-point matching [7], and grid-based matching [10]. These methods either do not scale well over large data sets, or only offer limited flexibility over shape variations. Recently, local shape descriptors [5], [16], [22] have been proposed and used in a bag-of-visual-words (BoW) framework for shape-based retrieval. Such methods can scale sub-linearly with appropriate search structures.

Automatic Maya glyph retrieval has been addressed in [22], where the histogram-of-orientation shape context (HOOSC) descriptor was developed. HOOSC combines the underlying formulation of the shape context (SC) [7] with the benefits that the histogram of oriented gradients (HOG) [8] provides. It was also applied in [5] for Egyptian hieroglyph analysis.

Previous studies have framed the Maya hieroglyph retrieval problem without considering any specific structure of Maya writing as a language. In contrast, language modeling has been widely used in machine translation and speech recognition. Glyph context information has recently been applied in [5] for Egyptian hieroglyph analysis with limited performance improvement. To the best of our knowledge, our previous work [17] was the first to incorporate glyph context information in Maya hieroglyph retrieval with significant accuracy improvement. However, the language model applied in [17] was at an incipient stage. It contained incomplete binary co-occurrence information of glyph pairs, extracted from the classic Thompson Maya hieroglyph catalog [25] dating from the 1960s. In this work, we extract a refined glyph co-occurrence model and test it on data sets of two different resources.

DATA SOURCES AND DIGITAL REPOSITORY

Two main data sources are considered in our work: the ancient Maya hieroglyphic books (codices) and monumental inscriptions. In this section, we first introduce the two data sources and then explain the novel data processing approach that we proposed to produce high-quality representation and annotation of ancient codical data.

DATA SOURCES

Given the inherent difficulties in the direct study and examination of the original hieroglyphic codex materials, the codical data sources comprise the existing primary documentation of the three extant ancient codices known as the Dresden, Madrid, and Paris codex, respectively. This documentation consists of reproductions, facsimiles, photographs, digital scans, and online resources as described in [2] and [15]. The Dresden codex is held in the state library in Dresden, Germany [4]; the Madrid codex is stored at the Museo de América in Madrid, Spain; and the Paris codex resides at

the Biblioteque Nationale de France [3]. While the exact provenience and dating of the Maya codices remains uncertain, most contemporary scholars consider that they were made within the northern Yucatan Peninsula during the late Post-Classic period.

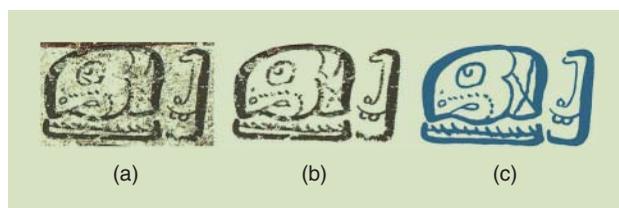
The monumental inscription data sources comprise a variety of carved stone monuments and architectural elements. Besides differences in media and format, it has distinctive historical characteristics as compared to the codical sources. Their dating falls several centuries earlier than the codices, and they stem from several parts of the Maya region, whereas the codices are restricted to the northern Yucatan Peninsula. Furthermore, monumental inscriptions number in the thousands as opposed to only three extant codices. Thus, the monumental sign repertoire is far better represented than the codical one.

DIGITAL MULTIMEDIA REPOSITORY

In this section, we introduce our data processing approach to generate high-quality digitization and annotation of the three Maya codices, which we refer to as the *digital multimedia repository (DMR)* of Maya hieroglyphic texts and icons. Ancient Maya scribes usually divided codex pages into smaller sections by red bands/lines; these sections are referred to by modern scholars as *t'ols*; each *t'ol* being further divided into frames relevant to the specific dates, texts, and imagery depicted. Frames contain glyph blocks (organized in a grid-like pattern), calendric glyphs, captions, and icons. Figure 1 shows an example *t'ol* (register) from the Dresden codex “segmented” into main elements. The DMR approaches the codices at different levels of detail:

- entire codex overview
- thematic sections
- almanacs
- *t'ols*
- frames
- individual elements (main text glyph blocks, calendric glyph blocks, captions, icons)
- individual signs or individual iconic elements.

With several software applications, we generate high-quality digitization from the raw image data. Specifically, we first conduct image enhancement, noise-reduction, and up-sizing of images to 400% of their original size; the enhanced text area is then cropped into glyph blocks [Figure 2(a)]; we generate clean-raster images from the cropped blocks by separating the cultural information (brushstrokes) from background noise and preservation accidents [Figure 2(b)]; we then generate high-quality



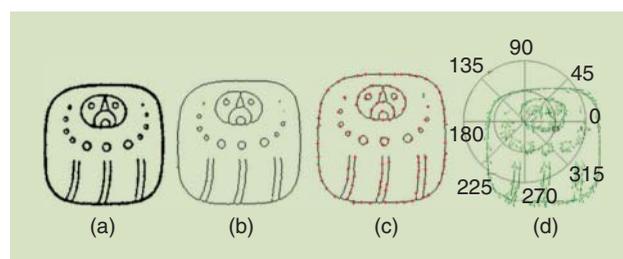
[FIG2] (a) A cropped glyph block. (b) A clean raster image. (c) A reconstructed vectorial representation. (Vectorization by Carlos Pallán.)

vectorial images in various formats, and by reconstructing the broken lines and missing strokes through thorough comparative analysis, we also generate reconstructed forms [Figure 2(c)]. Epigraphers require 15–30 minutes to generate a clean raster for a block, depending on complexity and preservation factors of the original data, and one to two hours more to further produce the high-quality reconstructed vectorial glyphs.

For the annotation of these visualizations, we developed an online server-like capture tool (relying on FileMaker Pro), allowing all partners real-time annotation and feedback capabilities. Among the several interconnected tables and templates of this tool, it is worth mentioning the glyph-concordance tool that we developed and that radically reduces the amount of time required to annotate glyph-strings within the codices under several different taxonomical systems. It provides automatic field translation and autocompletion functionalities for any individual sign (grapheme) or sign-string, maintaining maximum consistency between the annotations in four scholarly glyph catalogs [12], [20], [25], [27] that we have incorporated and cross-correlated. By enabling advanced multiple queries involving cross-referencing among the aforementioned catalogs, the system allows us to overcome the inherent limitations of having a language model based solely on the Thompson catalog. It also increases compatibility with one of our partner’s Web site [2], which contains annotations for a glyph-string of up to four symbols occurring within the codices and that was used to build a co-occurrence model (see the section “Glyph Co-Occurrence Model”). Finally, several tables and layouts of the DMR are currently being translated into computer science specifications to develop an advanced interface for data parsing and visualization.

GLYPH REPRESENTATION

Traditional shape descriptors [7], [21] may not be sufficient to capture the visual richness of Maya glyphs. We rely on the



[FIG3] (a) An original glyph; (b) thinning; (c) the pivot points (in red); and (d) HOOSC spatial quantization of a given pivot. (a) is used with permission from [25].

T0501	T0502	T0668	T0757	T0102	T0103
/b'al	/mal	/chal	/b'al	/kil	/tal

[FIG4] Thompson numbers, visual examples, and syllabic values of glyph pairs. Each pair contains two different signs of similar visual features. All images used with permission from [25].

HOOSC descriptor, which performed well for Maya hieroglyph analysis [17], [22]. We now describe the preprocessing and feature extraction steps that we follow.

Maya glyphs are often composed of strokes with different degrees of thickness. Thus, contour extractors sometimes generate “double” contours from the internal and external stroke lines, which can result in noisy descriptors. Therefore, we apply a thinning algorithm [23] to preprocess the binary shape [Figure 3(b)]. To reduce computation cost, we compute the local descriptor only at a set of points (called *pivots*) obtained through uniform sampling of the points along the thinned contour [see Figure 3(c)].

HOOSC was proposed in [22] as a robust shape representation for Maya hieroglyphs. It combines the strength of HOG [8] with circular split binning from the SC descriptor [7].

For each pivot point, the HOOSC is computed on a local circular space centered at the pivot’s location, partitioned into eight orientations and two rings as illustrated in Figure 3(d). The radius of the outer ring is the average pairwise distance between each pair of points along the contour; the inner ring covers half of this distance. An eight-bin HOG is calculated within each spatial region. This results in a richer representation than [7] and [21], where a simple counting of points or a sum of the unit gradient vectors of all points falling within a region is computed. The HOOSC descriptor for a given pivot point is then the concatenation of histograms of the 16 regions, which forms a 128-dimensional feature vector. As suggested in [22], we apply a per-ring normalization.

AUTOMATIC MAYA HIEROGLYPH RETRIEVAL

In this section, we present an automatic Maya glyph retrieval system, combining shape and glyph context information. Experimental results show the performance of our method.

SHAPE-BASED GLYPH RETRIEVAL

We adapt the BoW model for glyph retrieval, which has been established as a framework for scalable image retrieval [24]. Specifically, we apply *k*-means clustering on the set of HOOSC descriptors extracted from all glyphs in the database. The resulting *k* clusters are referred to as *visual words* and define the vocabulary of the system. A histogram representing the count of each visual word is then computed as a global descriptor for each glyph.

Given a query *G* and a database glyph *D*, represented by histograms H^G and H^D generated from the BoW system, we compute the city-block distance to measure the dissimilarity between *G* and *D*:

$$d(G, D) = \sum_{1 \leq i \leq k} |H^G(i) - H^D(i)|, \tag{1}$$

where each histogram is normalized so that $\sum_{1 \leq i \leq k} H(i) = 1$.

GLYPH CO-OCCURRENCE MODEL

Using shape alone to distinguish different glyphs is often problematic for many reasons. First, different signs often share



[FIG5] (a) Six glyph block examples from the codex data set; (b) four glyph strings segmented from the raster (black) and vectorial (blue) representation of the first two blocks shown individually in (a); (c) six glyph blocks from the monument data set; and (d) three glyph strings cropped from the first three blocks shown separately in (c).

similar visual features; see Figure 4. Furthermore, glyphs of the same sign category vary with time, location, and individual styles. For example, observe the last two blocks in Figure 5(a). Although visually different, the two blocks are actually composed of two same glyphs (T0668 and T0102; see Figure 4 for images of the two glyphs separately), pronounced “cha-ki” and representing the name of the rain god. Figure 6 shows six glyph pairs; each pair represents two glyphs of a same category but with different visual features. Finally, the surviving Maya scripts often lose their visual quality over time. We propose using glyph co-occurrence to complement visual information and help improve retrieval accuracy.

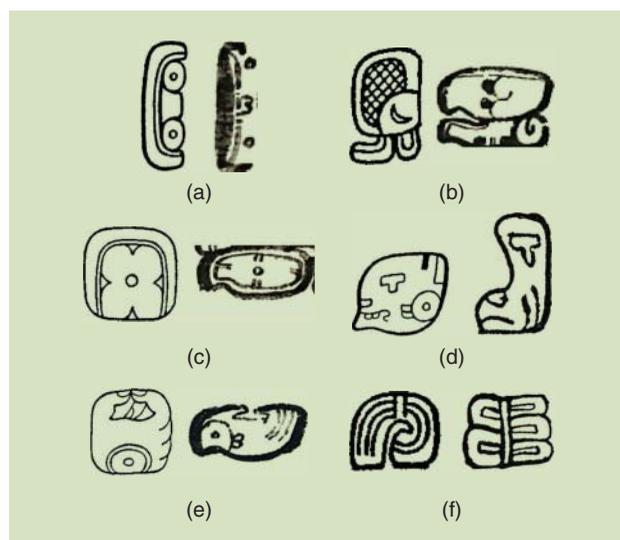
Maya glyph blocks were frequently composed of combinations of individual signs. Glyph co-occurrence within single blocks could therefore encode valuable information. While the reading order within a block usually follows the basic rule of left-to-right and top-to-bottom, several exceptions occur, particularly in the Madrid and Paris codices. Our methodology converts each glyph block into a linear string of individual signs, according to the reading order determined by our team’s epigraphers, as shown in Figure 5. We consider the first-order co-occurrence of neighboring glyphs as the glyph context information to build a statistical Maya language model.

Two glyph co-occurrence models (i.e., the Thompson co-occurrence model and the Vail co-occurrence model) are extracted from different sources. Their general form is

$$C(S_i, S_j) = \begin{cases} f_n & \text{if sign } S_i \text{ appears before sign } S_j, \\ \alpha & \text{otherwise,} \end{cases} \quad (2)$$

where f_n represents the normalized frequency that sign S_i appears before sign S_j , and $\alpha \in [0, 1]$ is a smoothing factor that accounts for missing co-occurring glyph pairs in the two models, which we explore in the experiments. Note that $C(S_i, S_j) \neq C(S_j, S_i)$.

From a computational point of view, the difference between the Thompson and the Vail models are just variations of the



[FIG6] Six pairs of glyph signs: (a) T0001, (b) T0158, (c) T0544, (d) T0668, (e) T0671, and (f) T0145. The left one of each pair is from the Thompson catalog, the right one is an example from the codex data set.

co-occurrence table. However, they are really different from an archeological point of view, as one features the Classic period monumental glyphs while the other features the Post-Classic codices. In the section “Experimental Results,” we test both of them on two different query sets to investigate the impact that this difference of writing conventions between the codex and the monument data has on the retrieval results.

THOMPSON CO-OCCURRENCE MODEL

The Thompson co-occurrence model is extracted from the classic Thompson hieroglyphic catalog [25], which covers 892 signs extracted from codices and monuments. Despite its outmoded taxonomy, it remains one of the most comprehensive and widely used lists of Maya glyphs. Thompson categorizes signs into affixes, main signs, and portraits (this categorization is no longer used in recent research on Maya writing). Affixes often co-occur with main signs, portraits, or affixes to form blocks. In his book, Thompson provides two glyph co-occurrence tables for affix signs, distinguishing whether they are used as prefix or postfix. However, no frequency information is given in these tables (we set $f_n = 1$ for valid co-occurrence), and co-occurrence between main signs and portraits is not listed. There are 4,574 glyph pairs included in this model, which correspond to a sparsity of 99.4%.

VAIL CO-OCCURRENCE MODEL

This co-occurrence model is extracted from the online Maya hieroglyphic database [2] containing a state-of-the-art description and interpretation of the three surviving Maya codices. Using the database, we extract the co-occurrence information of neighboring glyphs within single blocks. The resulting model contains a complete set of co-occurring pairs that are known today from the three surviving codices, as well as their occurrence frequency. In total, the database contains 336 different glyph signs. There are 1,818 glyph pairs, which co-occur with frequencies f from one to 188, corresponding to a sparsity of 99.8% (since we consider 892 signs from the Thompson catalog). More than half of these pairs only co-occur once. Approximately 93% of the pairs appear fewer than ten times. We normalize f with the following function:

$$f_n = 1 + \log_{10}(f). \quad (3)$$

THOMPSON AND VAIL CO-OCCURRENCE MODEL

Additionally, we build a third model by considering all co-occurrence pairs from the two former models. We disregard the frequency information of the Vail model and generate a binary model, which we refer to as the *Thompson and Vail* co-occurrence. It contains 5,600 co-occurred glyph pairs, which correspond to a sparsity of 99.3%. We expect this model to perform better across different Maya data sets (e.g., codices and monument).

INCORPORATING CONTEXT INFORMATION

We now explain how to incorporate this information in the shape-based retrieval system.

SEQUENCE MODEL

Denote by $G_{1:n} = [G_1, \dots, G_i, \dots, G_n]$ the observed glyph string, and by $S_{1:n}$ the sequence of recognized states, where S_i indicates the sign category annotated for glyph G_i . Considering the glyph string $G_{1:n}$ as a first-order Markov chain, the probability of labeling it to a sequence of states $S_{1:n}$ is

$$P(S_{1:n} | G_{1:n}) \propto P(G_1 | S_1) \prod_{2 \leq i \leq n} (P(G_i | S_i) P(S_i | S_{i-1})), \quad (4)$$

where $P(S_i | S_{i-1})$ denotes the transition probability. Here we directly use $C(S_{i-1}, S_i)$ to approximate this probability. $P(G_i | S_i)$ refers to the likelihood of glyph G_i being labeled as sign S_i . To encode this term we use the visual similarity between G_i and the glyph example of S_i in the database, and define $P(G_i | S_i) \propto e^{-d(G_i, S_i)/\lambda}$, where $d(G_i, S_i)$ is computed using (1), and λ is a scale factor empirically set to the average distance of the top 50 ranked results for all queries.

RETRIEVAL SYSTEM

When only shape information is considered, the score of a query glyph G_i being labeled by sign D , is computed by their shape likelihood

$$\text{Score}^{\text{sh}}(S_i = D) \propto P(G_i | S_i = D). \quad (5)$$

In our model, we propose ranking the glyphs according to

$$\text{Score}^{\text{sh+context}}(S_i = D) = \max_{S_{1:i-1}, S_{i+1:n}} P(S_{1:n} | G_{1:n}), \quad (6)$$

which means the following: given $S_i = D$, find the sequence of labels $S_{1:n}$ that provides the maximum probability to label $G_{1:n}$, under the model in (4) and use this probability as score to rank the database glyphs. This can be efficiently computed using the Viterbi algorithm. Thus, the score of the glyph G_i being recognized as $S_i = D$ now takes into account all observed glyphs in the string, with the effect that a glyph D that normally co-occurs with glyphs that are visually likely at neighboring positions will receive a higher weight.

EXPERIMENTAL RESULTS

We present the data sets, experimental setting, and retrieval results.

GLYPH DATA SETS

Two data sets, the codex and monument data sets, were used as query sets to retrieve from a common database.

Codex Data Set

This data set was produced by epigraphers in our team (see the section “Data Sources and Digital Repository”) and is available for download from [1]. It contains glyph blocks from the three surviving Maya codices, along with their segmented individual glyphs and corresponding annotations; see Figure 5(a) and (b), for example. To study the impact of visual data quality on the retrieval

performance, we considered two subdata sets. The first one is called a *codex small* data set, composed of 151 glyphs segmented from 60 blocks, for which we have both the vectorial and raster representations. Remember that producing high-quality vectorial representations (including reconstruction of missing parts) is time-consuming compared to raster images, which reflect the actual visual content but are affected by degradations. The second subset is intended to assess the validity of the reported performance by using a larger corpus (termed *codex large*) comprising only the raster representation of 587 glyphs from 224 blocks.

Monument Data Set

The monument data set contains 127 glyphs of 40 blocks extracted from stone monument data and is derived from a quite different data source than the codex data in terms of Maya historical period, media type, and glyph generation process. Samples are shown in Figure 5(c) and (d). The data consisted of line drawings of glyphs manually traced on top of multiple layers of enhanced photographs taken at sites at night under raking-light illumination to bring out different levels of detail. Given each block, we manually drew a tight bounding box around individual glyphs to extract query glyphs. The queries may be affected by adverse effects, like background noise, additional strokes from neighboring glyphs, or partially missing strokes due to glyphs overlapping within blocks.

Retrieval Database

We scanned and segmented all the glyphs from the Thompson catalog [25] to form the retrieval database. In total, it contains 1,487 glyph images belonging to 892 different categories. Thus, a category is usually represented by a single glyph image, and sometimes by multiple ones, each representing a different visual instance of the glyph category. Figure 4 shows glyph images of six different categories.

EXPERIMENTAL SETTING

For each glyph query, we extract the rank of the true glyph in the retrieval results and use the average of these ranks over all queries as performance measure (the lower the average ranking, the higher the accuracy).

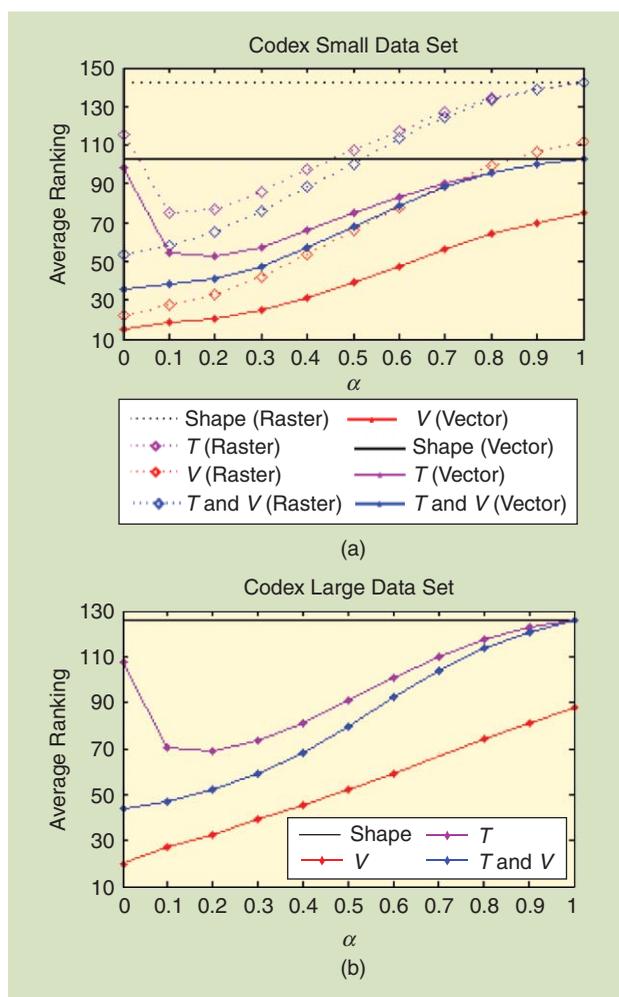
We studied the impact of several factors on the performance, including the vocabulary size of the BoW representation, the smoothing factor α used to build the co-occurrence models [see (2)], and the co-occurrence models.

RESULTS AND DISCUSSION

Results are presented in Figures 7 and 8, in which “T,” “V,” and “T&V” refer to the “Thompson,” “Vail,” and “Thompson and Vail” co-occurrence models, respectively.

Shape-Based Glyph Retrieval

Looking first at the impact of data origin and quality on glyph retrieval performance when only shape information is considered, the following observations can be made. First, as expected, higher quality vectorial representations result in higher retrieval accuracy [103 for vectorial versus 142 for raster images; see the plain and

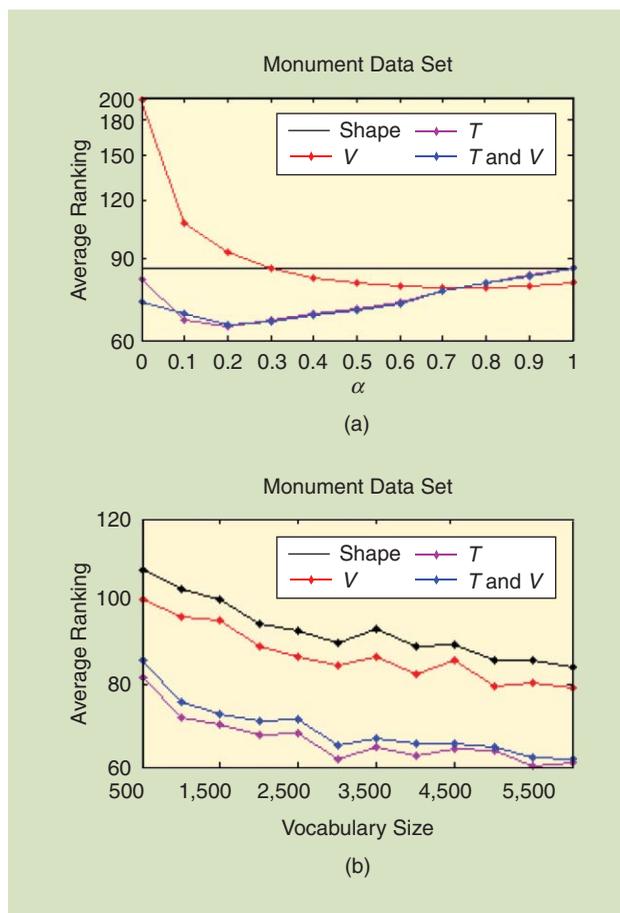


[FIG7] The average ranking on the codex data sets with varying smoothing factor α [see (2)] and fixed vocabulary size 5,000: (a) “codex small” and (b) “codex large” data sets.

dotted horizontal lines in Figure 7(a)]. Second, by comparing the shape-based retrieval results of the monument data [see the horizontal line in Figure 8(a)] and the codex data (see the plain and dotted horizontal lines in Figure 7), we can see that, despite the presence of distracting elements (background noise, line strokes from neighboring glyphs, etc.), the shape retrieval accuracy on the monument data (86) is higher than on the codex data. This reflects the higher visual similarity between the glyphs in the monument data set and those from the retrieval database. As glyphs in the retrieval database are extracted from the Thompson catalog, which largely relies on monument data, as compared to signs from the codex data, which often exhibit more variability as illustrated in Figure 6. For the same reason, using finer HOOSC descriptor quantization (i.e., using a larger BoW vocabulary) consistently improved the results on the monument data [Figure 8(b)], whereas it had no impact on the codex data (curves not shown).

Incorporating Context Information

As can be seen from Figures 7 and 8, the average retrieval rankings obtained using different co-occurrence models and smoothing

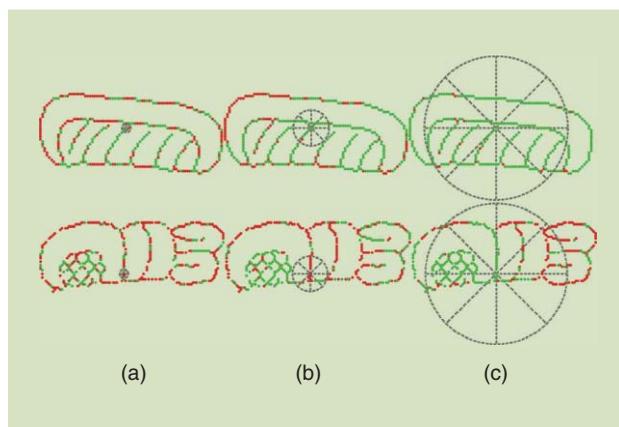


[FIG8] The average ranking on the monument data set (a) with varying smoothing factor α [in (2)] and fixed vocabulary size 5,000; (b) with varying vocabulary size and fixed $\alpha = 0.2$ for the "T" and "T&V" models and $\alpha = 1$ for the "V" model.

factors are usually significantly lower than when using only shape. For instance, on the small codex data set, the Vail model (with $\alpha = 0$) can reduce by as much as 130 and 90 the average ranking for the raster and vectorial representation, respectively, whereas on the monument data [Figure 8(a)], the gain is smaller (around 20 for the Thompson model with $\alpha = 0.2$).

Differences across models as well as the optimal smoothing factors mainly reflect the agreement between the source of the block queries and the data used to build the co-occurrence models. Thus, on one hand, the Vail model achieves the best accuracy on codex data sets (Figure 7), but underperforms on the monument data (Figure 8), even degrading the shape-only results for low smoothing factors. Since this model is purely built from the codices, this may imply that the Maya writing on codices and monuments follows different glyph co-occurrence conventions. On the other hand, the Thompson model, built from a mixed source of monument and codex data, offers a much smaller gain when applied to the codex data, but still performs well on monument data. Altogether, these two models are advantageously combined in the more versatile "T&V" model.

No smoothing factor ($\alpha = 0$) is needed when applying the Vail model on codex data, since it covers all known co-occurrence



[FIG9] Classified pivots using HOOSC one-ring with position and spatial context (a) 1/16, (b) 1/4, and (c) 1, respectively. Green (respectively, red) points indicate correctly (respectively, incorrectly) classified pivots.

instances of the codices; whereas the Thompson model that relies only on incomplete data sources misses some co-occurrence pairs and thus requires a smoothing factor (typically $\alpha = 0.2$). In general, all the aforementioned remarks remain valid when considering the large codex data [Figure 7(b)].

As a final remark, one can notice in Figure 7(a) that the retrieval performance differences between the vectorial and raster representation becomes less important when using a co-occurrence model. In this context, the raster representation can be used as a compromise between data production efficiency and retrieval accuracy.

SHAPE-BASED GLYPH CLASSIFICATION

There are use cases in which inferring the correct category of a glyph is important. In this section, we study the single-glyph classification task, first presenting the classification methods, and then discussing the obtained results.

GLYPH CLASSIFICATION METHODS

The objective is to build a classifier that categorizes a test shape into one of the N_G categories. As a baseline, we use the method of [23], where glyphs are represented using the global BoW representation. A test glyph gets the class label of its nearest neighbor [using the BoW city-block distance in (1)] in the training set.

As an alternative, we propose a method that categorizes an unknown glyph by first identifying the category of its individual local pivot points. Specifically, for a given glyph, we first compute the HOOSC descriptor at each pivot point and classify it using a K -nearest neighbor method. In a second step we classify the glyph as the category that receives the largest number of votes from the individual pivots.

EXPERIMENTAL RESULTS

DATA SET

We used a subset of glyphs from the monumental inscriptions that were used in [23]. We only consider glyph categories that contain more than 30 glyphs. The resulting data set is

composed of ten glyph categories with 25 training images per class and 125 test images in total. The groundtruth of the glyph category is provided by our team scholars.

EXPERIMENTAL SETTING

We used 300 equidistant pivots where we compute the HOOSC descriptor. Note that here, we extracted the orientation from the raw images preprocessed by a continuous Gaussian orientation filter, as this produced more stable results than applying the thinning preprocessing. We considered three settings to compute the HOOSC descriptor:

- two rings and eight radial bins; see Figure 3(d);
- one ring and eight radial bins; see Figure 9
- case (2) with added position information, i.e., the HOOSC descriptor is augmented with the relative position (defined within $[0, 1] \times [0, 1]$) of the pivot point within the glyph bounding box.

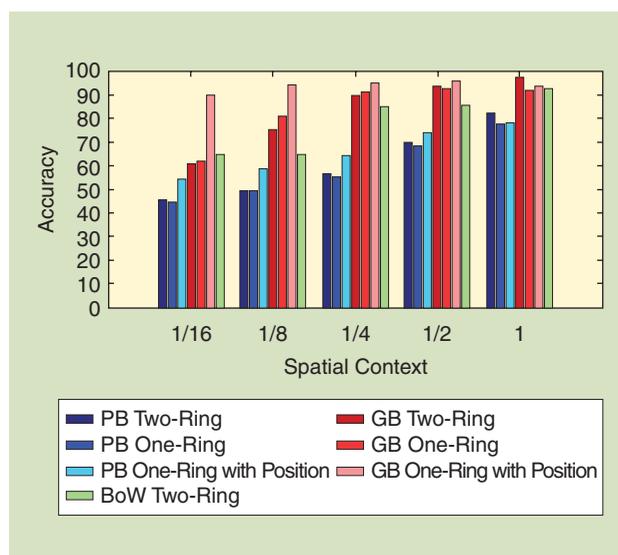
Furthermore, for each of the three settings, we considered five spatial contexts (defined by the radius of the outer ring in HOOSC computation): 1/16, 1/8, 1/4, 1/2, and 1, all defined as a proportion to the mean of the pairwise distance between pivot points (see the gray circles in Figure 9), as we are interested in studying the impact of the spatial scope used to compute the HOOSC descriptor on the classification performance. Indeed, while large scopes used in previous works (and our retrieval experiments in the “Retrieval System” section) led to good results when dealing with clean glyph inputs, there are situations where smaller scopes would be useful, e.g., when dealing with damaged glyph shapes (the damage will affect most of the descriptors when using a large scope), or if we wanted to identify which local part of the glyph is a “diagnostic” feature, i.e., a discriminant visual element that scholars rely on to distinguish a glyph.

RESULTS AND DISCUSSION

Figure 10 shows the classification results obtained using the BoW method and the proposed method (“glyph-based” results, denoted *GB*) for different spatial context sizes and partition settings. To better understand the proposed method, we also show the “pivot-based” (denoted *PB*) classification accuracy, i.e., the percentage of pivot points whose descriptor is correctly classified as the category of its associated glyph.

First, from the results of the “pivot-based” method (blue bars), we can notice that the performance degrades almost linearly as the spatial context decreases, but remains well above chance level (10%) even for a small spatial extent (1/16). Interestingly, as this context gets smaller, the incorporation of the spatial position (PB one-ring with position) allows to boost performance by 10% as compared to the case without position (PB one-ring). Furthermore, while two rings are useful as the spatial context is large, it is not superior than one ring in terms of PB performance and actually degrades the GB performance when smaller spatial context is considered (e.g., 1/4 to 1/16).

Second, the performance with respect to spatial context at the glyph level (red bars) does not decrease as dramatically as at the pivot level, indicating that misclassified points, even if they



[FIG10] Classification accuracy of the BoW method (green bar), and the proposed method at the pivot level [pivot-based (PB) results (blue bars)] and the glyph level [glyph-based (GB) results (red bars)] with various settings to compute the HOOSC descriptor.

dominate, usually get distributed over all other classes rather than a single one. Hence, the pivots predicted with true labels may win in the voting phase. For GB one-ring with position, the classification remains as high as 94% with a spatial context of 1/8. Note that this is not the case in the BoW approach (green bars), whose performance degrades as the spatial context decreases, performing worse than the proposed approach with spatial radius larger than 1/4, and cannot keep up with the one-ring with position results at smaller spatial scopes.

Figure 9 illustrates the pivot classification results for two glyphs over three spatial context levels. We can see that the number of pivots classified correctly increases with the spatial context. It also shows that while some local structures are recognized at most scales (diagonal lines for the top glyph, hatches for the bottom one), there are structures that still remain confusing among glyph classes, even at the larger contexts (e.g., the pivots near the “ears” in the bottom glyph).

We can conclude that a two-step approach where class-information is used to categorize the descriptor (rather than simple quantization in BoW) brings more robustness as the spatial context decreases (and may bring even more robustness when dealing with partially damaged glyphs), and that incorporating the relative position of pivots is important, as the same local shape structure might be observed at different positions for different glyph categories.

CONCLUSIONS AND FUTURE WORK

This article presented an approach to capture and produce high-quality multimedia data from valuable historical codex data sources, upon which we propose algorithms for automatic Maya hieroglyph analysis.

We defined consistent conventions to generate high-quality representations of the ancient Maya hieroglyphs, as well as a data

model that not only provides a convenient platform for epigraphers to annotate and analyze data, but also serves as a bridge between epigraphers and computer scientists for data parsing and analysis. A novel codex data set is contributed under the proposed system.

We then addressed two automatic glyph analysis tasks with value to support epigraphers' daily work, namely glyph retrieval and glyph classification. Regarding retrieval, two Maya language models were extracted from different data sources and incorporated into a shape-based automatic glyph retrieval framework. Our study showed that glyph co-occurrence encode valuable information of the Maya writing system, which can be used to complement the visual automatic analysis. The retrieval results also showed that the Maya writing on codices and monuments follows different glyph co-occurrence conventions. Finally, we studied the effect of shape representation choices in the classification task.

Our future work includes automatic Maya text area detection, as well as detection and segmentation of blocks and glyphs, which will facilitate the daily work of scholars when more data becomes available. In another direction, we are working on designing a visualization interface to allow manipulation of Maya data in a systematic and flexible way. Data will be displayed as clusters in various feature spaces (from low-level visual features to high-level semantic spaces); analyzed with different levels of context information (within block co-occurrence, surrounding text, icons); and visualized in various resolutions and positions by zooming and panning. We expect the traditional Maya hieroglyph decipherment to benefit from such functionalities.

ACKNOWLEDGMENTS

We thank the Swiss National Science Foundation and the German Research Foundation for their support, through the Multimedia Analysis of Ancient Yucatan Annotations project.

AUTHORS

Rui Hu (rhu@idiap.ch) is a postdoctoral researcher at the Idiap Research Institute in Switzerland.

Gulcan Can (gcan@idiap.ch) is a Ph.D. candidate at the Idiap Research Institute and EPFL in Switzerland.

Carlos Pallán Gayol (s5capall@uni-bonn.de) is a research associate in the Department of Anthropology of the Americas at the University of Bonn, Germany.

Guido Krempel (guidokrempel@gmail.com) is a research associate in the Department of Anthropology of the Americas at the University of Bonn, Germany.

Jakub Spotak (spotak.jakub@gmail.com) is a Ph.D. candidate in the Faculty of Arts at Comenius University in Bratislava, Slovakia.

Gabrielle Vail (gvail@ncf.edu) is a research scholar in the Division of Social Sciences, New College of Florida in Sarasota.

Stephane Marchand-Maillet (stephane.marchand-maillet@unige.ch) is an associate professor at the University of Geneva, Switzerland.

Jean-Marc Odobez (odobez@idiap.ch) is the head of the Perception and Activity Understanding Group at the Idiap Research Institute, Switzerland, and Maitre d'Enseignement et de Recherche at EPFL, Switzerland.

Daniel Gatica-Perez (gatica@idiap.ch) is the head of the Social Computing Group at the Idiap Research Institute, Switzerland, and Professeur Titulaire at EPFL, Switzerland.

REFERENCES

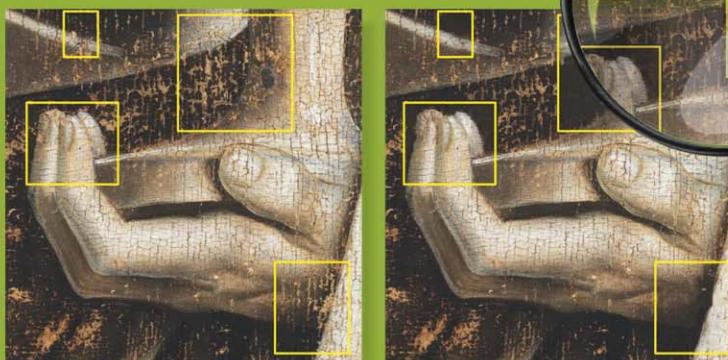
- [1] Codex dataset. [Online]. Available: www.idiap.ch/dataset/maya-codex
- [2] The Maya hieroglyphic codices. [Online]. Available: <http://www.mayacodices.org/>
- [3] Paris Codex. [Online]. Available: http://gallica.bnf.fr/ark:/12148/btv1b8446947j/f1.zoom.r=Codex_anus.langDE
- [4] Saxon State and University Library Dresden (SLUB) library. [Online]. Available: <http://digital.slub-dresden.de/werkansicht/dlf/2967/1/>
- [5] M. Franken and J. v. Gemert, "Automatic Egyptian hieroglyph recognition by retrieving images as texts," in *Proc. ACM Multimedia*, 2013, pp. 765–768.
- [6] N. Arvanitopoulos and S. Süssstrunk, "Seam carving for text line extraction on color and grayscale historical manuscripts," in *Proc. Int. Conf. Frontiers in Handwriting Recognition*, 2014, pp. 726–731.
- [7] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," in *Proc. Pattern Anal. Machine Intell.*, 2002, pp. 509–522.
- [8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. Comput. Vision Pattern Recognition*, 2005, pp. 886–893.
- [9] A. Del Bimbo and P. Pala, "Visual image retrieval by elastic matching of user sketches," in *Proc. Pattern Anal. Machine Intell.*, 1997, pp. 121–132.
- [10] M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa, "A descriptor for large scale image retrieval based on sketched feature lines," in *Proc. 6th Eurographics Symp. Sketch-Based Interfaces and Modeling*, 2009, pp. 29–36.
- [11] M. Eitz, R. Richter, T. Boubekeur, K. Hildebrand, and M. Alexa, "Sketch-based shape retrieval," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 31:1–31:10, 2012.
- [12] E. B. Evrenov, Y. Kosarev, and B. A. Ustinov, *The Application of Electronic Computers in Research of the Ancient Maya Writing*. USSR: Novosibirsk, 1961.
- [13] A. Fischer, H. Bunke, N. Naji, J. Savoy, M. Baechler, and R. Ingold, "The His-Doc project: Automatic analysis, recognition, and retrieval of handwritten historical documents for digital libraries," in *Proc. Int. Interdisciplinary Aspects of Scholarly Editing*, 2012, pp. 81–96.
- [14] M. J. Fonseca and J. A. Jorge, "Towards content-based retrieval of technical drawings through high-dimensional indexing," in *Proc. 1st Ibero-American Symp. Computer Graphics*, 2002, pp. 263–270.
- [15] D. Gatica-Perez, C. Pallán-Gayol, S. Marchand-Maillet, J. M. Odobez, E. Roman-Rangel, G. Krempel, and N. Grube, "The MAAYA project: Multimedia analysis and access for documentation and decipherment of Maya epigraphy," in *Proc. Digital Humanities Conf.*, 2014.
- [16] R. Hu and J. P. Collomosse, "A performance evaluation of gradient field hog descriptor for sketch based image retrieval," in *Proc. Computer Vision Image Understanding*, 2013, pp. 790–806.
- [17] R. Hu, C. Pallán-Gayol, G. Krempel, J. M. Odobez, and D. Gatica-Perez, "Automatic Maya hieroglyph retrieval using shape and context information," in *Proc. ACM Multimedia*, 2014, pp. 1037–1040.
- [18] H. Kettunen and C. Helmke, *Introduction to Maya Hieroglyphs—XVII European Maya Conf.*, Helsinki Univ., Helsinki, Finland, 2012.
- [19] W. H. Leung and T. Chen, "Trademark retrieval using contour-skeleton stroke classification," in *Proc. Int. Conf. Multimedia Expo*, 2002, pp. 517–520.
- [20] M. J. Macri and G. Vail, *The New Catalog of Maya Hieroglyphs (Vol. II, The Codical Texts)*. Univ. of Oklahoma Press, Norman, OK, 2009.
- [21] G. Mori, S. J. Belongie, and J. Malik, "Efficient shape matching using shape contexts," in *Proc. Pattern Anal. Machine Intell.*, vol. 27, no. 11, pp. 1832–1837, 2005.
- [22] E. Roman-Rangel, C. Pallán-Gayol, J. M. Odobez, and D. Gatica-Perez, "Analyzing ancient Maya glyph collections with contextual shape descriptors," *Int. J. Comput. Vision*, pp. 101–117, 2011.
- [23] E. Roman-Rangel, C. Pallán-Gayol, J. M. Odobez, and D. Gatica-Perez, "Searching the past: An improved shape descriptor to retrieve Maya hieroglyphs," in *Proc. ACM Multimedia*, 2011, pp. 163–172.
- [24] J. Sivic and A. Zisserman, "Video Google: A text retrieval approach to object matching in videos," in *Proc. Int. Conf. Comput. Vision*, 2003, pp. 1470–1477.
- [25] J. Eric and S. Thompson, *A Catalog of Maya Hieroglyphs*. Univ. of Oklahoma Press, Norman, OK, 1962.
- [26] C. Wang, J. Zhang, B. Yang, and L. Zhang, "Sketch2Cartoon: Composing cartoon images by sketching," in *Proc. ACM Multimedia*, 2011, pp. 789–790.
- [27] G. Zimmermann, *Die Hieroglyphen der Maya Handschriften. Abhandlungen aus dem Gebiet der Auslandskunde*. Band 62-Reihe B. Universität Hamburg Cram, De Gruyter & Co., 1956.



[Ramya Srinivasan, Conrad Rudolph, and Amit Roy-Chowdhury]

Computerized Face Recognition in Renaissance Portrait Art

© GHEENT, KATHEDRALE KERKFABIEK; PHOTO COURTESY OF KIK-IRPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

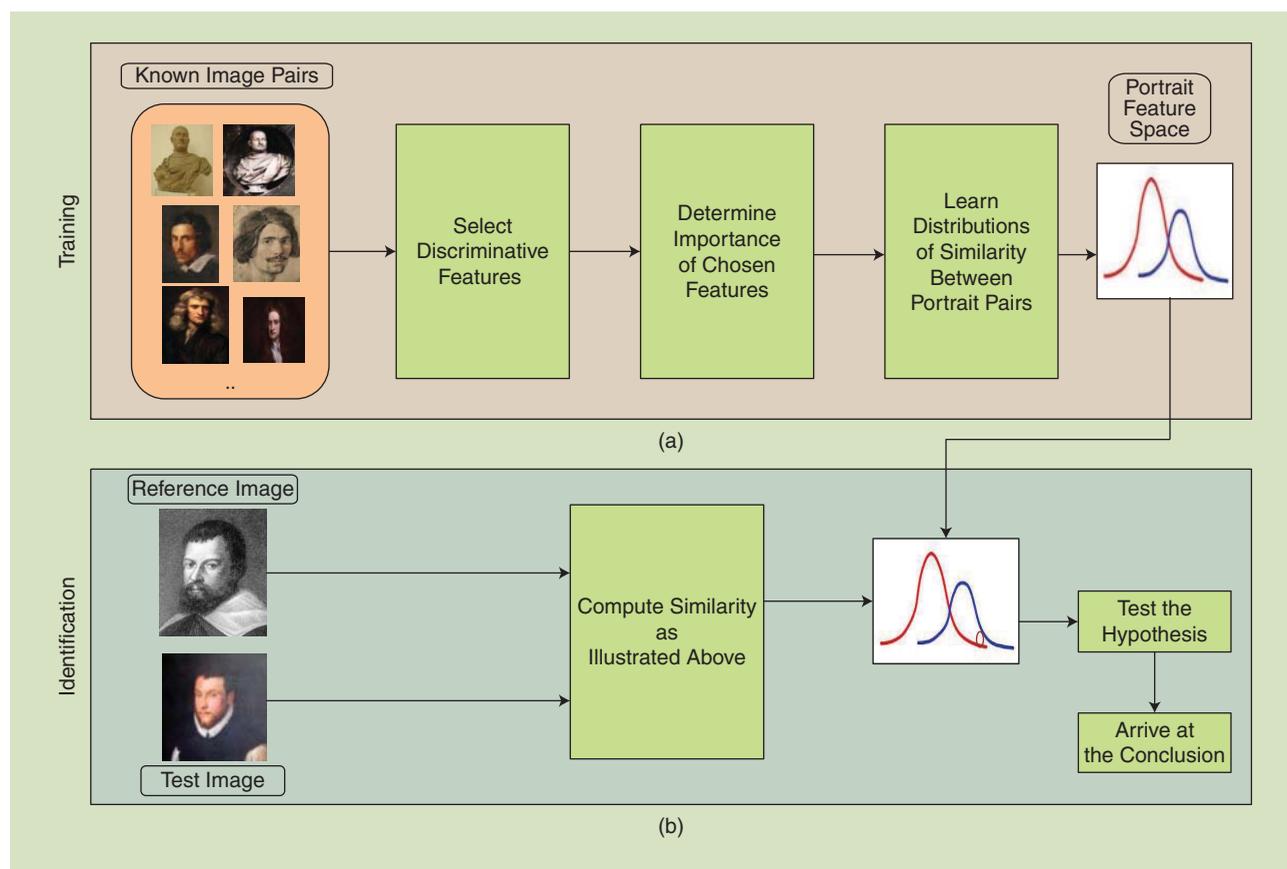
[A quantitative measure for identifying uncertain subjects in ancient portraits]

This article explores the feasibility of face-recognition technologies for analyzing works of portraiture and, in the process, provides a quantitative source of evidence to art historians in answering many of their ambiguities concerning identity of the subject in some portraits and in understanding artists' styles. Works of portrait art bear the mark of visual interpretation of the artist.

Digital Object Identifier 10.1109/MSP.2015.2410783

Date of publication: 15 June 2015

Moreover, the number of samples available to model these effects is often limited. Based on an understanding of artistic conventions, we show how to learn and validate features that are robust in distinguishing subjects in portraits (sitters) and that are also capable of characterizing an individual artist's style. This can be used to learn a feature space called *portrait feature space (PFS)* that is representative of quantitative measures of similarities between portrait pairs known to represent same/different sitters. Through statistical hypothesis tests, we analyze uncertain portraits against known identities and explain the significance of the results from an art historian's perspective. Results are shown



[FIG1] An illustration of the (a) training and (b) identification framework.

on our data consisting of over 270 portraits belonging largely to the Renaissance era.

INTRODUCTION

Renaissance portraits were depictions of various important people of those times. These encompass a wide range of art works such as sculptures, death masks, mosaics, etc. Apart from being used for a variety of dynastic and commemorative purposes, they were used to depict individuals often to convey an aura of power, beauty, or other abstract qualities [1]. A large number of these portraits, however, have lost the identities of their subjects through the fortunes of time.

The analysis of faces in portraits can offer significant insights into the personality, social standing, etc. of the subject they represent. However, this is not a simple task, since a portrait can be “subject to social and artistic conventions that construct the sitter as a type of their time” [1], thus resulting in large uncertainty regarding the identity of many of these portraits. Traditionally, identification of many of these portraits has been limited to personal opinion, which is often quite variable. The project Faces, Art, and Computerized Evaluation Systems (FACES) was conceptualized to evaluate the application of face recognition technology to portrait art and, in turn, aid art historians by providing a quantitative source of evidence to help answer questions regarding subject identity and artists’ styles. This article will describe the challenges

inherent in face recognition in art images, and summarize the results obtained in this project over the last two years. Some preliminary results have been presented in [12].

There have been lingering ambiguities about the identity in some portraits—henceforth referred to as *test images*. The question has been whether they might represent a certain known identity, which we call *reference images*. As an example, the test image in Figure 1 is a portrait painted perhaps around 1590, and is believed by some to represent Galileo. Through computerized face recognition technologies, we try to provide an alternate and quantitative source of evidence to art historians in answering such questions.

We leverage upon a number of portrait pairs that are known to represent a certain person as shown in Figure 1(a). The task then is to train the computer in identifying discriminative features that can not only distinguish one sitter from another, but also learn the importance of the chosen features depending on the amount of emphasis given to that feature by an artist. Using the learned features, quantitative measures of similarity between portrait pairs known to represent the same person can be computed to yield what we call *match scores*. Analogously, similarity scores between portrait pairs not known to represent the same person yield *non-match scores*. The resulting match (blue curve) and nonmatch scores (red curve) together constitute what we refer to as the PFS. Subsequently, using hypothesis tests, the similarity score between test and reference image, shown by the brown ball amid

the similarity score distributions in Figure 1(b), is analyzed with respect to the learned PFS to arrive at appropriate conclusions of a possible match or nonmatch. If both match or nonmatch happen to be likely, then no decision can be made.

We begin by describing the challenges involved in the face recognition of portraits. Apart from the typical challenges associated with face recognition systems such as variations in pose, expression, and illumination, face recognition in portraits comes with additional challenges. Some of these are described below.

- *Modeling artists' styles:* Since portraits bear the mark of the visual interpretation of an artist, styles of individual artists characterizing their aesthetic sensibilities (often biased by their sociocultural backgrounds) have to be modeled. Thus, portraits of the same sitter can vary from artist to artist. This results in considerable variability in the renditions, which has to be accounted for by the face recognition algorithms.
- *Lack of sufficient training data:* Many existing feature selection methods rely on the availability of a significant amount of training data. This is rarely the case in our problem domain due to the following reasons: 1) lack of a significant body of images, the authenticity of which is well established; and 2) we need to logically choose a set of related images directed toward a particular demonstrative end and adhering to a particular period style.
- *Choice of features:* Given the aforementioned constraints, we need to choose features that best justify an artist's rendition and possess high discriminative power in distinguishing the sitter from others. Although there has been some preliminary work on this [2], there is little to no elaborate work on understanding how to model style in face portraiture. This leads to interesting questions in machine learning on combinations of various algorithms that are pertinent here.

RELATED WORK

We review some image processing techniques employed for art analysis and also provide a survey of the state of the art in computerized face recognition.

IMAGE ANALYSIS IN ARTWORKS

The analysis of paintings using sophisticated computer vision tools has gained popularity in recent years [5]. Computer analysis has been used for identifying the artist [25] and for studying the effect of lighting in artworks [26], among others. A recent paper has explored application of computer-based facial-image analysis [6] using three-dimensional (3-D) shape information to identify one subject, Leonardo da Vinci, in four artworks. The present work involves multiple sitters (both genders) by different artists portrayed across different media such as paintings, death masks,

THE PROJECT FACES, ART, AND COMPUTERIZED EVALUATION SYSTEMS (FACES) WAS CONCEPTUALIZED TO EVALUATE THE APPLICATION OF FACE RECOGNITION TECHNOLOGY TO PORTRAIT ART AND, IN TURN, AID ART HISTORIANS BY PROVIDING A QUANTITATIVE SOURCE OF EVIDENCE REGARDING SUBJECT IDENTITY AND ARTISTS' STYLES.

etc. Some preliminary results have been presented in our earlier paper [12] on a small set of data. In this article, extensive results are shown on a much richer data set and using a more sophisticated feature-extraction algorithm. Also, for the present analysis, shape information was found to be less discriminative when compared to other features such as anthropometric distances (ADs) and local features (LFs). This can be partly attributed to the evidence that artists often focused on LFs and took some liberties with shape [13].

COMPUTERIZED FACE RECOGNITION

A survey of still- and video-based face recognition research is provided in [3]. A vast majority of face recognition applications address surveillance and entertainment. These approaches can be classified into three categories: holistic methods, feature-based structural matching methods, or a combination of both depending on the representation in feature space. Three-dimensional modeling approaches such as [24] have also been studied. Recent research efforts have focused on cross spectral face recognition for comparing images taken in heterogeneous environments [8]. Such methods are not applicable for our study. First, since the images in the present scenario are obtained from museums across the world, we have no control on the kind of sensors used to capture them. Second, the quality of the image is not an issue here; the challenge is choice of appropriate features.

Some works [9], [28] model style factors such as a facial pose, and expression and separate it from content, i.e., the identity of the person, and show promising results for face pose estimation, among others. In [11], the authors use attributes like “chubby,” “attractive,” etc. for face verification tasks. While models for separating style (e.g., an artist's rendition) from content (sitter's identity) can be useful for the present study, all of the existing methods require hundreds of images. Some works have looked at face recognition from sparse training data [10]. In [7], the authors leverage upon much larger mug shot gallery images or composite sketches for training. In [27], the authors evaluate the probability that two faces have the same underlying identity cause for recognition. However, these methods do not model style. In this article, we explore artist's style from the available sparse data.

DISCRIMINATIVE FEATURE SELECTION

A portrait is a visualization of an artist's aesthetic sensibilities blended with the sitter's personality. We therefore begin by understanding the relevant features for analysis based on a study of artistic trends during the period under study.

FACE AS SEEN BY ARTISTS

It is evident from [13] that while drawing a human body, much emphasis was laid upon maintaining the proportions of various

parts. It is purported that the principles for the canons of human body may have been defined by Egyptian artists, who divided the entire body into different parts and provided base-lines for their measurement. The importance of anthropometric ratios/distances was preserved even during the Renaissance era. According to da Vinci, in a well-proportioned face, the size of the mouth equals the distance between the parting of the lips and the edge of the chin, whereas the distance from chin to nostrils, from nostrils to eyebrows, and from eyebrows to hair-line are all equal, and the height of the ear equals the length of the nose [14].

A historical appraisal of facial anthropometry from antiquity upto Renaissance has been provided in [15] to compare artists' concept of human profile. A flattened nose, tilted forehead, and prominent upper lip were some of the features prevalent in Renaissance art works. In fact, prominent facial landmarks of a person were retained in works of the sitter by different artists as illustrated in [1].

CHOICE OF FEATURES

From the examples described above, it is clear that ancient Renaissance artists laid emphasis on two aspects in their renderings, which we use for our analysis.

LOCAL FEATURES

We use a set of 22 fiducial points to represent each face, these being (1, 2) forehead tips (left, right), (3) forehead center, (4) chin bottom, (5) nose top, (6) nose bottom, (7, 8) points on temple (left, right), (9, 10) chin and ear corners (left and right), (11, 12) points on chin (left and right), (13, 14) cheekbones (left and right), (15, 16) mouth corners (left and right), (17, 18) iris (left and right), (19, 20) left eye corners (right and left eye), and (21, 22) right eye corners (right and left eye). The precise location of these points is determined by registering a generic mesh on the face. Gabor jets corresponding to five frequencies and eight orientations are evaluated at each of these fiducial points. At a fiducial point n and for a particular scale and orientation j , the corresponding jet coefficient J_{n_j} is given by

$$J_{n_j} = a_{n_j} \exp(i\phi_{n_j}), \quad (1)$$

where a_{n_j} is the magnitude and ϕ_{n_j} is the phase.

ANTHROPOMETRIC DISTANCES

All images are normalized with respect to scale and orientation. A set of 11 salient distances is used to represent each face:

- distance between forehead tips
- distance between forehead center and chin bottom
- distance between nose top and bottom
- distance between points on the temples
- distance between chin and ear corners
- distance between points on the chin
- distance between each iris
- distance between cheekbones
- distance between mouth corners

- the width of the nose
- distance between the center of the forehead and bottom of the nose.

FEATURE EXTRACTION

Different artists are likely to depict and emphasize the aforementioned features in different ways. We wish to learn those features that are characteristic of an artist's style. We employ a method called the random subspace ensemble learning as it is capable of handling deficiencies of learning in small sample sizes [16]. Small sample sizes is very relevant to the present problem as we have very few works by an artist at our disposal (see the "Introduction" section). The random subspace method randomly samples a subset of the aforementioned features and performs training in this reduced feature space.

More specifically, we are given Z training portrait pairs and D features. Let L be the number of individual classifiers in the ensemble. We choose $d_i \leq D$ (without replacement) to be the number of features used in the i th classifier. For each classifier, we determine the match and nonmatch scores (as appropriate) using the d_i features as follows. We compute

$$s_{LF}(I, I') = \frac{1}{d_i} \sum_{n=1}^{d_i} s_n(J, J'), \quad (2)$$

where $s(J, J')$ is an average LF similarity measure between n corresponding Gabor jets computed across salient points in image pair (I, I') . To compute $s_n(J, J')$, we use the normalized similarity measure mentioned in [4] given by

$$s_n(J, J') = \frac{\sum_j a_{n_j} a'_{n_j}}{\sqrt{\sum_j a_{n_j}^2 \sum_j a'^2_{n_j}}} \quad (3)$$

Similarly, we compute AD similarity between image pairs (I, I') using the equation

$$s_{AD}(I, I') = e^{-\beta y}, \quad (4)$$

where y is the two-dimensional Euclidean distance between the AD vectors \vec{m} , \vec{n} that characterize images I, I' respectively (we use only those distances as selected by the random subspace classifier) and β is a coefficient that is chosen suitably to obtain a discriminative dynamic range of values. In our experiments, we set β to be five.

To identify features that give the highest separation between match and nonmatch scores, we then compute the Fisher linear discriminant function for each classifier. We choose the union of features from those classifiers that give the top k Fisher linear discriminant values as our style features.

IMPORTANCE OF THE CHOSEN FEATURES

Not all features identified by the previously mentioned method are equally important in representing an artist's style. To understand the importance of the chosen features, we consider the nonparametric statistical permutation test [17]. The permutation test helps in assessing what features are the same across all the instances belonging to an artist. Thus, features that are more invariant across the portraits by an artist can be perceived

to be more characteristic of that artist and hence be assigned greater importance. Permutation tests have been used to determine invariant features in artworks [2].

PERMUTATION TEST

The null hypothesis (H_0) is chosen to indicate that two portrait groups G_1, G_2 have the same average value in a particular feature; the alternate hypothesis (H_1) indicates that the average value of that feature is different in the two groups. Thus,

$$H_0: \mu_{G1} = \mu_{G2}; H_1: \mu_{G1} \neq \mu_{G2}, \quad (5)$$

where μ is the average value of a particular feature v under consideration in the two groups.

If the null hypothesis is true, then it should not matter when this feature v is randomly assigned among images in the group. For instance, let us assume that there is a certain way that the mouth corner is portrayed by Italian artist Bernini, whose works are included in our data set. On an average, if this appearance is the same across all images by Bernini, then the principle behind this test is that there will not be a significant difference if the mouth tips are randomly assigned across images in the group, i.e., assigning the feature of one sitter to the corresponding feature of another sitter.

Specifically, if there are N_s images by an artist Y , then we can divide these N_s images into two subgroups consisting of N_{s_1} and N_{s_2} images depicting different sitters. Let the feature values for the first group be $[v_1, v_2, \dots, v_{N_{s_1}}]$ and in second group be $[v_{N_{s_1}+1}, v_{N_{s_1}+2}, \dots, v_{N_{s_2}}]$. The permutation test is done by randomly shuffling $[v_1, \dots, v_{N_s}]$ and assigning the first N_{s_1} values, $[v_{(1)}, v_{(2)}, \dots, v_{(N_{s_1})}]$ to the first group and the rest N_{s_2} values $[v_{(N_{s_1}+1)}, \dots, v_{(N_{s_2})}]$ to the other group.

For the original two groups we compute,

$$\delta_0 = \left| \frac{1}{N_{s_1}} \sum_{i=1}^{N_{s_1}} v_i - \frac{1}{N_{s_2}} \sum_{i=1}^{N_{s_2}} v_{N_{s_1}+i} \right|, \quad (6)$$

where δ_0 denotes the variation in the feature v by artist Y as exhibited by various instances I_1, \dots, I_N in the two groups G_1 and G_2 . Thus, $\delta_0 = |\mu_{G1} - \mu_{G2}|$. For any two permuted groups we compute

$$\delta_s = \left| \frac{1}{N_{s_1}} \sum_{i=1}^{N_{s_1}} v_{(i)} - \frac{1}{N_{s_2}} \sum_{i=1}^{N_{s_2}} v_{(N_{s_1}+i)} \right|, \quad (7)$$

where δ_s denotes the variation in the feature v by artist Y after assigning v as depicted in I_i to an image not necessarily depicting the sitter in I_i .

We repeat this random shuffling of features among the images under consideration multiple times. The proportion of times $\delta_s > \delta_0$ is the p value. This value reflects the variation of the feature in the two groups. Smaller p denotes stronger evidence against

TO UNDERSTAND THE IMPORTANCE OF THE CHOSEN FEATURES, WE CONSIDER THE NONPARAMETRIC STATISTICAL PERMUTATION TEST. THE PERMUTATION TEST HELPS IN ASSESSING WHAT FEATURES ARE THE SAME ACROSS ALL THE INSTANCES BELONGING TO AN ARTIST.

the null hypothesis, meaning that the feature differed considerably in the two groups and thus less characteristic of the artist's style. We compute p values for each feature as described above. The computed p values are used as scaling factors (weights) in estimating the similarity scores (s_p) in (2) and (4). It is to be noted that this method can be employed when we have ≥ 12 images by an artist [21]; in cases where enough images/artists are

not available or when the artist is unknown, we use all 22 LF and 11 AD features with equal weight (of one assigned to all the features) in obtaining the LF/AD similarity scores.

FEATURE COMBINATION

The similarity scores obtained from LF and AD features may not be equally important in determining the similarity between portrait pairs. Furthermore, since the number of LF/AD features used are different, the scores need to be fused in a way such that the resulting distribution of match and nonmatch scores are as peaked and disjoint as possible. We employ the following algorithm toward this.

- 1) We consider a convex combination of the scores from the two measures LF and AD, i.e.,

$$\text{score} = \lambda s_{LF} + (1 - \lambda) s_{AD}, \quad (8)$$

with λ being varied from 0 to 1 in steps of 0.1.

- 2) For every λ , we evaluate the mean and standard deviation of match and nonmatch scores using the random sampling consensus algorithm [18] to prune outliers.

- 3) At each λ , we evaluate $J = (S_b/S_w)$, where S_b is between class variance and S_w is within class variance. We choose that value of $\lambda = \lambda_{opt}$ that gives the maximum value of J . This is essentially computing the Fisher linear discriminant [20].

Using the procedure described above, we compute similarity scores between portrait pairs that are known to depict the same sitters and different sitters to get match and nonmatch scores respectively. The resulting set of match and nonmatch scores, computed across various artists and sitters, are modeled as two Gaussians distributions (one for match scores and another for nonmatch scores). The mean and standard deviations of these distributions are estimated from training data. We refer to these match/nonmatch score distributions as the *PFS*.

VALIDATION OF THE LEARNED FEATURES

We wish to ascertain if the learned features are good representations of the portraits considered. To verify this, we perform twofold cross validation of the similarity scores.

VALIDATION OF ARTIST-SPECIFIC SIMILARITY SCORES

If the chosen features are robust representations of an artist Y , then the obtained match/nonmatch scores divided into two folds (groups),

[TABLE 1] HOW p VALUES ARE USED TO MAKE A DECISION ABOUT IDENTITY OF A PORTRAIT.

REFERENCE		DISTRACTERS		CONCLUSION
MATCH	NONMATCH	MATCH	NONMATCH	
$p > \alpha$	$p < \alpha$	$p < \alpha$	$p > \alpha$	MATCH
$p < \alpha$	$p > \alpha$	$p < \alpha$	$p > \alpha$	NO MATCH
$p > \alpha$	$p > \alpha$	NA	NA	NO DECISION
$p < \alpha$	$p < \alpha$	NA	NA	NO DECISION
$p > \alpha$	$p < \alpha$	$p > \alpha$	$p < \alpha$	NO DECISION

NA stands for "not applicable."

say A, B , should more or less be "similar" in that they come from the same artist. For this, we employ the Siegel–Tukey statistical test [23].

Siegel–Tukey Test

This is a nonparametric statistical method to test the null hypothesis (H_0) that two independent scores come from the same population (e.g., artist) against the alternative hypothesis (H_1) that the samples come from populations differing in variability or spread. Thus,

$$H_0 : \sigma_A^2 = \sigma_B^2, Me_A = Me_B ; H_1 : \sigma_A^2 \geq \sigma_B^2, \quad (9)$$

where σ^2 and Me are the variance and medians for the groups A and B . The test is entirely distribution-free. The absence of any normality assumption is an important feature of the test, because its parametric alternative, the F test for variance differences, is quite sensitive to departures from normality [22]. The p value obtained from this test, p_{st} , is given by

$$p_{st} = \Pr[X \leq U], \quad (10)$$

where U_A, U_B are the U statistics for groups A, B and $X \sim$ Wilcoxon (r, m) [21]. This is a measure of the confidence associated with the scores. Thus, if the learned features are good representations of an artist's style, they should be associated with a higher p_{st} value than the p_{st} value associated with scores obtained using all features.

VALIDATION OF THE PORTRAIT FEATURE SPACE

To validate the PFS computed across various artists/sitters, we randomly divide the known instances into two groups to perform two-fold cross validation. In fold 1, we use group one to learn the PFS and use group 2 to validate and vice versa in fold 2. Ideally, the learned PFS from the two folds should have the same statistics.

IDENTIFICATION FRAMEWORK

The goal of this article is to aid art historians by providing an alternate source of evidence in verifying uncertain portraits against a reference image by providing a quantitative measure of similarity. We use hypothesis testing for this purpose.

HYPOTHESIS TESTING

This is a method for testing a claim or hypothesis about a parameter in a population [19]. Next, we summarize it with respect to the learned PFS.

- A Null hypothesis claims that the match distribution accounts for the test's similarity score with reference better than nonmatch distribution. The alternate hypothesis is that nonmatch distribution models the score better.

- We set the level of significance α , i.e., the test's probability of incorrectly rejecting the null hypothesis, as 0.05, as per the behavioral research standard.

- We compute the test statistic using one independent nondirectional z test [19], which determines the number of standard deviations the similarity score deviates from the mean similarity score of the learned match/nonmatch distributions.

- We compute p values, which are the probabilities of obtaining the test statistic that was observed, assuming that the null hypothesis is true. If $p < \alpha$, we reject the null hypothesis.

IDENTITY VERIFICATION

To examine the validity of the chosen approach, we consider similarity scores of the test image with artworks known to depict persons different from the one depicted in the reference image. We call these images *distracters*. In cases where enough works of the same artist are not available, we consider similar works of other artists. If a test image indeed represents the same sitter as in the reference image, not only should its score with the reference image be modeled by the match distribution, but also its scores with distracter faces should be modeled by the nonmatch distribution.

ANALYSIS SCENARIOS

Following the procedure outlined previously, we compute the similarity scores of test cases with the corresponding reference image and with distracters. Table 1 lists various hypothesis test scenarios that can arise [19] and the corresponding conclusions that one can infer. Match and nonmatch cases are straightforward to infer from Table 1. In cases where both match and nonmatch distributions are likely to account for the score in the same way as in third and fourth rows of Table 1, it can be said that the learned PFS cannot accurately describe the test data. If the match distribution is more likely to account for both test as well as distracters (bottom row in Table 1), it can be inferred that the chosen features do not possess sufficient discriminating power to prune outliers. Thus, in these scenarios, it is not possible to reach any conclusion.

DATA SET

CHOICE OF IMAGES

We have employed a set of images belonging to Western Europe between the 15th and early 18th centuries. These images have been logically chosen by art historians to address different tasks such as 1) to test the relation of an unmediated image of the subject, e.g., a death mask to a work of portrait art like a painting, 2) to analyze a number of portraits of different sitters by the same artist to model artist's style, and 3) to verify if the identity of the ambiguous subject in a given image is same as that of a known subject in a reference image. The images belong to different media such as drawings, prints, paintings, sculptures,

death masks, etc. The data set consists of works by over 35 artists including Gian Lorenzo Bernini, Alessandro Algardi, and François Clouet, among others.

DESCRIPTION

The data set consists of approximately 271 images where the identity of the subject is known beyond a doubt. There are 20 test paradigms (with each having multiple image pairs to be compared) where the identity of the subject is in question and has to be compared against the reference image given in that paradigm. Table 2 provides a detailed description of the distribution of images in terms of the specific sitter and artist. Figure 2 provides an illustration of the data set. We encourage the interested reader to refer to the supplementary material in IEEE *Xplore* (<http://ieeexplore.ieee.org>) for a description of the sources for the portraits shown throughout this article.

EXPERIMENTS

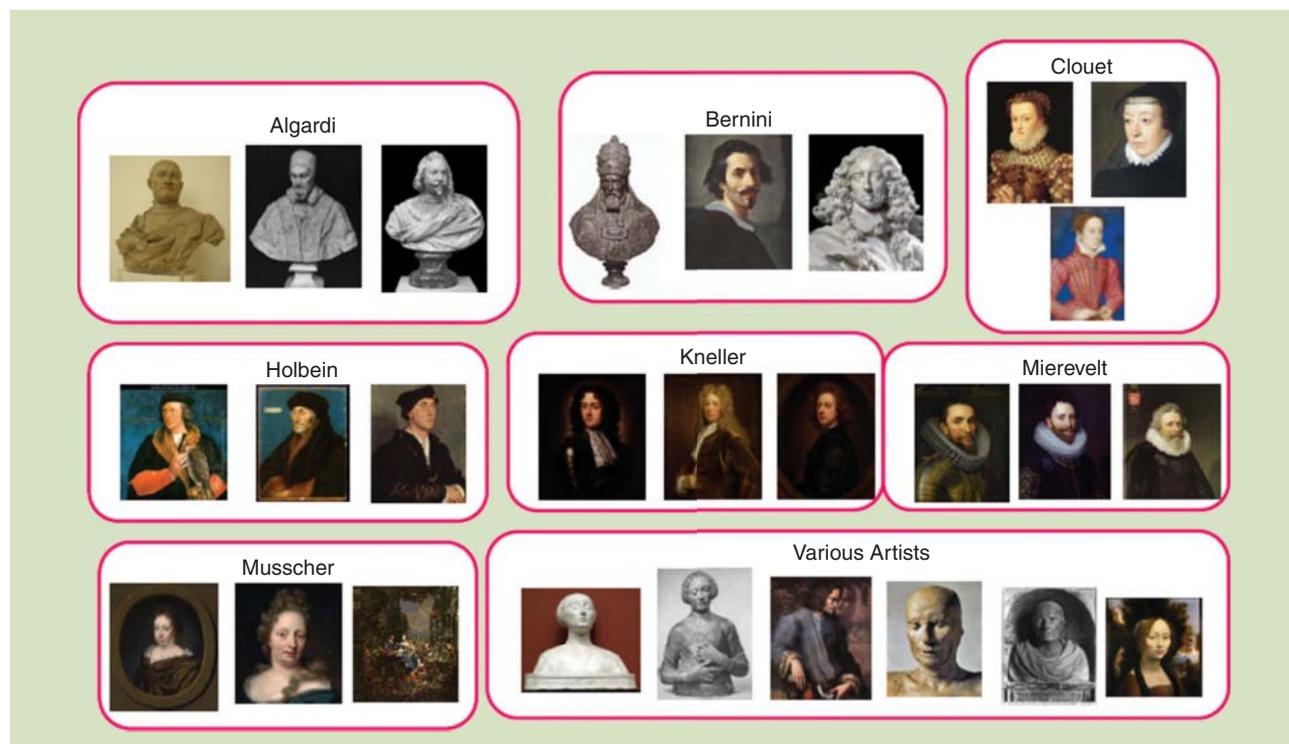
STYLE MODELING RESULTS

We first extracted the 22 LF and 11 AD for all the images. For those artists for which we had enough images to model their style, we learned the features characteristic of their style. In Figure 3(a), characteristic LFs with dots denoting the relative importance of the feature as per the p value of permutation test are depicted. AD features representative of the style was similarly determined for these artists; these being AD features 4, 8, 3, 7, and 2 for Algardi (see the section “Choice of Features” for a description of numbers); 1, 10, 7, 5, and 8 for Bernini; 2, 1, 8, 9,

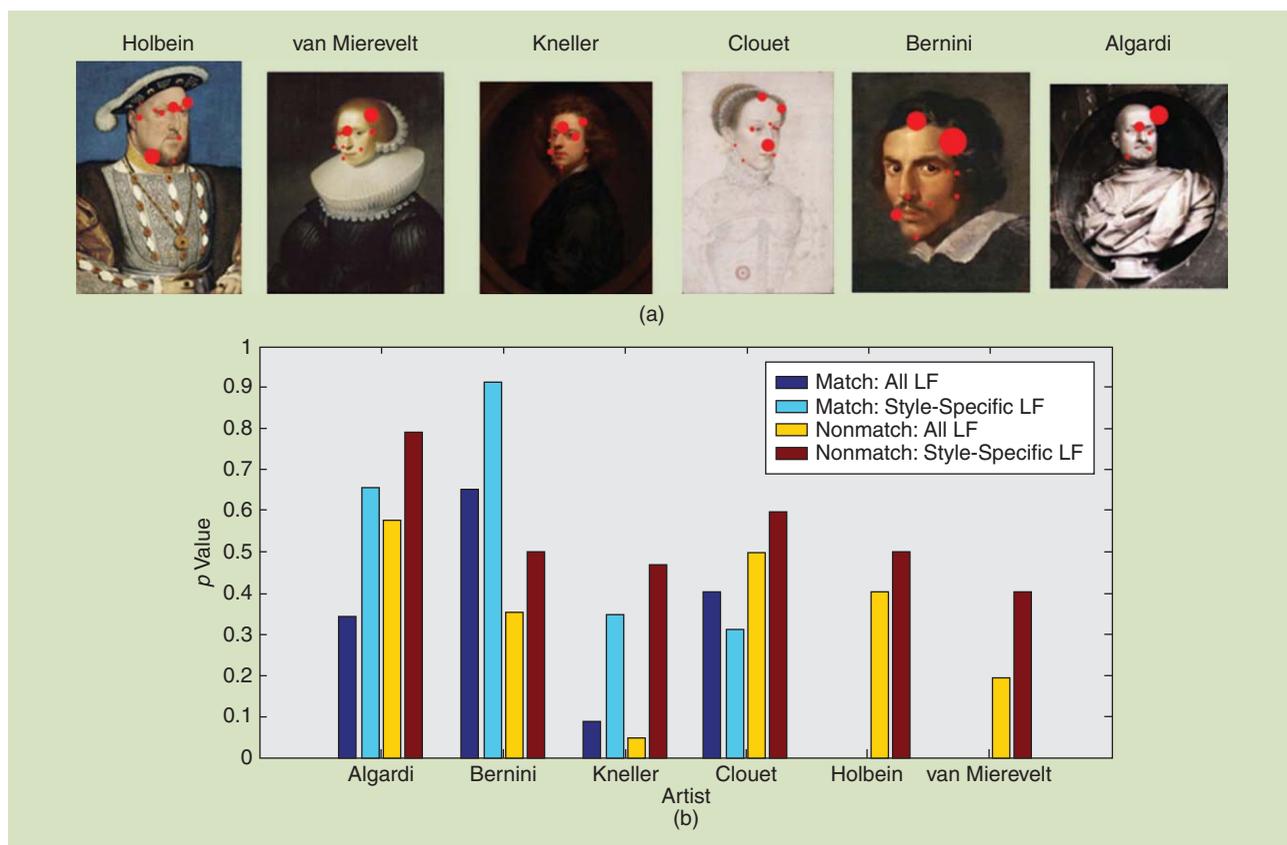
[TABLE 2] AN ILLUSTRATION OF IMAGE DISTRIBUTION: THE NUMBER OF IMAGES PER ARTIST.

ARTIST	NUMBER OF IMAGES	ARTIST	NUMBER OF IMAGES
ALGARDI	14	GIOTTO	6
BANDINI	1	HANSEN	3
BERNINI	33	HOLBEIN	45
BOTTICELLI	9	KNELLER	19
BRONZINO	5	LANGEL	1
BUGGIANO	2	LAURANA	10
CAFA	2	MANTENGA	3
CAMPIN	4	MASACCIO	4
CLOUET	14	RAPHAEL	5
DA FIESOLE	5	SIGNORELLI	5
DA VINCI	7	SITTOU	4
DE CHAMPAIGNE	7	STRINGA	4
DE BENINTENDI	3	THRONHILL	3
DEL CASTAGNO	3	TORRIGIANO	1
DELLA FRANCESCA	4	VAN MIEREVELT	24
VASARI	4	VAN MUSCCHER	18
GHIRLANDAIO	5	VERROCCHIO	6

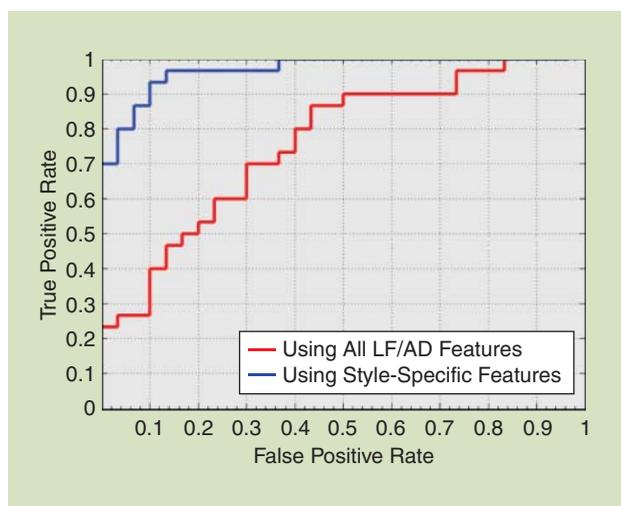
10, 5, and 4 for Godfrey Kneller; 5, 11, 2, and 7 for Clouet; 4, 6, 11, 7, and 3 for Michiel Jansz. van Mierevelt; and 2, 8, 11, and 3 for Hans Holbein The Younger. Features are listed in decreasing order of importance for each artist. We verified the validity of these features using the p_{st} value computed from Siegel–Tukey test. As illustrated in Figure 3(b), for almost all cases, the confidence of the similarity scores increased upon using only the style features, thus validating the chosen LF. Similar results were obtained for AD features. It is to be noted that the Siegel–Tukey test validates both style-specific match and nonmatch



[FIG2] An illustration of the data set across individual/multiple artists depicting different sitters.



[FIG3] (a) The importance of chosen features with bigger dots indicating more important features. (b) The validation of style through the Siegel-Tukey test.



[FIG4] The ROC curve for pairwise sitter validation upon using style features.

scores; whenever there are not enough images to obtain match scores, only the available nonmatch scores are validated. The receiver operating characteristic (ROC) curve shown in Figure 4 compares the performance for pairwise sitter validation upon using 1) style features and 2) all LF/AD features. The ROC demonstrates the improvement in pairwise validation upon using style features.

SIGNIFICANCE OF STYLE MODELING

These results could possibly aid art historians in attributing works to an artist that was not attributed to him/her before. It could also help in identifying unrecognized portraits by these artists more confidently. It might also be possible to understand the adherence to artistic canon and individual variations in art practices.

VALIDATION WITH KNOWN SITTERS

From the set of known identities, we obtained match and non-match scores. It is to be noted that wherever an artist's style could be modeled, we used only those (weighted) features in obtaining the LF/AD similarity scores and otherwise used all the LF/AD features followed by the feature combination strategy to fuse the similarity scores. The weight for the LF similarity score was found to be 0.55 and the weight for the AD similarity score was 0.45. Experiments showed that there was improvement in the performance upon fusing scores from LF and AD as against using any one of them. The values of mean of the PFS were 0.7246 (match) and 0.5926 (nonmatch) with standard deviations 0.043 and 0.052 respectively (see Figure 5).

IDENTITY VERIFICATION

We want to provide quantitative measures of similarity to uncertain test paradigms provided to us by art historians. In this article, we do not claim to provide the incontestable identity of the sitter

in question, but to only provide a complementary viewpoint, which could serve the art history community.

SIGNIFICANCE OF RESULTS FROM ART PERSPECTIVE

In these identification tests, support was given to previous scholarly opinion on a number of important cases. Among these were the posthumous bust of Battista Sforza by Laurana in the Bargello and a death mask cast also by Laurana in the Louvre Museum in Paris, France, shown in column 1 of Figure 6. A match suggests that, as was thought, the mask was that of Battista. It also supports the idea that the cast was quite closely followed by Laurana as a model, rather than, say, Piero della Francesca's profile portrait of Battista. A match was also indicated for Botticelli's *Portrait of a Lady at the Window* (c. 1475; widely thought to be a rendering of Smeralda Brandini) and Verrocchio's *Lady with Flowers* (c. 1475), the two portraits also sometimes being suggested by some to represent the same sitter, thus lending objective support to this position despite the two distinctly different personas conveyed in the images.

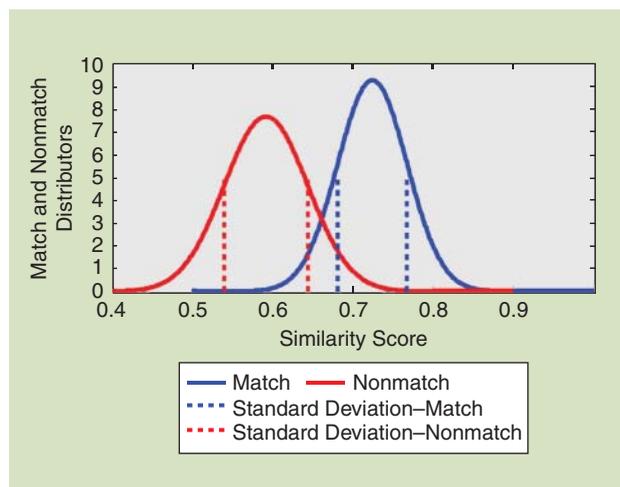
Tests strongly support the traditional supposition that Nicholas Hilliard's *Young Man Among Roses*, said to be perhaps the most famous miniature ever painted, represents Robert Devereux, second earl of Essex. The results of test scores between a portrait of a woman at the National Portrait Gallery in London thought by some to represent Mary Queen of Scots and eight other portraits known to be of Mary were almost startling in their support for the identification of the unknown portrait as Mary. Results also lend new support to previous opinion that the portrait at the National Portrait Gallery thought by some to depict James Scott, Duke of Monmouth, First Duke of Buccleuch, does portray Monmouth lying in bed after his beheading for treason.

The portrait shown in the bottom row of column six in Figure 6 was sent to us by the Italian astronomer Paolo Molaro, of what he believes may be the earliest known likeness of Galileo Galilei, perhaps painted around 1590. When tested against a chronological spectrum of eight other known portraits of Galileo, the results gave decreasing similarity scores within the match range for the chronologically three closest likenesses

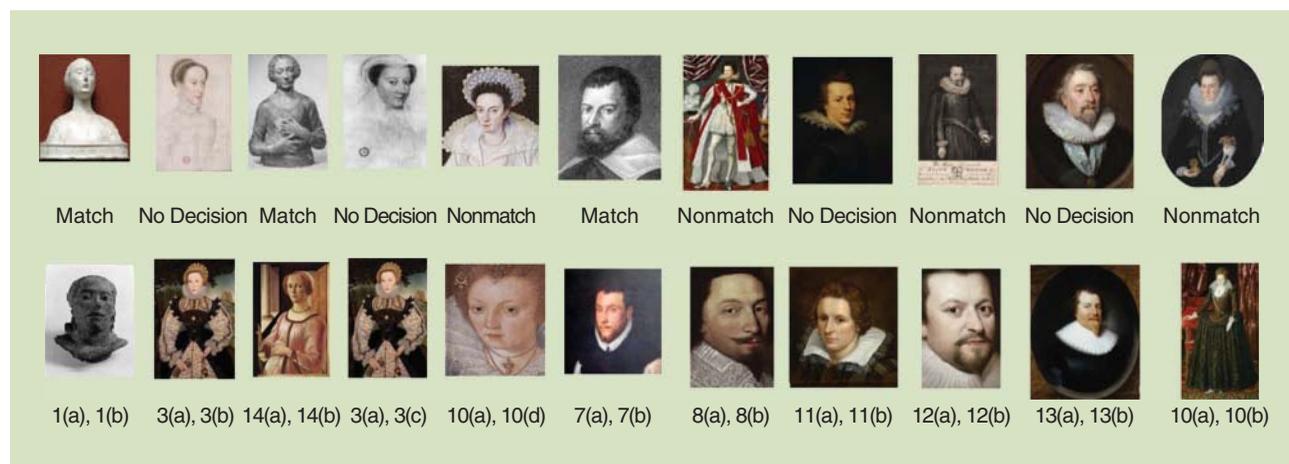
(1601–c. 1612). Thus, the test gives support to the identification of a previously unrecognized portrait as Galileo—possibly the earliest known portrait of Galileo. While age remains a challenge for FACES and requires more research, age differences of around ten years or so have not been too much of an obstacle.

A comparison between an unknown painting attributed to de Neve and a known portrait of George de Villiers, First Duke of Buckingham (column 7 of Figure 6), and a comparison between an unknown portrait and a known portrait of Lady Arabella Stuart (column five of Figure 6) gave nonmatch scores. A list of identification paradigms with results is provided in the supplementary material in *IEEE Xplore*. For a detailed description of these results from the art perspective, see [29].

The results of FACES are only as dependable as the images tested. Areas that would benefit from further research include modeling wide age differences, strong angle views (including profile images), and even the use of different media (e.g., terracotta as opposed to marble, chalk in contrast to oil, etc.).



[FIG5] The PFS showing the distribution of match and nonmatch scores along with their standard deviations.



[FIG6] Illustrations of identification tests with conclusions in the center. The bottom row shows images whose identity is uncertain; the numbers refer to corresponding images in the supplementary material supplied for this article in *IEEE Xplore*.

CONCLUSIONS

We presented a work that explores the feasibility of computer-based face analysis for portraiture. After a careful understanding of artistic conventions, we arrived at relevant features for analysis. Subsequently, using machine-learning tools, we learned a feature space describing the distribution of similarity scores for cases known to match/not match and also validated the same. We proposed a novel method to model artists' styles and to analyze uncertain portrait pairs. We believe that these results can serve as a source of complementary evidence to the art historians in addressing questions such as verifying authenticity, recognition of uncertain subjects, etc. For future work, we would like to explore modeling age variations in portraits and building family trees of artists/sitters.

ACKNOWLEDGMENTS

We would like to express our thanks to the National Endowment for the Humanities for two grants HD-51625-12 and HD-51735-13, which allowed us to establish proof of concept of our project.

AUTHORS

Ramya Srinivasan (rmalu001@ucr.edu) received the B.E. degree from the RV College of Engineering, Bangalore, India; the M.S. degree from the University of California, Santa Barbara; and the Ph.D. degree from the University of California, Riverside. Her research interests include computer vision, pattern recognition, machine learning, and soft biometrics, with a focus on face recognition.

Conrad Rudolph (conrad.rudolph@ucr.edu) is a professor of medieval art history at the University of California, Riverside. He has received fellowships and grants from the Solomon R. Guggenheim Museum, J. Paul Getty Museum, Andrew W. Mellon and Kress Foundations, as well as from the National Endowment for the Humanities and the College Art Association. He is the author of a number of books and articles, the latest being a study (and digital construction) of *The Mystic Ark*, a 12th-century image of all time, all space, all matter, all human history, and all spiritual striving, an image that was meant to be constructed again and again.

Amit Roy-Chowdhury (amitr@ee.ucr.edu) leads the Video Computing Group at the University of California, Riverside, with research interests in computer vision/image processing, pattern recognition, and statistical signal processing. His group is involved in research projects related to camera networks, human behavior modeling, face recognition, and bioimage analysis, with application domains that include homeland security, commercial multimedia, home infrastructure, computational biology, and digital arts. His research has been supported by various federal agencies and private companies. He is the author of *Camera Networks: The Acquisition and Analysis of Videos Over Wide Areas*.

REFERENCES

[1] S. West, *Portraiture*. Oxford, U.K.: Oxford Univ. Press, 2004.
 [2] J. Li, L. Yao, and J. Wang, "Rhythmic brushstrokes distinguish Van Gough from his contemporaries: Findings via automated brushstrokes extraction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1159–1176, 2012.

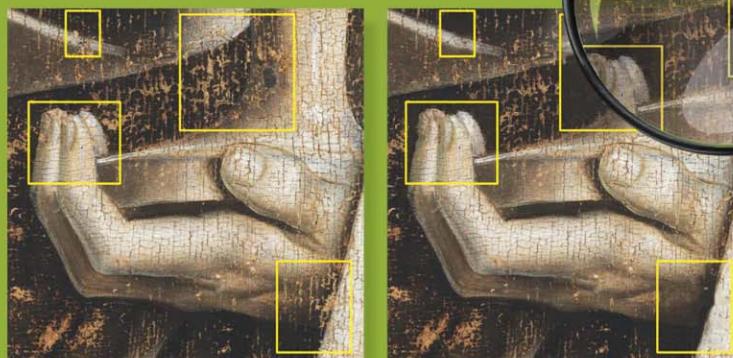
[3] W. Zhao and R. Chellappa, *Face Processing: Advanced Modeling and Methods*. New York: Academic Press, 2006.
 [4] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition by elastic graph bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, 1997.
 [5] D. Stork, "Computer vision and computer graphic analysis of paintings and drawings: An introduction to the literature," in *Computer Analysis of Images and Patterns* (Lecture Notes in Computer Science), vol. 5702, pp. 9–24, 2009.
 [6] C. Tyler, W. Smith, and D. Stork, "In search of Leonardo: Computer based facial image analysis of Renaissance artworks for identifying Leonardo as subject," *Hum. Vis. Electron. Imaging*, 2012.
 [7] B. Klare, Z. Li, and A. K. Jain, "Matching forensic sketches to mug shot photos," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 639–646, 2011.
 [8] N. D. Kalka, T. Bourla, B. Cukic, and L. Hornak, "Cross-spectral face recognition in heterogeneous environments: A case study on matching visible to short-wave infrared imagery," in *Proc. Int. Joint Conf. Biometrics*, 2011, pp. 1–8.
 [9] J. B. Tenenbaum and W. T. Freeman, "Separating style and content with bilinear models," *Neural Computat.*, vol. 12, no. 6, pp. 1247–1283, 2000.
 [10] W. Deng, J. Hu, and J. Guo, "Extended SRC: Undersampled face recognition via intraclass variant dictionary," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1864–1870, 2012.
 [11] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar, "Attribute and simile classifiers for face verification," in *Proc. IEEE Int. Conf. Computer Vision*, 2009, pp. 365–372.
 [12] R. Srinivasan, A. Roy-Chowdhury, C. Rudolph, and J. Kohl, "Recognizing the royals—Leveraging computerized face recognition for identifying subjects in ancient artworks," in *Proc. 21st ACM Int. Conf. Multimedia*, 2013, pp. 581–584.
 [13] A. Perrig, "Drawing and the artist's basic training from the 13th to the 16th century," in *The Art of the Italian Renaissance: Architecture, Sculpture, Painting, Drawing*. h.f. Ullmann Publishing GmbH: Cologne, Germany, 2007, pp. 416–441.
 [14] F. Vegter and J. Hage, "Clinical anthropometry and canons of the face in historical perspective," *Hist. Fac. Anthropol.*, vol. 106, no. 5, pp. 1090–1096, 2005.
 [15] L. G. Farkas, P. Sohm, J. C. Kolar, M. J. Katic, I. R. Munro, "Inclinations of facial profile: Art versus reality," *Plastic Reconstruct. Surg.*, vol. 75, no. 4, pp. 509–518, 1984.
 [16] T. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, 1998.
 [17] P. I. Good, *Permutation, Parametric, and Bootstrap Tests of Hypothesis*, 3rd ed. New York: Springer, 2009.
 [18] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *ACM Commun.*, vol. 24, no. 6, pp. 381–395, 1981.
 [19] M. Asadoorian, *Essentials of Inferential Statistics*. Lanham, MD: Univ. Press of America, 2005.
 [20] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. New York: Wiley-Interscience, 2001.
 [21] H. Mann and D. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Ann. Math. Statist.*, vol. 18, no. 1, pp. 50–60, 1947.
 [22] A. Gayen, "The distribution of the variance ratio in random samples of any size drawn from non-normal universe," *Biometrika*, vol. 37, nos. 3–4, pp. 236–255, 1950.
 [23] S. Siegel and J. Tukey, "A nonparametric sum of ranks procedure for relative spread in unpaired samples," *J. Amer. Statist. Assoc.*, vol. 55, no. 291, Sept. 1960.
 [24] A. Roy-Chowdhury, R. Chellappa, and H. Gupta, "3D face modeling from monocular video sequences," in *Face Processing: Advanced Modeling and Methods*. New York: Academic Press, 2005.
 [25] C. R. Johnson, Jr., E. Hendriks, I. J. Bereznyoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, and J. Z. Wang, "Image processing for artist identification—Computerized analysis of Vincent van Gogh's painting brushstrokes," *IEEE Signal Processing Mag.*, vol. 25, no. 4, pp. 37–48, 2008.
 [26] D. Stork and M. Johnson, *Recent Developments in Computer-Aided Analysis of Lighting in Realist Fine Art*, Image processing for art investigations. New York: Museum of Modern Art, 2010.
 [27] L. Peng, Y. Fu, U. Mohammed, J. H. Elder and S. J. D. Prince, "Probabilistic models for inference about identity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 144–157, Jan. 2012.
 [28] M. Vasilescu, O. Alex, and D. Terzopoulos, "Multilinear analysis of image ensembles: Tensorfaces," in *Proc. European Conf. Computer Vision*, 2002, pp. 447–460.
 [29] C. Rudolph, A. Roy-Chowdhury, R. Srinivasan, and J. Kohl, "FACES: Faces, art, and computerized evaluation systems—A feasibility study of the application of face recognition technology to works of portrait art," to be published.



[David Picard, Philippe-Henri Gosselin, and Marie-Claude Gaspard]

Challenges in Content-Based Image Indexing of Cultural Heritage Collections

© GHENT, KATHEDRALE KERKFABRIEK; PHOTO COURTESY OF KIK-IRPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

[Support vector machine active learning with applications to text classification]

Cultural heritage collections are being digitized and made available through online tools. Due to the large volume of documents being digitized, not enough manpower is available to provide useful annotations. In this article, we discuss the use of automatic tools to both automatically index the documents (i.e., provide labels) and

search through the collections. We detail the challenges specific to these collections as well as research directions that must be followed to answer the questions raised by these new data.

INTRODUCTION

Digitization of cultural heritage collections has recently become a topic of major interest. Large campaigns of digitization have been launched by several institutional and private entities to allow instant access to billions of documents. Thanks to these new portals, anyone can see these collections that were usually stored in

Digital Object Identifier 10.1109/MSP.2015.2409557

Date of publication: 15 June 2015



[FIG1] Examples of images and their corresponding labels (in French) from the BnF collection. Observe the variety in the keywords, from generic like “animal representation” in (a) and (d) to very specific like “Ptolémée V” in (c) and “Cuir Ornemental” in (b).

archives under restricted access [see, e.g., the online collection of the Library of Congress (<http://www.loc.gov/pictures>), or that of the Metropolitan Museum of Art (<http://www.metmuseum.org/collection/the-collection-online>)]. These campaigns are both rich in the number of digitized documents and targeted subjects. Together, these points make them appealing as a research topic and as a new tool for the final user.

While the idea of an open access to digital copies of all kinds of historical contents is very appealing, the size and variety of these new data lead to a wide range of new problems. In particular, the pace at which historical artifacts are digitized greatly exceeds the manpower needed to manually index them. By indexing, we mean labeling using carefully chosen keywords for all documents so as to ease the search through the entire collection. The online and open aspect of digitized collections raises new questions with respect to the uses emerging from the size of available data as well as from the variety of users searching them.

In this article, we consider automatic labeling and interactive search challenges. In automatic labeling, the goal is to automatically infer a set keywords for each newly digitized artifact. The thesaurus of all possible keywords can be very large and can contain concepts with varying semantic degrees. The main goal of automatic labeling is to ease the work of specialists searching throughout the entire collection by querying very precise keywords. In an interactive search, results of a query (either starting from an example or from a keyword) are graphically shown to the user. These results can be refined thanks to

user feedback, for instance, through the highlighting or the removal of some elements. The goal of an interactive search is to find documents that cannot be retrieved using keywords in a minimum amount of interaction.

CULTURAL HERITAGE DIGITAL IMAGE COLLECTIONS

This section examines the problem of indexing cultural heritage collections. For this purpose, we first present the example of a labeled subset of the Bibliothèque Nationale de France (BnF) image collection. Based on this presentation, we then detail the expected difficulties that are to be tackled when indexing such collections.

THE BnF IMAGE COLLECTION

Examples of images digitized by the BnF and their associated labels are shown in Figure 1. The collection, which is currently online, contains an estimated 275,000 images. Approximately 14% of them are labeled with one or several keywords. The images are pictures from any kind of cultural heritage artifact such as paintings, coins, tapestry, a manuscript, etc. The corresponding labels vary from very generic terms [e.g., “animal representation” in Figure 1(a) and (d)] to very specific ones [e.g., “Cuir Ornemental” in (b) and “Ptolémée V” in (c)].

The images and their keywords can be accessed online at <http://images.bnf.fr>. The advance search feature can be used to reveal the hierarchical architecture of the thesaurus, as well as the statistics of occurrences of the keywords. The main level of

the hierarchy is divided in broad categories stating the image acquisition geometry, the picture genre, its time and geographical information, and its subject in terms of objects and people.

In this article, we focus on a small subset of about 4,000 labeled images. In Figure 2, we show the distribution of the number of labels per image. As we can see, only around 120 images have only one label. The vast majority of images have between two and 15 labels.

In Figure 3, we show the distribution of the number of images per class. There are many classes with only one corresponding image, and these classes usually correspond to very specific labels. Examples of classes with one image are: *Philippe IV le Bel, époque Louis XIII, Papouasie-Nouvelle-Guinée, la guerre, électricité, marteau d'armes, générosité*. The greater the number of images per class, the fewer the number of classes (the peak at 100 in Figure 3 refers to the cumulative tail of the distribution). This distribution gives hints to the specificity of the classes found in cultural heritage collections.

Contrary to the Rijksbureau voor Kunsthistorische Documentatie (RKD) challenge presented in [1], indexing the BnF is much more complex. The tasks in the RKD challenge are to predict the author, the type of work, the material, and the time of creation. The number of samples available for these classes fairly outnumbers the one for the classes in the BnF. Furthermore, apart from the author identification, the classes of the RKD challenge are based on physical properties and not on semantic visual content. Nonetheless, the RKD provides a Web search engine allowing users to search its entire digitized collection with keywords (<http://www.rkd.nl>). Although these keywords are highly semantic like those of the BnF, they were not retained for the RKD challenge. Therefore, there is a need for a public data set that encompasses the full difficulty of indexing cultural heritage collections and allowing researchers to assess their tools.

OPEN QUESTIONS

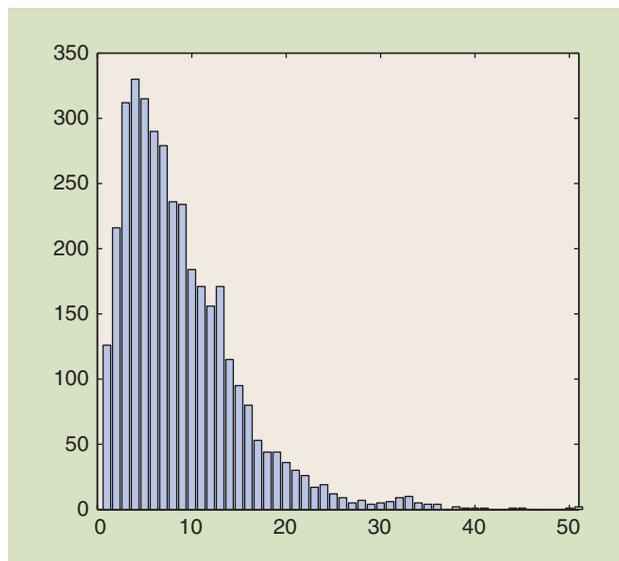
Compared to the generic image collections used in current computer vision benchmarks, labeling cultural heritage collection is much more difficult, due to the wide range of expected labels and to the very specific knowledge required to understand them. As we can see in Figure 1, some of the labels are sufficiently common to be inferred by everyone, but other ones require specific knowledge in history or in material science. To better evaluate the difficulty of labeling such collections, we propose to divide the labels in several domains:

- visual characteristic (shape, color, etc)
- semantic content (objects within the image, category of art: portrait, landscape, manuscript, etc)
- physical properties (canvas, marble, wood, paper, etc)
- historical information (production period, name of a character, style, etc.)
- geographical information (geographic name, towns, regions, map, etc.).

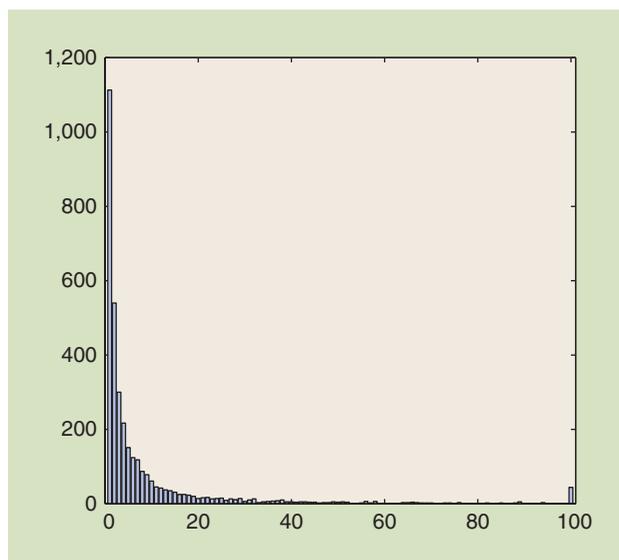
In this list, only the visual characteristic is effectively tackled by current content-based indexing methods. Semantic content has currently promising results thanks to the recent development in computer vision. There is, as far as we know, very few works on the remaining domains.

The main difficulty induced by this list is that methods need to be based on specific properties of the signal to perform well on some very specific classes. For example, predicting the type of paper might use different image characteristics (different scales, different modalities) than predicting the style. Designing a generic system that can automatically select the signal properties adapted to each specific class, without being a complex combination of ad hoc methods, is the main challenge of the proposed tasks.

The second problem induced by the wide class diversity is the scarcity of labeled samples. Indeed, some labels have by nature very few examples (e.g., “King Louis XI of France”) and very few of them



[FIG2] A histogram of the number of labels per image in the BnF collection. The last value corresponds to all images with more than 50 labels.



[FIG3] A histogram of the number of images per class in the BnF collection. The last value corresponds to all classes with more than 100 images.

are labeled. Unlike generic image collections, as the cultural heritage image collections grow, it is unlikely that the number of samples per class increases much for the vast majority of classes. Instead, it is more likely that the number of classes grows, while the number of samples per class remains almost constant. This arises challenging learning problems, since we have to build a large set of similarity measures and classifiers with very few samples.

CONTENT-BASED IMAGE SIMILARITY

The main technical challenge associated with the indexing of image collections is to be able to compute a numerical representation of each image and its associated similarity measure. This similarity measure aims at being as close as possible to human visual perception.

Generally speaking, the design of such representations and their associated metric is not an easy task, as it has to bridge the semantic gap [2]. Several families of similarity measure have been defined in recent years, in many cases with a very specific goal in mind. Indeed, the design of such measures is highly dependent on the application (e.g., object recognition, scene understanding), since it allows the incorporation of prior knowledge which usually boosts the performances. In the following, we list most of the existing families of content-based image similarity, detailing their original application, and showing some of their adaptations to art collections.

GLOBAL APPROACHES

Historically, the first approaches in defining a similarity between images were using a global index. A global index means that the numerical representation computes statistics on the properties of the signal at the scale of the whole image. Simple examples of such techniques include color histograms [3] or texture histograms [4].

With respect to cultural heritage images, global descriptors such as GIST have been used for image alignment and registration [5]. However, they suffer the same drawbacks as for the general image labeling and retrieval tasks: They are not able to handle classes discriminated by local visual properties. In particular, when the goal is to retrieve a specific instance of an object (e.g., a specific Roman emperor coin from the category coins), it becomes obvious that statistics at the image level are not sufficient to discriminate this specific instance from the other of the same category.

LOCAL DESCRIPTORS

To solve the precision problem of global descriptors, local keypoint matching techniques have successfully been proposed. The key idea behind the keypoint matching techniques is to select a set of salient regions in the image [denoted *region of interest (ROI)*], to compute a description of the content of the region and then to perform a pairwise matching of the keypoint descriptions between two images. The more keypoints match between images, the more they are considered similar.

The ROI detection step is based on salience measures like corner detectors [6] or a blob detector [7]. A good overview of the keypoint detection techniques can be found in [8]. Recently, it has been found that a dense extraction of keypoints leads to even better similarity measures, at the cost of a more complex matching step.

The ROI description is, in a sense, very similar to that of the global indexes, except that it is only computed on a small region of the image. The most used descriptors in current systems are scale invariant feature transform (SIFT) [7] or histogram of oriented gradients (HOG) [9], which basically computes a histogram of the gradient orientations in cells spanning the ROI. With such descriptors, the shape of the edges in the ROI is encoded.

Once descriptors are extracted, measuring the similarity between two images is akin to counting the number of matching pairs of descriptors. Given a descriptor d of the first image, its nearest neighbor $1NN(d)$ in the second image is considered as a match if their distance is less than a threshold relative to the distance with the second nearest neighbor $2NN(d)$:

$$d(x, 1NN(x)) < \lambda d(x, 2NN(x)) \quad (1)$$

with typically $\lambda = 0.6$ [7]. The assumption is that a given descriptor in the first image has a unique corresponding descriptor in the second image; together these form the closest pair.

To extend this matching scheme to an entire image collection, we consider the set $B = \cup_i B_i$ of all descriptor sets B_i of all images i in the collection. Then, for each descriptor from the query q , its k -nearest neighbors are retrieved from the entire collection. Every image receives as many votes as nearest neighbors it contains. Votes are summed up for all descriptors of the query and the image with the highest number of votes is the most similar.

$$s(q, i) = \sum_{d \in B_q} (kNN_B(d) \cap B_i). \quad (2)$$

However, such a matching scheme is unable to scale with a larger collection and larger sets of extracted descriptors. To run scalable searches, most of the accelerating schemes are based on an approximate nearest neighbor search in high-dimensional spaces, such as inverted files [10] or locality sensitive hashing [11].

With respect to cultural heritage collections, the main assumption driving local descriptor matching is relevant for duplicate or near duplicate retrieval. Searching for a seal, a coin, or a specific printed pattern are clear examples where the assumption holds. More semantic queries, such as author identification or time estimation can also be tackled using these approaches, depending on the scale of the images. For example, the brush stroke of a painter creating a specific pattern, such as the ear of a character, is a highly distinctive ROI that can be matched in another painting. In [12], the authors proposed a matching scheme to perform object detection in paintings while training on natural images and show matching local parts of an object improve the performances. However, they stay with very generic categories such as “dog” or “chair.”

The main challenge of these approaches in art-related collections is the selection of the right detector/descriptor couple to obtain satisfying results for a specific application. Unfortunately, no generic local detector/descriptor couple is able to tackle all of the similarities that can be defined in such collections due to the large variability of scale (from canvas threads to scene layout) and materials (parchment, canvas, marble, metal, etc.).

AGGREGATING METHODS

While local descriptors matching usually leads to very high performances, accelerating schemes are not sufficiently efficient to deal with very large collections. In particular, since all descriptors are kept, the amount of data to be stored grows with the number of extracted descriptors per image and the number of different modalities.

To overcome this problem, aggregating methods have been proposed to reduce the representation of an image from a set to a single vector. To perform the aggregation, most methods use a dictionary D of M prototypical descriptors $D = \{\mu_c\}_{1 \leq c \leq M}$. The set of descriptors B_i of an image i can then be described in terms of statistics over D . The first such method, called a *bag of visual words* (BoW), assigns each descriptor of the image to its closest entry in the dictionary and computes the histogram of such assignments [10]. Formally, the assignment corresponds to a quantization function q that returns a vector filled with zeros, except for a one at the component corresponding to the closest prototype,

$$q(x) = [\delta_{ik}]_i, k = \text{INN}(x). \quad (3)$$

The signature is simply the sum of all these vectors

$$x_i = \sum_{x \in X} q(x). \quad (4)$$

The interesting part of the BoW method is that it corresponds to a matching function between two sets of descriptors, where a match is found if and only if the two descriptors are assigned to the same prototype

$$s_{\text{BoW}}(q, i) = \left(\sum_{d \in B_q} q(d) \right)^\top \left(\sum_{p \in B_i} q(p) \right) \quad (5)$$

$$= \sum_{d \in B_q} \sum_{p \in B_i} \delta_{kl}, k = \text{INN}(d), l = \text{INN}(p). \quad (6)$$

More recently, refinement in the encoding of the descriptors have been proposed. For instance, instead of simply quantizing each descriptor to its closest prototype, sparse coding and related methods [13] propose viewing the encoding as a constrained reconstruction problem

$$q(x) = \underset{\alpha}{\text{argmin}} \|x - D\alpha\|^2 + \lambda\Omega(\alpha), \quad (7)$$

where $\Omega(\alpha)$ is a regularizer, typically the ℓ_1 norm to enforce a sparsity pattern in the coefficients α or a locality constraint to ensure descriptors are encoded by nearby prototypes in the case of locality constraint linear coding [13]. The incentives behind such encoding schemes are that the reconstruction term reduces information loss when compared to hard quantization approaches. Furthermore, the aggregation of codes introduces a minimum averaging effect due to the sparsity pattern in the codes. Since D is an overcomplete dictionary and using the generalized Parseval identity, there is a relationship between the dot product of two descriptors in the descriptors space and the dot product of their coding coefficients. As a consequence, the dot product of two signatures is related to a matching scheme where the descriptors are compared using the dot product.

The idea of using a matching scheme that can be linearized into a single vector has been proposed in several methods. In vectors of locally aggregated descriptors (VLADs) [14], the authors assign each descriptor to its closest prototype and then encode the differences between the descriptor and the prototype

$$q(d) = [\delta_{ik}(d - \mu_i)]_i, k = \text{INN}(d). \quad (8)$$

The aggregation is simply the sum of all codes, like in BoW. The corresponding matching scheme compares only descriptors assigned to the same prototype and computes the match using the dot product of the descriptors centered on their respective prototypes

$$s_{\text{VLAD}}(q, i) = \sum_{d \in B_q} \sum_{p \in B_i} \delta_{kl} \langle d - \mu_k, p - \mu_l \rangle, \quad (9)$$

$$k = \text{INN}(d), l = \text{INN}(p). \quad (10)$$

By looking at VLADs, we can clearly see the bridge between matching schemes and the comparison of different statistics over D . In that sense, the BoW is a piecewise constant matching scheme and corresponds to a zero-order statistic, while VLADs are a piecewise linear matching scheme and corresponds to a first-order statistic.

To improve the discriminative capability of the similarity measure, higher orders have been proposed. In particular, Fisher vectors [15] consider second-order information by computing the dictionary as a Gaussian mixture model. Then, it assigns the descriptors to all components of the mixture proportionally to their likelihood. Finally, it computes the first- and second-order moments of the descriptors with respect to each component. Using a hard assignment, vectors of locally aggregated tensors (VLATs) [16] computes higher-order moments using the tensor products of descriptors. The final signature is then the concatenation of all orders. The authors show that the dot product between the signatures is equivalent to a matching scheme using the dot product between the descriptor raised to the power t , which in turn is an approximation of a Gaussian matching kernel between the descriptors using a Taylor expansion of order t .

Considering cultural heritage images, aggregating approaches are likely to generalize in the labeling task, as shown with Fisher vectors in [17]. In an interactive search, their relation to keypoint matching is also likely to obtain a good accuracy thanks to the discriminative power of such schemes.

DEEP ARCHITECTURES

In contrast with the two steps of the local descriptor aggregation approaches, deep architectures stacking several layers of encoding have also been proposed. While deep neural networks have been proposed for a long time [18], their recent success in generic image classification benchmarks revived recent development of such methods.

The greatest advances were made with the use of convolutional neural networks (CNNs), which alternate layers of convolutional filters and layers of pooling operations [19]. The weights of the neurons on the convolution layers correspond to the coefficient of filters and can be trained in two steps. First, a pretraining step minimizing the reconstruction error trains a preliminary set of autoencoding filters. Second, the filters coefficients are tuned by

[TABLE 1] RESULTS (% mAP) OF THE LABELING TASK FOR DIFFERENT FEATURES.

	FISHER VECTORS	CPVLAT	CNNs
SEMANTIC	16.9	14.4	16.7
VISUAL	35.4	25.2	32.1
HISTORICAL	18.8	16.8	20.2
GEOGRAPHICAL	34.7	28.1	31.3
PHYSICAL	31.2	23.4	28.7
COMBINED	27.4	21.6	25.8

back propagation of the classification error. This second step helps locate the right combination of filters.

This repetition of convolution, nonlinear operation, and pooling bears some similarity with the multiscale analysis performed by wavelets. Indeed, the authors of [20] propose a deep architecture composed of layers of wavelet filters combined with a nonlinear operator (modulus) to obtain invariance to certain transforms.

Although most CNNs are used directly in classification tasks, it was empirically shown that the layers before the classification provide very good image representation that can be used in almost any image similarity related task [21]. It has also been recently shown that stacking many layers [22] further improves the performances, which raises the question of the structure rules to follow when designing a deep neural network.

Since CNNs provide a strong baseline for image features, they were already used with paintings in [17]. The authors propose classifying paintings using a training set of natural images taken

from Google images. They achieve surprisingly good performances considering the discrepancy between the objects in natural images and their depiction in paintings. However, the experiments are limited to a small number of generic classes (e.g., boat, car, horse), and cannot be easily extended to the very specific classes we consider in this work.

APPLICATIONS AND EXPERIMENTS

In this section, we present baseline results on the BnF collection to show the complexity of the challenging applications. We set up a benchmark with rigorous evaluation procedure allowing the comparison of different visual features (the benchmark can be downloaded at http://perso-etis.ensea.fr/~picard/bnf_bench/). First, we present results on automatic labeling and then on an interactive search.

AUTOMATIC LABELING

In these experiments, we only considered classes with more than ten images to be able to compute relevant statistics. We found 569 classes corresponding to this criterion, with the following repartition: Semantic (459), visual (62), historical (26), geographical (14), and physical (8). We used a standard classification approach consisting in a single midlevel feature per image combined with a linear support vector machine (SVM). This setup compares to most pipeline used in current academic challenges. We trained the SVM on a one-versus-rest mode for each class. Using fivefold cross-validation, we computed the average precision (AP) for each class.



[FIG4] Examples of (a) poor and (b) well performing classes for various features.

We compare three different types of midlevel features: Fisher vectors [15], compact preprocessed vector of locally aggregated tensors (CPVLAT) [23], and deep CNN-based features taken from the penultimate layer of CNNs in [24]. These three different features allow to assess the behavior of different families of methods, in particular probabilistic models for Fisher vector, keypoint matching for CPVLAT, and deep neural networks for CNNs. Contrary to other methods, CNNs are trained on a much larger data set (ImageNet) not related to cultural heritage.

Table 1 shows the results of each feature with respect to the different categories. Fisher vectors almost consistently outperform the other methods, which shows that classical computer vision methods provide a strong baseline for visual similarity even with a large variety of classes. As we can see, the combined mean AP (mAP) is less than 28%, which shows the complexity of the task. Semantic is the most difficult category, whereas visual is the easiest, although these categories are the ones containing most classes and thus most samples.

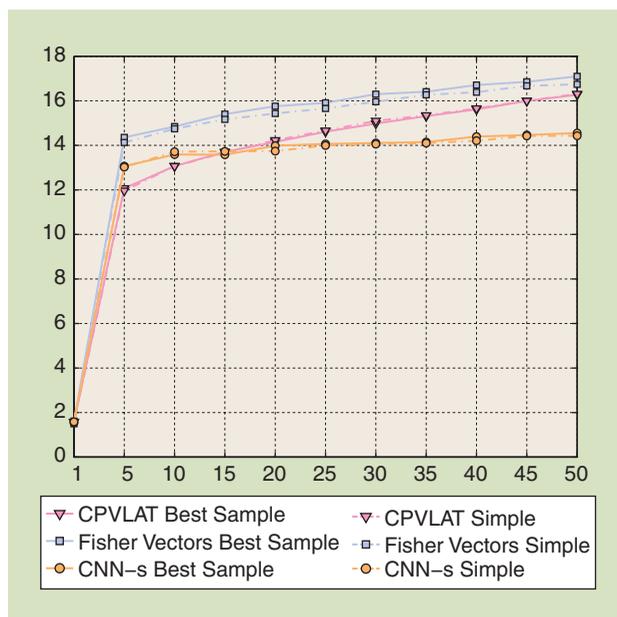
We show in Figure 4 detailed results for some of the best and worst classes for each feature. While it is easy to understand why some classes are difficult, like “drum” or “whip,” which are tiny, blurry details combined with few training samples (between ten and 15), it is worth remarking that most of the easy classes obtain good results mainly for statistical reasons. For instance, all the images with the label “Danish” are from the same data set of sketches representing the Danish army during the 18th century. Likewise, the sewer class contains only maps of the sewers of Paris that are visually very similar.

INTERACTIVE SEARCH

In an interactive search, we consider the scenario where a user is searching for a subset of the collection corresponding to a specific concept. This can be the case when a new class is inserted in the thesaurus, for example. In that case, the user wants to retrieve all of the images belonging to the concept with the minimum amount of interaction with the system (i.e., the minimum number of clicks). A typical session is as follows: starting from a single image belonging to the class, we interactively select five images, label them and add them to the pool of known images; retrain the classifier; and present the results for the next round of interaction until 50 images are labeled.

To evaluate performances in this setup, we compared the performances of two strategies of interaction by measuring the mAP against the number of labeled images (averaging ten sessions per class). The “BestSample” strategy selects the most relevant sample as evaluated by the current classifier (i.e., $\max_x f(x)$ with f the current classifier), while the “simple” strategy [25] selects the sample closest to the margin (i.e., $\min_x |f(x)|$). Note that the labeled images are counted when computing the mAP, which biases the results compared to the classical labeling task. However, counting the labeled samples is coherent with the end-user application that should provide all correct results to the user, including those seen during the interaction.

We show in Figure 5 the mAP against the number of labeled images. As we can see, both strategies perform about the same



[FIG5] The mAP against the number of labeled images for two active strategies.

regardless of the features, with a slight advantage for the BestSample strategy. This can easily be explained by the nature of our data, where most classes are very small and diverse. In such a case, samples close to the margin are likely to be negative ones mainly due to the low appearance probability of the relevant class. Fisher vectors also outperform other features in this task. Contrary to the labeling task, CPVLAT offers better performances than CNNs, which means this feature is more able to discriminate and less likely to generalize, which is consistent with the retrieval task. This is probably due to the strong relation of CPVLAT with keypoint matching schemes.

Furthermore, it is interesting to note that the best mAP performance after 50 labeled images is only around 17%. Considering that there are few classes with more than 50 images and recalling that the labeled images are counted in the mAP, a good active learning strategy should be able to obtain much higher mAP.

CONCLUSIONS

The main conclusions of this article are threefold:

- 1) We discussed the availability of large cultural heritage image collections that are currently being digitized, and which we believe will be a major topic of interest in the content-based indexing community.
- 2) By carefully looking at how these collections are currently manually indexed, we detailed specific tasks that are out of the scope of current content-based indexing problems, although they are of great interest for the users of these cultural heritage collections.
- 3) By performing a review of currently available techniques of content-based indexing, and testing a baseline method on the BnF collection, we show that there is still a lot of research to be done to achieve satisfactory results.

The main open questions concerning the design of similarity measures specifically tailored for cultural heritage collections are with respect to the wide range in type and scale of signal properties to be encoded in the signature. In particular, it is very difficult with the presented techniques to design a numerical representation that can encode both microscopic properties, such as canvas patterns or brush strokes in painting, and macroscopic properties, such as a scene layout or a direction of illumination. Designing a similarity measure that can tackle all these different types and scales of similarities is probably the biggest challenge in the indexing of cultural heritage collections.

The second problem arises from the paradoxical scarcity of the data. While data are massively available, including cultural heritage images, examples of specific categories may not. For example, only a few examples of an antic coin may be available. In such circumstances, methods that require a large amount of data to train their parameters are hindered and may not be able to obtain satisfactory results. Designing methods able to perform well on very precise similarity tasks with only few relevant training examples is the second challenge in the indexing of cultural heritage collections.

ACKNOWLEDGMENT

This work is part of the ASAP research project funded by the PATRIMA French Labex and the "Fondation pour les Sciences du Patrimoine."

AUTHORS

David Picard (picard@ensea.fr) received the M.Sc. degree in electrical engineering in 2005 and the Ph.D. degree in image and signal processing in 2008. He joined the ETIS Laboratory at the École Nationale Supérieure de l'Électronique et ses Applications, France, in 2010 as an associate professor within the Multimedia Indexing and Data Integration team. His research interests include image processing and machine learning for visual information retrieval with a focus on kernel methods for multimedia indexing and distributed algorithms.

Philippe-Henri Gosselin (gosselin@ensea.fr) received the Ph.D. degree in image and signal processing in 2005 in Cergy, France. After two years of postdoctoral positions at the LIP6 Laboratory, Paris, France, and at the ETIS Laboratory, Cergy, France, he joined the Multimedia Indexing and Data Integration team in the ETIS Laboratory as an assistant professor and was promoted to full professor in 2012. His research focuses on machine learning for online multimedia retrieval. He is involved in several international research projects with applications to image, video, and three-dimensional objects databases.

Marie-Claude Gaspard (marie-claude.gaspard@bnf.fr) is a computer science engineer. She joined the Bibliothèque Nationale de France (BnF) in the Information Technology Department after several years in service companies. She then moved to the National Bibliographic Agency, where she started working in the field of library science. She completed her training with a master's degree in information and communication technologies. She then headed the team in charge of the BnF

Image Bank within the Reproduction Department. She is currently an IT project manager.

REFERENCES

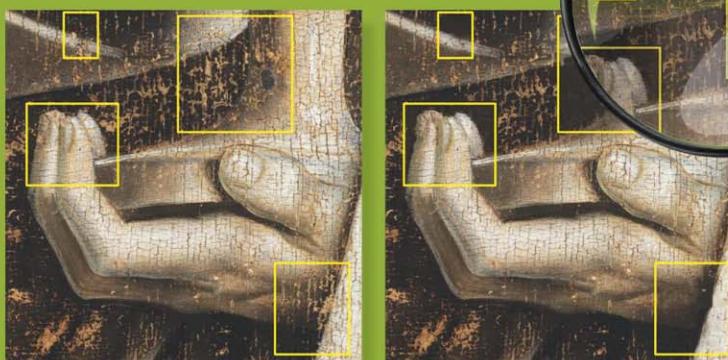
- [1] T. Mensink and J. van Gemert, "The Rijksmuseum challenge: Museum-centered visual recognition," in *Proc. ACM Int. Conf. Multimedia Retrieval (ICMR)*, 2014.
- [2] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.
- [3] J. R. Smith and S.-F. Chang, "VisualSEEk: A fully automated content-based image query system," in *Proc. 4th ACM Int. Conf. Multimedia Proc. (MULTIMEDIA '96)*, New York, pp. 87–98. [Online]. Available: <http://doi.acm.org/10.1145/244130.244151>
- [4] B. S. Manjunath and W.-Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, 1996.
- [5] B. C. Russell, J. Sivic, J. Ponce, and H. Dessales, "Automatic alignment of paintings and photographs depicting a 3D scene," in *Proc. IEEE Int. Conf. Computer Vision Workshops (ICCV)*, 2011, pp. 545–552.
- [6] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vision Conf.*, Manchester, U.K., 1988, vol. 15, p. 50.
- [7] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [8] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR 2005)*, 2005, vol. 1, pp. 886–893.
- [10] J. Sivic and A. Zisserman, "Video Google: A text retrieval approach to object matching in videos," in *Proc. 9th IEEE Int. Conf. Computer Vision*, 2003, pp. 1470–1477.
- [11] M. Datar, N. Immorlica, P. Indyk, and V. S. Mirrokni, "Locality-sensitive hashing scheme based on p-stable distributions," in *Proc. ACM 20th Annu. Symp. Computational Geometry*, 2004, pp. 253–262.
- [12] E. J. Crowley and A. Zisserman, "The state of the art: Object retrieval in paintings using discriminative regions," in *Proc. British Machine Vision Conf.*, Sept. 2014.
- [13] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 3360–3367.
- [14] H. Jégou, F. Perronnin, M. Douze, J. Sánchez, P. Pérez, and C. Schmid, "Aggregating local image descriptors into compact codes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1704–1716, 2012.
- [15] F. Perronnin, J. Sánchez, and T. Mensink, "Improving the fisher kernel for large-scale image classification," in *Computer Vision—ECCV 2010*. New York: Springer, 2010, pp. 143–156.
- [16] D. Picard and P.-H. Gosselin, "Efficient image signatures and similarities using tensor products of local descriptors," *Comput. Vis. Image Understand.*, vol. 117, no. 6, pp. 680–687, 2013.
- [17] E. J. Crowley and A. Zisserman, "In search of art," in *Proc. Workshop Computer Vision for Art Analysis (European Conf. Computer Vision)*, 2014.
- [18] B. B. Le Cun, J. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," in *Proc. Advances in Neural Information Processing Systems*, 1990.
- [19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.
- [20] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1872–1886, 2013.
- [21] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2014, pp. 512–519. doi: 10.1109/CVPRW.2014.131
- [22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, arXiv preprint arXiv:1409.1556, 2014.
- [23] R. Negrel, D. Picard, and P.-H. Gosselin, "Web-scale image retrieval using compact tensor aggregation of visual descriptors," *IEEE Multimedia*, vol. 20, no. 3, pp. 24–33, 2013.
- [24] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," in *Proc. British Machine Vision Conf.*, 2014.
- [25] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *J. Mach. Learn. Res.*, vol. 2, pp. 45–66, Nov. 2002.



[François Lozes, Abderrahim Elmoataz, and Olivier Lézoray]

PDE-Based Graph Signal Processing for 3-D Color Point Clouds

© GHEENT, KATHEDRALE KERKFABRIEK. PHOTO COURTESY OF KIK-IRPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

[Opportunities for cultural heritage]

With the advance of three-dimensional (3-D) scanning technology, the cultural heritage community is increasingly interested in 3-D scans of cultural objects such as antiques, artifacts, and heritage sites. Digitization of these objects is commonly aimed at heritage preservation. Since 3-D color scanning has the potential to tackle a variety of traditional documentation challenges, the use of signal processing techniques on such data can be expected to yield new applications

that are feasible for the first time with the aid of captured 3-D color point clouds. Our contributions are twofold. First, we propose a simple method to solve partial differential equations (PDEs) on point clouds using the framework of partial difference equations (PdEs) on graphs. Second, we survey several applications of 3-D color point cloud processing on real examples for which signal processing researchers can develop tools that can be valuable for museum specialists. The results of these methods have been reviewed by experts in the arts and found promising.

INTRODUCTION

Historians, archaeologists, and museum curators are interested in the preservation of pieces of art and want to make them available

Digital Object Identifier 10.1109/MSP.2015.2408631

Date of publication: 15 June 2015

to as wide an audience as possible. While direct encounters with the material pieces of art will always have an essential value for museum specialists, one important concern for museum curators is the preparation of catalogs for the purpose of object identification and examination. Traditionally, museum specialists have used two-dimensional (2-D) imaging techniques such as hand drawings, photographs, or document scans; see [1] and [2]. Unfortunately, these standard formats have the limitation of being selective and insufficient to record nuanced information about the complete shape, color, and texture of the object. In addition, some objects might undergo important

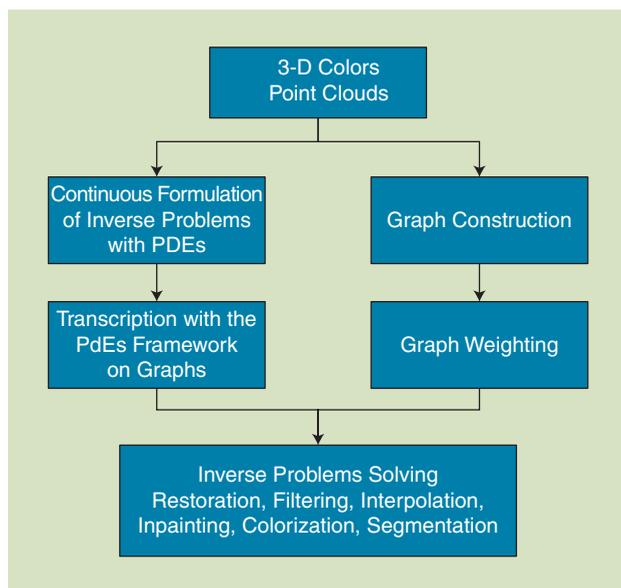
WITH THE ADVANCE OF 3-D SCANNING TECHNOLOGY, THE CULTURAL HERITAGE COMMUNITY IS INCREASINGLY INTERESTED IN 3-D SCANS OF CULTURAL OBJECTS SUCH AS ANTIQUES, ARTIFACTS, AND HERITAGE SITES.

changes over time, and in this case, the preservation of the object state (at a given temporal snapshot) is important. These issues can now be addressed by using 3-D color scanning, which has the ability to record the complete object in great detail. The production of digital 3-D copies of pieces of art has therefore gained considerable attention from the

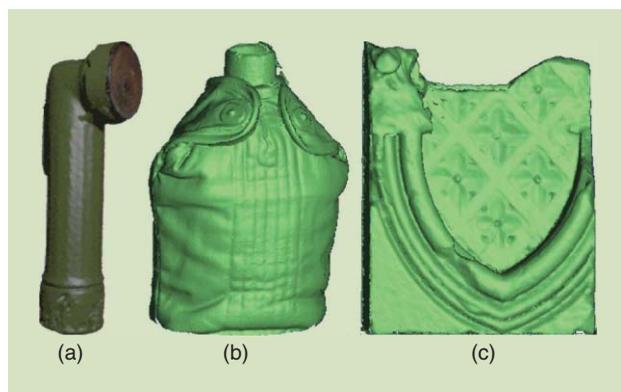
cultural heritage community [3]. The recent proliferation of commercial 3-D digital scanning devices has led to the establishment of 3-D scanning as a practical reality in the field of cultural heritage preservation [4]. The interest is evident: it provides a high-precision digital reference of a cultural object and makes possible mass distribution and consultation.

Most traditional signal processing methods in art investigation, such as filtering, PDE-based processing, and wavelets, are designed for data defined on regular Euclidean grids; see [2] and [5]. With 3-D color point clouds, the data takes the form of unstructured raw point samples without any attached geometry. Traditional signal processing tools cannot be directly applied on raw 3-D color point clouds since there is no structuring information. In addition, it is not desirable to transform point clouds to meshes since this requires sampling points with a loss of accuracy and this is not compatible with the goal of high-fidelity conservation. Therefore it is essential to investigate a framework for the adaptation of signal processing tools for 3-D point cloud processing.

Our contributions are twofold. First, we propose a simple method for adapting and solving PDEs on point clouds. This method relies on the framework of PdEs on weighted graphs [6] as shown in Figure 1. Second, we survey several applications in 3-D color point cloud processing where signal processing tools can be applied to images of interest to art and museum specialists. From a signal processing point of view, most of these problems can be formulated as inverse problems for graph signals [7] (signals living on graphs representing 3-D color point clouds as in Figure 1).



[FIG1] The PdEs framework on weighted graphs can be applied to solve PDEs on 3-D point clouds.



[FIG2] Some objects from World War II acquired with our 3-D scanner at the Mémorial de Caen. (a) A lamp. (b) A canteen. (c) Part of a destroyed church.

3-D COLOR POINT CLOUDS

This section details how 3-D images of historic or art items can be acquired. We first provide a review of the benefits and disadvantages of each of the current technologies. In cultural heritage, two main approaches are used to digitize a 3-D object.

The first approach is photogrammetry. This is an image-based modeling technique where multiple photos are taken to retrieve 3-D data coordinates. Photogrammetry requires only cameras and is the least expensive approach. However, it works only with small objects that have regular geometric shapes. The second approach is laser scanning. In this approach, the scanner acquires many 3-D points of the object to digitize. The precision is much better but both the cost and the acquisition time are greater. In this article, we focus on this latter type of 3-D raw point clouds.

Figure 2 shows the acquisition of several World War II items, acquired by a Konica Minolta VIVID-9i noncontact 3-D laser

scanner with a lens of $f = 25$ -mm focal length at a distance of 0.6 m. According to the specification of the scanner, this corresponds to an accuracy of digitizing of ± 0.05 mm.

A point cloud is composed of raw data $\mathcal{D} = \{d_1, \dots, d_n\}$. One associates to each $d_i \in \mathcal{D}$ a 3-D coordinate vector $\mathcal{P} = \{p_1, \dots, p_n\}$ with $p_i(p_i^x, p_i^y, p_i^z)^T \in \mathbb{R}^3$, and a color vector $\mathcal{C} = \{c_1, \dots, c_n\}$ with $c_i(c_i^R, c_i^G, c_i^B)^T \in \mathbb{R}^3$. We define for a given point cloud a function $f^*: \mathcal{D} \rightarrow \mathbb{R}^3$, which is either $f^*: \mathcal{D} \rightarrow \mathcal{P}$ or $f^*: \mathcal{D} \rightarrow \mathcal{C}$. The function f^* is difficult to process and to analyze with classical signal processing techniques because the underlying point cloud lacks of any topological structure. Therefore, classical 2-D signal processing techniques cannot be considered and new dedicated graph signal processing techniques are needed.

GRAPH CONSTRUCTION FROM POINT CLOUDS

The proposed approach to process point clouds starts by creating a weighted graph from a given point cloud. This section explains how to build such a graph. This latter method is composed of two parts: first the topology of the graph is defined and then the edges of the graph are weighted.

GRAPH CONSTRUCTION

This section discusses how to build a weighted graph from 3-D point clouds. The creation of a graph requires several steps.

- 1) Vertices are created from the raw data.
- 2) The vertices are connected with edges to build a proximity graph based on geometrical structure of the point cloud.
- 3) Weights associated with each edge are calculated.

Weights of edges are deduced from values associated to vertices (the graph signal) and patches can be used to compute a better similarity value that takes account of local neighborhood similarities.

Graph creation from point clouds is challenging. Indeed, the structuring information is lacking, and the data is not naturally organized in a manifold. Therefore, the set of edges cannot be easily determined. Given a point cloud P with associated coordinates $\{\mathbf{p}_1, \dots, \mathbf{p}_n\} \in \mathbb{R}^3$, there are many ways of associating a graph to such a data set. Since point clouds data exhibit a geometrical structure, proximity graphs are preferable: if two data points satisfy certain geometric requirements, the corresponding vertices in the graph are connected by an edge. To each raw data point $d_i \in \mathcal{D}$, one associates a vertex of a proximity graph \mathcal{G} to define a set of vertices $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$. Then, determining the edge set \mathcal{E} of the proximity graph \mathcal{G} requires defining the neighbors of each vertex v_i according to its embedding \mathbf{p}_i using the Euclidean distance. Among many possible choices, we choose to consider the symmetric k -nearest neighbor graph. An undirected edge (v_i, v_j) is added between two vertices v_i and v_j if the distance between \mathbf{p}_i and \mathbf{p}_j is among the k smallest distances from either \mathbf{p}_i or \mathbf{p}_j to all the other data points. The construction of such a graph being computationally expensive for large point clouds, a k -dimensional tree is used [8] to speed up the k -nearest neighbor search.

IN CULTURAL HERITAGE, TWO MAIN APPROACHES ARE USED TO DIGITIZE A 3-D OBJECT.

GRAPH WEIGHTS

Once the graph has been created, it has to be weighted. If one wants to ignore the vertex similarities, the weight function w can be set to $w(v_i, v_j) = 1, \forall (v_i, v_j) \in \mathcal{E}$. To

account for the similarities between the graph signals associated with the vertices, it is possible to use similarity functions based on distances to define the edge weights. Given an initial function $\mathbf{f}^0: \mathcal{V} \rightarrow \mathbb{R}^m$, computing distances between vertices consists of comparing their features (as a function of \mathbf{f}^0). To this end, each vertex v_i is associated with a feature vector $\mathcal{P}(v_i) \in \mathbb{R}^q$. From this, a usual similarity measure is provided by the Gaussian kernel

$$w(v_i, v_j) = \exp\left(-\frac{\|\mathcal{P}(v_i) - \mathcal{P}(v_j)\|_2^2}{\sigma^2}\right). \quad (1)$$

Traditionally, one has simply $\mathcal{P}(v_i) = \mathbf{f}^0(v_i)$. However, in image processing an important feature vector is provided by image patches [9]. In [10], we have proposed a new definition of patches that can be used with any graph representation associated to meshes or point clouds. We detail the latter.

PATCH DEFINITION

Around each vertex we build a 2-D grid (the patch) describing the neighborhood. This grid is defined on the tangent plane of the point (i.e., the vertex). The patch is oriented and finally filled in with a weighted average of the graph-signal values in the local neighborhood. We detail these two steps next.

Orientation

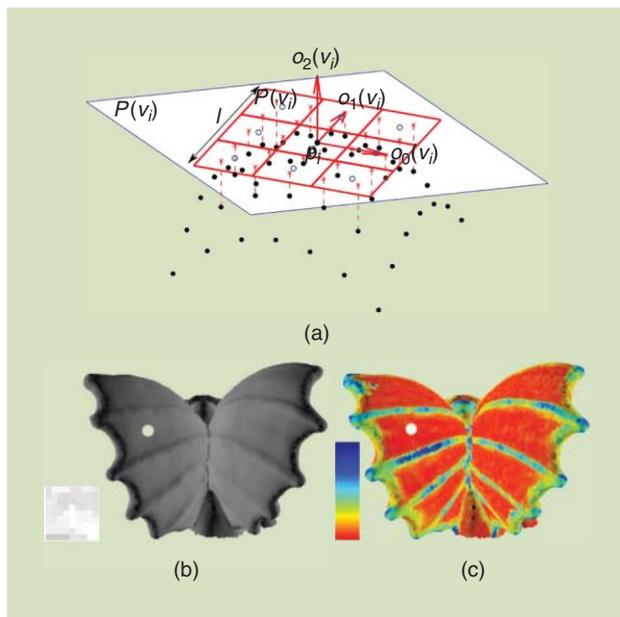
The first step consists of estimating the orientation of each patch. The algorithm first deduces a tangent vector $\mathbf{t}_1(v_i)$ from the normal vector $\mathbf{n}(v_i)$. We use a local principal component analysis on the coordinates p_i to estimate this normal vector: $\mathbf{t}_0(v_i) = \mathbf{n}(v_i)$. Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be the three axes of a 3-D space, the first tangent vector $\mathbf{t}_1(v_i)$ is computed with

$$\begin{cases} \mathbf{t}_1(v_i) = \mathbf{z} \times \mathbf{t}_0(v_i) & \text{if } |\mathbf{z} \cdot \mathbf{t}_0(v_i)| \neq 1 \\ \mathbf{t}_1(v_i) = \mathbf{x} \times \mathbf{t}_0(v_i) & \text{otherwise} \end{cases} \quad (2)$$

with \times the cross product and \cdot the dot product. The condition on the first line of (2) checks if the \mathbf{t}_1 and \mathbf{z} axis vector are collinear. If they are not collinear, \mathbf{t}_1 is calculated from the \mathbf{z} axis, otherwise the \mathbf{x} axis is used. Then a bitangent vector $\mathbf{t}_2(v_i)$ is computed by $\mathbf{t}_2(v_i) = \mathbf{t}_0(v_i) \times \mathbf{t}_1(v_i)$. The orientation vectors $\mathbf{o}_0(v_i), \mathbf{o}_1(v_i), \mathbf{o}_2(v_i)$ are then respectively assigned to $\mathbf{t}_1(v_i), \mathbf{t}_2(v_i), \mathbf{t}_0(v_i)$.

Patch Construction

The second step consists of constructing the patches. Given a point \mathbf{p}_i , a patch is defined for this point by constructing a



[FIG3] (a) The interpolation of the content of the patch. l is the patch length. $\mathbf{o}_0(v_i)$ and $\mathbf{o}_1(v_i)$ are the orientation of the patch $\mathcal{P}(v_i)$ at a point \mathbf{p}_i . Elements marked by “ \times ” symbol correspond to the projected neighbors \mathbf{p}'_j of the point \mathbf{p}_i on the patch. These projections are used to deduce values of each patch cell (a “ o ” symbol) by a weighted average of the associated graph signal values. (b) A point cloud with a selected vertex (in white), and the patch descriptor of that vertex. (c) A point cloud colored by the patch-based distance between all points and a given selected one, from most similar (in red) to least similar (in blue).

square grid around \mathbf{p}_i on its tangent plane in the orientation of the patch defined by $(\mathbf{o}_0, \mathbf{o}_1)$. We fix the patch length l manually. Let n be the number of cells on a row of the patch. A square lattice of n^2 cells is constructed around \mathbf{p}_i with respect to the basis obtained from the orientation computation. Each cell has a side length of l/n . A local graph is then considered that connects the vertex v_i to all the vertices v_j contained in a sphere of diameter $l\sqrt{2}$. Then, all the neighbors v_j of v_i are projected on the tangent plane of \mathbf{p}_i giving rise to projected points \mathbf{p}'_j . To fill the patch with values, these projected points \mathbf{p}'_j are associated with the cells with the closest center. The value of the cell is then deduced from a weighted average of the values $f^0(v_j)$ associated with the vertices v_j that were projected onto the patch cell. This value is a spectral value (the point's colors). The set of values inside the patch of the vertex v_i are denoted $\mathcal{P}(v_i)$. Let $C_k(v_i)$ denote the k th cell of the constructed patch around v_i with $k \in [1, n^2]$. With the proposed patch construction process, define the set $V_k(v_i) = \{v_j | \mathbf{p}'_j \in C_k(v_i)\}$ as the set of vertices v_j that were associated with the k th patch cell of v_i . Then, the patch vector is defined as :

$$\mathcal{P}(v_i) = \left(\frac{\sum_{v_j \in V_k(v_i)} w_p(\mathbf{c}_k, \mathbf{p}_j) f^0(v_j)}{\sum_{v_j \in V_k(v_i)} w_p(\mathbf{c}_k, \mathbf{p}_j)} \right)_{k \in [1, n^2]}^T, \quad (3)$$

with $w_p(\mathbf{c}_k, \mathbf{p}_i) = \exp(-\|\mathbf{c}_k - \mathbf{p}_i\|_2^2 / \sigma^2)$ where the \mathbf{c}_k are the coordinate vectors of the k th patch cell center. This weighting takes into account the distribution of the points in the cells of the patch by computing their mean feature vector. Figure 3(a) summarizes the method of patch construction. Figure 3(b) and (c) shows that points with similar geometric configurations are close with respect to the patch distance. With the following (parameters are respectively $\sigma = 0.3$ and $n = 5$).

GRAPH SIGNAL PROCESSING

This section provides the mathematical definitions needed to understand the construction of weighted graphs from point clouds.

PdEs ON WEIGHTED GRAPHS

The definitions presented here provide the basis on which it is possible to translate PDEs on graphs into PdEs on graphs. Most of these definitions are from [6].

DEFINITIONS

A weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ consists of a finite set $\mathcal{V} = \{v_1, \dots, v_N\}$ of N vertices and a finite set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ of weighted edges. Assume \mathcal{G} to be undirected, with no self-loops and no multiple edges. Let (v_i, v_j) be the edge of \mathcal{E} that connects two vertices v_i and v_j of \mathcal{V} . Its weight, denoted by $w(v_i, v_j)$, represents the similarity between its vertices. Similarities are usually computed by using a positive symmetric function $w: \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+$ satisfying $w(v_i, v_j) = 0$ if $(v_i, v_j) \notin \mathcal{E}$. The notation $v_i \sim v_j$ is also used to denote two adjacent vertices. The degree of a vertex v_i is defined as $\delta_w(v_i) = \sum_{v_j \sim v_i} w(v_i, v_j)$. Let $\mathcal{H}(\mathcal{V})$ be the Hilbert space of real-valued functions defined on the vertices of a graph. A function $f \in \mathcal{H}(\mathcal{V})$ assigns a real value $f(v_i)$ to each vertex $v_i \in \mathcal{V}$. We define the internal border of a set $A \subset \mathcal{V}$ as $\partial^- A = \{v_i \in A | \exists v_j \sim v_i, v_j \in \partial A\}$.

Similarly, we define the space $\mathcal{H}(\mathcal{E})$ of functions that are defined on the set \mathcal{E} of edges. Given a function $f: \mathcal{V} \rightarrow \mathbb{R}$, its ℓ_p and ℓ_∞ norms are given by

$$\|f\|_p = \left(\sum_{v_i \in \mathcal{V}} |f(v_i)|^p \right)^{1/p}, \quad \text{with } 1 \leq p < \infty, \quad (4)$$

$$\|f\|_\infty = \max_{v_i \in \mathcal{V}} |f(v_i)|, \quad \text{for } p = \infty. \quad (5)$$

DIFFERENCE OPERATORS ON WEIGHTED GRAPHS

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ be a weighted graph and $w: \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+$ a weight function that depends on the interactions between the vertices. The difference operator [6], denoted $d_w: \mathcal{H}(\mathcal{V}) \rightarrow \mathcal{H}(\mathcal{E})$, is defined for all $f \in \mathcal{H}(\mathcal{V})$ and $(v_i, v_j) \in \mathcal{E}$ by

$$(d_w f)(v_i, v_j) = \sqrt{w(v_i, v_j)} (f(v_j) - f(v_i)). \quad (6)$$

The weighted gradient operator of a function $f \in \mathcal{H}(\mathcal{V})$, at a vertex $v_i \in \mathcal{V}$, is the vector defined by

$$(\nabla_w f)(v_i) = ((d_w f)(v_i, v_j))_{v_j \in \mathcal{V}}^T. \quad (7)$$

The ℓ_p norm of this vector is defined, for $p \geq 1$, as

$$\|(\nabla_w f)(v_i)\|_p = \left(\sum_{v_j \sim v_i} w(v_i, v_j)^{p/2} |f(v_j) - f(v_i)|^p \right)^{1/p}. \quad (8)$$

The external and internal morphological directional difference operators are defined as in [11] to be $((d_w f)(v_i, v_j))^\pm$, with $(x)^+ = \max(x, 0)$ and $(x)^- = -\min(x, 0) = \max(-x, 0) = -(x)^+$. The corresponding discrete upwind weighted gradients are

$$(\nabla_w^\pm f)(v_i) = \left(((d_w f)(v_i, v_j))^\pm \right)_{v_j \in \mathcal{V}}^T. \quad (9)$$

The ℓ_p and the ℓ_∞ norms of these gradients are

$$\|(\nabla_w^\pm f)(v_i)\|_p = \left[\sum_{v_j \sim v_i} w(v_i, v_j)^{p/2} \left((f(v_j) - f(v_i))^\pm \right)^p \right]^{1/p}, \quad (10)$$

$$\|(\nabla_w^\pm f)(v_i)\|_\infty = \max_{v_j \sim v_i} \sqrt{w(v_i, v_j)} (f(v_j) - f(v_i))^\pm. \quad (11)$$

p-LAPLACE OPERATORS ON WEIGHTED GRAPHS

The isotropic weighted p -Laplace operator, with $p \in [1, +\infty[$, at a vertex $v_i \in \mathcal{V}$ is defined on $\mathcal{H}(\mathcal{V})$ by [6] as

$$(\Delta_{w,p}^i f)(v_i) = \sum_{v_j \sim v_i} \psi_{w,p}^i(v_i, v_j) (f(v_i) - f(v_j)), \quad (12)$$

where

$$\psi_{w,p}^i(v_i, v_j) = \frac{1}{2} w(v_i, v_j) (\|\nabla_w f(v_i)\|_2^{p-2} + \|\nabla_w f(v_j)\|_2^{p-2}). \quad (13)$$

The ∞ -Laplacian is defined by [12] as

$$(\Delta_{w,\infty} f)(v_i) = \frac{1}{2} \left[\|(\nabla_w^+ f)(v_i)\|_\infty - \|(\nabla_w^- f)(v_i)\|_\infty \right]. \quad (14)$$

From the definitions of these discrete difference operators on graphs, we are now in position to translate any PDEs that involves gradient, p -Laplacian or ∞ -Laplacian in their continuous formulation onto Euclidean domains. In the sequel, we will consider directly the discrete formulation on graphs; see [6] for further details.

REGULARIZATION ON GRAPHS

Here we present some PDEs on graphs and show a methodology to regularize the functions defined on the vertices of graphs. Let $f^0 \in \mathcal{H}(\mathcal{V})$ be a given function defined on the vertices of a weighted graph $G = (\mathcal{V}, \mathcal{E}, w)$. In a given context, this function represents an observation of a clean function $h \in \mathcal{H}(\mathcal{V})$ corrupted by an additive noise $n \in \mathcal{H}(\mathcal{V})$ such that $f^0 = h + n$. Regularizing functions on graphs using either isotropic or anisotropic p -Laplacian, was proposed in [6] and [13]. Recently, a new family of p -Laplace operators based on a divergence formulation, which unifies both the isotropic and anisotropic p -Laplacian on graphs, has been proposed in [10].

To recover the uncorrupted function h , the processing task is to remove the noise n from f^0 . A commonly used method is to seek a function $f \in \mathcal{H}(\mathcal{V})$, which is regular enough on G , and



[FIG4] The filtering of a noisy Mayan wall with a preservation of edges: (a) a noisy point cloud and (b) a denoised point cloud.

also close enough to f^0 . This can be formalized by the minimization of an energy functional, that involves a regularization term (or penalty term) plus an approximation one (or fitting term). This article considers the following model:

$$h \approx \underset{f: \mathcal{V} \rightarrow \mathbb{R}}{\operatorname{argmin}} J_{w,p}^\phi(f) + \frac{\lambda}{2} \|f - f^0\|_2^2 \quad (15)$$

$$\text{where } J_{w,p}^\phi(f) = \sum_{v_i \in \mathcal{V}} \phi(\|(\nabla_w f)(v_i)\|_p) \quad (16)$$

is a gradient-based functional, and $\lambda \in \mathbb{R}$ is a regularization parameter, called the Lagrange multiplier, that controls the tradeoff between the penalty term and the fitting term. The function $\phi(\cdot)$ is a positive convex function that penalizes large variations of f in the neighborhood of each vertex.

To solve (16), we consider the Euler-Lagrange equations

$$\frac{\partial J_{w,p}^\phi(f)}{\partial f(v_i)} + \lambda (f(v_i) - f^0(v_i)) = 0, \forall v_i \in \mathcal{V}, \quad (17)$$

where the first term denotes the variation of (16) with respect to f at a vertex v_i .

In [10], we have proven that the solution of (17) can be obtained with the following iterative algorithm:

$$f^{n+1}(v_i) = \frac{\lambda f^0(v_i) + \sum_{v_j \sim v_i} \beta_{v_i, v_j}^{\phi, p, f} f^n(v_j)}{\lambda + \sum_{v_j \sim v_i} \beta_{v_i, v_j}^{\phi, p, f}} \quad (18)$$

with

$$\beta_{v_i, v_j}^{\phi, p, f} = \alpha_{v_i, v_j}^{\phi, p, f} |f(v_j) - f(v_i)|^{p-2} \quad (19)$$

and

$$\alpha_{v_i, v_j}^{\phi, p, f} = w(v_i, v_j)^{p/2} \left(\frac{\phi'(\|(\nabla_w f)(v_i)\|_p)}{\|(\nabla_w f)(v_i)\|_p^{p-1}} + \frac{\phi'(\|(\nabla_w f)(v_j)\|_p)}{\|(\nabla_w f)(v_j)\|_p^{p-1}} \right). \quad (20)$$

Figure 4 shows such color filtering of a given 3-D colored point cloud (obtained from a laser scan of a Mayan temple wall).

APPLICATIONS

This section gives an overview of 3-D color point cloud applications. We illustrate the abilities of the proposed methods and algorithms

for signal processing on point clouds. (All of the 3-D point clouds we used are available at <https://lozes.users.greyc.fr/>.) The typical graph signals are point (respectively vertices) coordinates or colors. Given a weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ associated to a point cloud, consider an initial graph signal $\mathbf{f}^0: \mathcal{V} \rightarrow \mathbb{R}^m$ with $m = 3$. This signal will be either the vertices' coordinates, in which $\mathbf{f}^0(v_i) = \mathbf{p}_i$, or the vertices' colors, in which $\mathbf{f}^0(v_i) = (f^R(v_i), f^G(v_i), f^B(v_i))^T$, where $f^X(v_i)$ denotes the X color component of the color at the vertex v_i .

The approach we developed can be interesting within a cultural heritage documentation, analysis, and dissemination perspective (according to the evaluation done by art expert Livio De Luca, as mentioned in the "Acknowledgments" section). First,

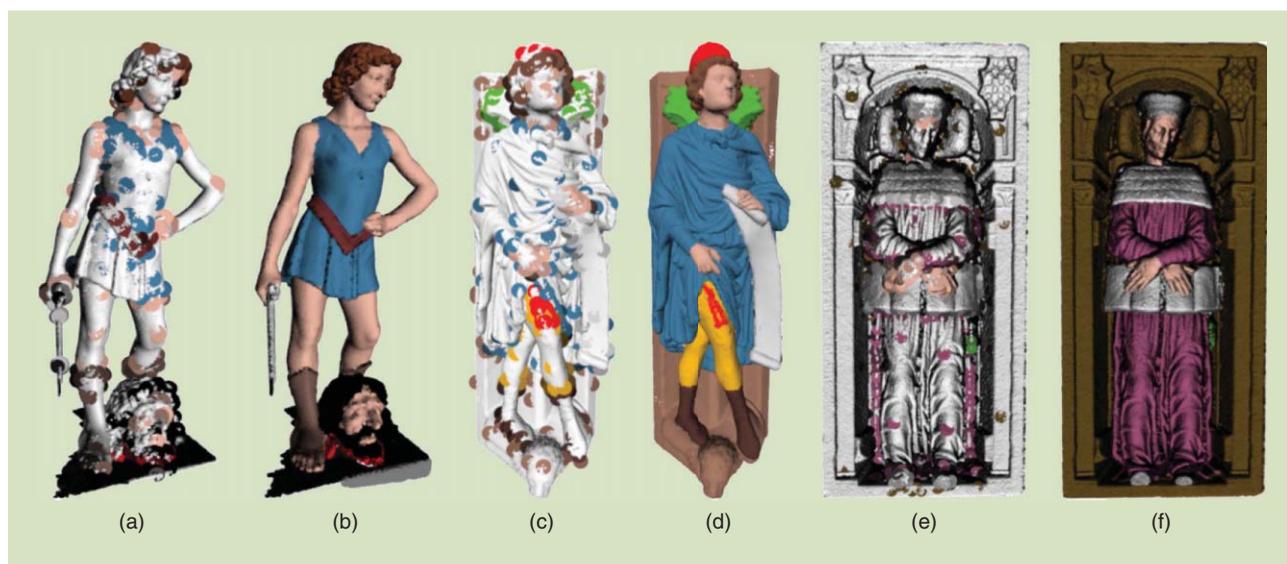
THE USE OF ADAPTIVE GEOMETRICAL WEIGHTS RELYING ON A GEOMETRIC FEATURE MAPS PATCHES ENABLES TO BETTER TAKE INTO ACCOUNT THE GEOMETRY OF THE POINT CLOUD DURING THE COLOR DIFFUSION AND BLENDING.

the idea to introduce a method for structuring color/texture information within a point-based 3-D representation is particularly relevant within the 3-D digitization and documentation purposes. Points clouds (especially if enriched by color information) include the essential geometric information useful for several applications scenarios

related to the heritage artifacts analysis (measurement, visualization, semantic annotation, etc.). The point clouds processing methods we suggest (colorization, filtering and simplification, and inpainting of missing data and segmentation) can find several effective applications within the digital driven process for documenting heritage buildings, archaeological sites, and museums' objects.



[FIG5] The colorization of the bishop castle (approximately 1.5 million points) and a Bas-relief (with 506,000 points): (a) the original point cloud with annotations, (b) the colored point cloud, (c) the colorless Bas-relief with annotations, and (d) the colored Bas-relief.



[FIG6] The colorization of some cultural heritage statues [sizes of point clouds: (a) and (b) 253,000, (c) and (d) 254,000, (e) and (f) 791,000 points]. (a) Scribbles, (b) result, (c) scribbles, (d) result, (e) scribbles, and (f) result.

COLORIZATION

Colorization is the process of adding colors on an uncolored object. Let $\mathbf{f}^0: \mathcal{V} \rightarrow \mathcal{C}$ be a function that assigns colors to vertices. Let $A \subset \mathcal{V}$ be the subset of vertices with unknown colors and ∂A the subset of vertices with known values. The purpose of interpolation is to find a function \hat{f} approximating f^0 in \mathcal{V} minimizing the following isotropic total variation energy for $v_i \in A$:

$$\min_{f \in H^1(\mathcal{V})} \left\{ R_{w,p}(f) = \sum_{v_i \in A} \|\nabla_w f(v_i)\|_p^p \right\} \quad (21)$$

with $f(v_i) = f^0(v_i)$, for $v_i \in \partial A$. This can again be solved using the Euler–Lagrange equations

$$\begin{cases} (\Delta_{w,p}^i f)(v_i) = 0 & \forall v_i \in A, \\ f(v_i) = f^0(v_i) & \forall v_i \in \partial A. \end{cases} \quad (22)$$

The solution of (22) can be obtained with the iterative algorithm in (18) with $\lambda = 0$, $\forall v_i \in \partial^- A$. The similarity function $w: \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ is computed from the comparison of patches of a geometric feature. This geometric feature is computed from the degree at a vertex using local height weights. These later weights are obtained from the similarity of height patches. Since the colorization starts from initial color annotations, not all the points are considered simultaneously. The colorization starts from the points that are neighbors to the annotated colors and the set of points to be colorized grows as the algorithm iterates. The colorization process stops when the set of vertices to colorize is empty and has converged to a stable solution. The use of adaptive geometrical weights relying on a geometric feature maps patches enables to better take into account the geometry of the point cloud during the color diffusion and blending.

Figure 5 shows the colorization of the bishop castle and of a Bas-relief (a type of sculpture in which shapes are carved). Figure 6 shows the colorization on some uncolored statues. Both cases show that the colorization allows restoration of colors not captured by the 3-D scanner (the bishop castle case), but also allows placement of colors on an initially uncolored object (statues cases) for a more realistic rendering.

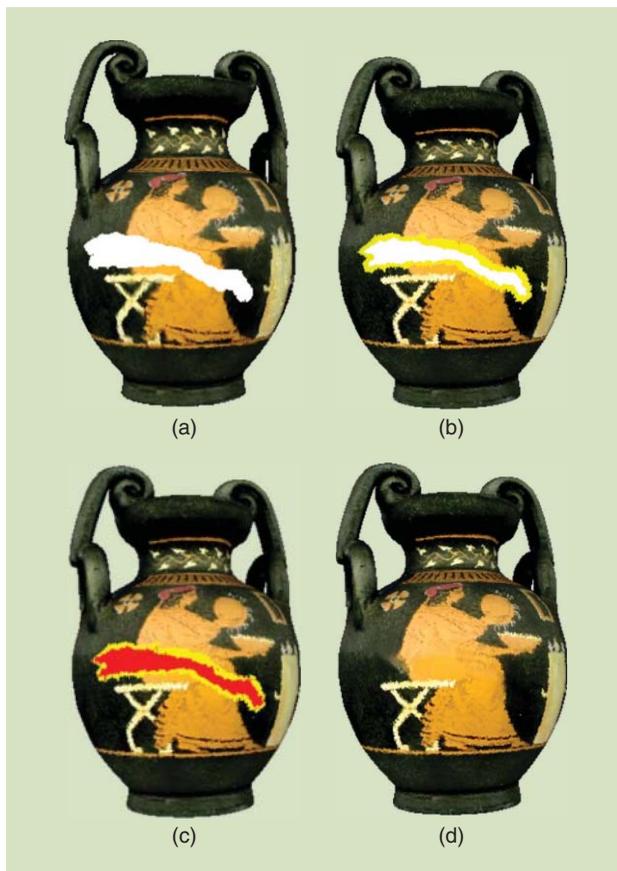
FILTERING AND SIMPLIFICATION

Modern 3-D scanners can generate large point clouds with several million or billion points. The processing of such data is difficult. Instead of downscaling the point clouds and losing details, we propose a way to remove fine details without modify the appearance of the point cloud. It can be interesting to filter and simplify complex and redundant point clouds for Web publication purposes.

Let $\mathbf{f}^0: \mathcal{V} \rightarrow \mathcal{P}$ be the vertices coordinates. The iterative algorithm of (18) allows filtering the geometry of 3-D objects using $\lambda = 0$. Figure 7 shows the simplification of a point cloud of the Saint Eligius statue with a reduction in the number of points by 77.6% using parameters $p = 0.1$ and $\phi(s) = s^p$. With such a simplified point cloud, the processing of any data on this point cloud, like colors, becomes computationally feasible. Indeed, reducing the quantity of data to process makes algorithms and renderings faster. Finally, the simplification could be leveraged in



[FIG7] The simplification of a point cloud of the Saint Eligius statue after 10,000 iterations. (a) The original point cloud (201,000 points) and (b) the point cloud simplified (45,000 points).



[FIG8] The restoration pipeline of a vase. The geometric part is first filled, then the missing color is inpainted: (a) the denoised vase, (b) the hole to fill in (boundary in yellow), (c) the filled-in hole, and (d) the inpainted hole.

a meshing process by producing more regular polygons.

INPAINTING OF MISSING DATA

Inpainting consists of constructing new values for missing data in coherence with a set of known data. This can be the reconstruction of deteriorated parts of a 3-D object represented as a point cloud. This can be of interest for extending color information to missing parts (e.g., for virtual restoration purposes). Recent inpainting work tends to unify local and nonlocal approaches under a variational formulation (see [14] and references therein for more details). We consider that data are defined on a general domain represented on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$. Let $f^0: \mathcal{V} \rightarrow \mathcal{C}$ be a function that assigns colors to vertices. Let $\mathcal{A} \subset \mathcal{V}$ be the subset of vertices with unknown values and $\partial\mathcal{A}$ the subset of vertices with known values. The purpose of the interpolation is to find a function \hat{f} approximating f^0 in \mathcal{V} and that corresponds to solving

$$\begin{cases} (\Delta_{w, \infty} f)(v_i) = 0 & \forall v_i \in \mathcal{A}, \\ f(v_i) = f^0(v_i) & \forall v_i \in \partial\mathcal{A}. \end{cases} \quad (23)$$

INPAINTING CONSISTS OF CONSTRUCTING NEW VALUES FOR MISSING DATA IN COHERENCE WITH A SET OF KNOWN DATA.

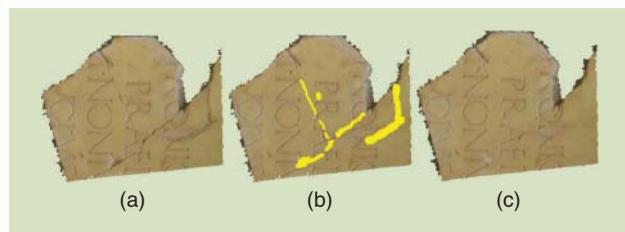
The infinity Laplacian is used here for interpolation since it provides better inpainting results than the isotropic p -Laplacian interpolation (see [15]). Works in [10] have shown that this interpolation problem has a unique solution that can

be obtained using the iterative algorithm presented in [12]. At the end of each iteration the set $\partial\mathcal{A}$ is updated by $\partial\mathcal{A}^{(n+1)} = \partial\mathcal{A}^{(n)} \cup \partial^- \mathcal{A}^{(n)}$, and $\partial^- \mathcal{A}^{(n+1)}$ is updated from $\partial\mathcal{A}^{(n+1)}$. The algorithm stops when the set of vertices to inpaint is empty. Figure 8 shows the restoration of a broken vase. The geometric part is corrected first, then the hole is filled; finally, the color is inpainted. One major objective of the cultural heritage is the preservation of an object. This inpainting algorithm provides a plausible virtual reconstruction of the original state of an object. The algorithm can also be used to remove and subvert parts of a damaged object, as in Figure 9.

SEGMENTATION

Segmentation is the process of partitioning a 3-D object into multiple regions. Let $f: \mathcal{V} \rightarrow \mathbb{R}$ be the function on a graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, w)$ to segment. The segmentation problem can be formulated as a PDE to be solved on weighted graphs. This latter formulation is based on front propagation using the eikonal equation to compute general distances on graphs. Let $L = \{l_1, \dots, l_n\}$ be the set of labels and let $S^0 = S_1^0 \cup \dots \cup S_n^0$ be the corresponding set of labeled vertices. The goal of label propagation is to label each vertex $u \in \mathcal{V}$ under the condition that u is a neighbor of S_i^0 . Such a label propagation can be formulated as solving the eikonal equation on a graph \mathcal{G} as

$$\begin{cases} \|(\nabla_w^- \phi)(v_i, t)\|_p = P_i(x) & v_i \in \mathcal{V} \\ f(v_i) = \phi(v_i) & x \in \gamma \subset \mathcal{V}, \end{cases} \quad (24)$$



[FIG9] The inpainting of a part of a damaged wall: (a) the object to inpaint, (b) the inpainting marker, and (c) the inpainting result.



[FIG10] Segmentation of a vase by resolving the eikonal equation on graphs [770,000 points for (a) and 200,000 points for (d)]: (a) a vase with labels, (b) the result, (c) a vase with labels, and (d) the result.

where $F_i: \mathcal{V} \rightarrow \mathbb{R}^3$ is a speed function, $P_i(v_i) = (1/F_i(v_i))$, and $\phi: \mathcal{V} \rightarrow \mathbb{R}$ represent the set of initials labels. The resolution of the eikonal equation (24) on a weighted graph has been investigated in [16].

To segment a mesh or a point cloud according to the vertices colors, compute the similarity function $w: \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+$ by taking into account the colorimetric distance as $w = e^{-\|f^0(v_i) - f^0(v_j)\|/\sigma^2}$, where $f^0: \mathcal{V} \rightarrow \mathcal{C}$ are the colors associated at the node $v_i \in \mathcal{V}$. The function $\phi: \mathcal{V} \rightarrow \mathbb{R}$ associates an initial label to each vertex v_i . The parameters are set as follows: $p = 2$, $F_i(v_i) = 1$, $\forall v_i \in \mathcal{V}$. Figure 10 shows some segmentation results on point clouds after the resolution of the eikonal equation. Segmentation results are 3-D subsets of the original points clouds. These subsets can then be processed by applying the previous graph signal techniques such as filtering, simplification, inpainting, and colorization.

CONCLUSIONS

This article has proposed an approach for the processing of functions on point clouds represented as graphs. We have shown how to translate PDEs using the framework of PdEs. The approach allows processing of signal data on point clouds (e.g., spectral data, colors, coordinates, and curvatures). We have applied this approach for cultural heritage purposes on examples aimed at restoration, denoising, hole-filling, inpainting, object extraction, and object colorization. The results demonstrate the many potentials of the point cloud approach to the processing of cultural heritage 3-D scanned objects.

ACKNOWLEDGMENTS

Livio De Luca is the director of the Models and Simulations for Architecture and Cultural Heritage Laboratory, a French research unit of the National Centre for Scientific Research and the Ministry of Culture and Communication working on the application of informatics technologies for the study of architecture considered in its heritage, constructive, urban, and landscape dimensions. We thank him for his evaluation of the potential applications of these methods mentioned in the “Applications” section.

AUTHORS

François Lozes (francois.lozes@unicaen.fr) received his Ph.D. degree in computer science at the Université de Caen Basse-Normandie, France. His research mainly concerns the analysis of data on 3-D meshes and 3-D point clouds using graph signal processing.

Abderrahim Elmoataz (abderrahim.elmoataz-billah@unicaen.fr) is a full professor of computer science at the Université de Caen Basse-Normandie, France. His research concerns partial differential equations (PDEs) on 3-D surfaces and points clouds, PDEs on graphs of arbitrary topology with applications in image processing, and machine learning.

Olivier Lézoray (olivier.lezoray@unicaen.fr) received the M.Sc. and doctoral degrees in computer science from the Université de Caen Basse-Normandie, France, in 1996 and 2000, respectively. From 1999 to 2000, he was an assistant professor in the Computer Science Department at the same university. From 2000 to 2009, he was an associate professor at the Cherbourg Institute of Technology in the Communication Networks and Services Department. Since 2010, he has been a full professor at the Cherbourg Institute of Technology. His research is focused on color image segmentation and filtering (graph-based variational and morphological methods) and machine-learning techniques for image mining (neural networks and support vector machines).

REFERENCES

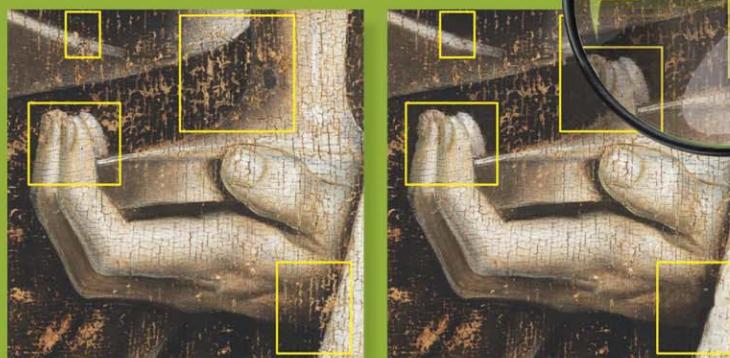
- [1] L. J. Van der Maaten and E. O. Postma, “Texton-based analysis of paintings,” in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2010, pp. 77980H–77980H.
- [2] P. Abyr, S. Jaffard, and H. Wendt, “When van Gogh meets Mandelbrot: Multifractal classification of painting’s texture,” *Signal Process.*, vol. 93, no. 3, pp. 554–572, 2013.
- [3] M. Hess and S. Robson, “3D colour imaging for cultural heritage artefacts,” in *ISPRS*, vol. 5, 2010, pp. 288–292.
- [4] *Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks*, (Digital Imaging and Computer Vision Series), F. Stanco, S. Battiato, and G. Gallo, Eds. New York: Taylor and Francis, 2011.
- [5] B. Cornelis, T. Ružić, E. Gezels, A. Dooms, A. Pižurica, L. Platiša, J. Cornelis, M. Martens, M. D. Mey, and I. Daubechies, “Crack detection and in painting for virtual restoration of paintings: The case of the Ghent Altarpiece,” *Signal Process.*, vol. 93, no. 3, pp. 605–619, 2013.
- [6] A. Elmoataz, O. Lezoray, and S. Bougleux, “Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing,” *IEEE Trans. Image Processing*, vol. 17, no. 7, pp. 1047–1060, 2008.
- [7] D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Processing Mag.*, vol. 30, no. 3, pp. 83–98, 2013.
- [8] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, “An optimal algorithm for approximate nearest neighbor searching fixed dimensions,” *J. ACM*, vol. 45, no. 6, pp. 891–923, Nov. 1998.
- [9] A. Buades, B. Coll, and J.-M. Morel, “Image denoising methods. A new nonlocal principle,” *SIAM Rev.*, vol. 52, no. 1, pp. 113–147, 2010.
- [10] F. Lozes, A. Elmoataz, and O. Lezoray, “Partial difference operators on weighted graphs for image processing on surfaces and point clouds,” *IEEE Trans. Image Processing*, vol. 23, no. 9, pp. 3896–3909, Sept. 2014.
- [11] V.-T. Ta, A. Elmoataz, and O. Lézoray, “Nonlocal PDEs-based morphology on weighted graphs for image and data processing,” *IEEE Trans. Image Processing*, vol. 26, no. 2, pp. 1504–1516, June 2011.
- [12] A. Elmoataz, X. Desquesnes, and O. Lezoray, “Non-local morphological pdes and p-Laplacian equation on graphs with applications in image processing and machine learning,” *IEEE J. Sel. Top. Signal.*, vol. 6, no. 7, pp. 764–779, 2012.
- [13] O. Lezoray, V.-T. Ta, and A. Elmoataz, “Partial differences as tools for filtering data on graphs,” *Pattern Recogn. Lett.*, vol. 31, no. 14, pp. 2201–2213, 2010.
- [14] P. Arias, G. Facciolo, V. Caselles, and G. Sapiro, “A variational framework for exemplar-based image inpainting,” *Int. J. Comput. Vision*, vol. 93, no. 3, pp. 319–347, 2011.
- [15] M. Ghoniem, A. Elmoataz, and O. Lézoray, “Discrete infinity harmonic functions: towards a unified interpolation framework on graphs,” in *Proc. IEEE Int. Conf. Image Processing*, 2011, pp. 1361–1364.
- [16] X. Desquesnes, A. Elmoataz, and O. Lézoray, “Eikonal equation adaptation on weighted graphs: Fast geometric diffusion process for local and non-local image and data processing,” *J. Math. Imaging Vis.*, vol. 46, no. 2, pp. 238–257, 2013.



[Aleksandra Pižurica, Ljiljana Platiša, Tijana Ružić, Bruno Cornelis, Ann Dooms, Maximiliaan Martens, Hélène Dubois, Bart Devolder, Marc De Mey, and Ingrid Daubechies]

Digital Image Processing of *The Ghent Altarpiece*

© GHENT, KATHEDRALE KERKFABRIEK; PHOTO COURTESY OF KIK-IRPA, BRUSSELS. MAGNIFYING GLASS—IMAGE LICENSED BY INGRAM PUBLISHING.



Signal Processing for Art Investigation

[Supporting the painting's study and conservation treatment]

Hanging in the Saint Bavo Cathedral in Ghent, Belgium, is *The Ghent Altarpiece*, also known as *The Adoration of the Mystic Lamb* (see Figure 1). According to an inscription on the outer frames, it was painted by brothers Hubert and Jan van Eyck for Joos Vijd and his wife Elisabeth Borluut in 1432. It is one of the most admired and influential paintings in the history of art and has given rise to many intriguing questions that have been puzzling art historians to date [11]. Moreover, the material history of the panels

is very complicated. They were hidden, dismantled, moved away, stolen, and recovered during riots, fires and wars. The recovery of the panels by the U.S. Army in the Nazi hoards deep in the Altaussee salt mines has particularly marked memories. One panel was stolen in 1934 and never recovered. Besides varying conservation conditions, the panels underwent numerous restoration treatments and were even partially painted over.

One of the most important unresolved questions related to this painting goes back to its creation: the division of hands between the two brothers and their respective workshops. The meticulous study of the painting technique, its different layers and materials, as well as the underdrawings and perhaps even numerous intriguing palm

Digital Object Identifier 10.1109/MSP.2015.2411753

Date of publication: 15 June 2015



[FIG1] *The Ghent Altarpiece*: (a) open and (b) closed. Image copyright Ghent, Kathedrale Kerkfabriek, Luksaweb.

and fingerprints could bring us closer to answering that question. However, *The Ghent Altarpiece* hides many other secrets, like the meaning of inscriptions that are difficult to decipher, such as the text in the book depicted in the panel *Virgin Annunciate*.

A major conservation and restoration campaign carried out by the Royal Institute for Cultural Heritage (KIK-IRPA), which is expected to take at least six years, started in October 2012. One of the questions of the treatment, supported by an international commission of experts, concerns uncovering van Eyck's original paint to the extent which can be safely carried out. Indeed, the paintings were covered over centuries with disfiguring retouchings, overpainting, and varnishes. Certain decisions regarding the restorations benefit from multidisciplinary research and signal processing could help in this regard.

In this article, we show progress in certain image processing techniques that can support the physical restoration of the painting, its art-historical analysis, or both. We first introduce a multi-modal crack detection algorithm, which gives a clear improvement over earlier reported crack detection results on *The Ghent Altarpiece*. We then show how a relatively simple analysis of the crack patterns could indicate possible areas of overpaint, which may be of great value for the physical restoration campaign, after further validation. Next, we explore how digital image inpainting can serve as a simulation for the restoration of losses (missing areas in one or more layers of the painting, often caused by abrasion or mechanical fracture and revealed after the cleaning process). As a separate problem, we address crack inpainting by outlining the main challenges and proposing a solution that improves upon earlier reported results on this painting [7]. Finally, we explore how the statistical analysis of the relatively simple and frequently recurring objects (such as pearls in this masterpiece) may characterize the consistency of the painter's style and thereby aid both art-historical

interpretation and physical restoration campaign. We carry out our analysis on a recently released high-resolution data set and on some images taken during the current treatment of the altarpiece.

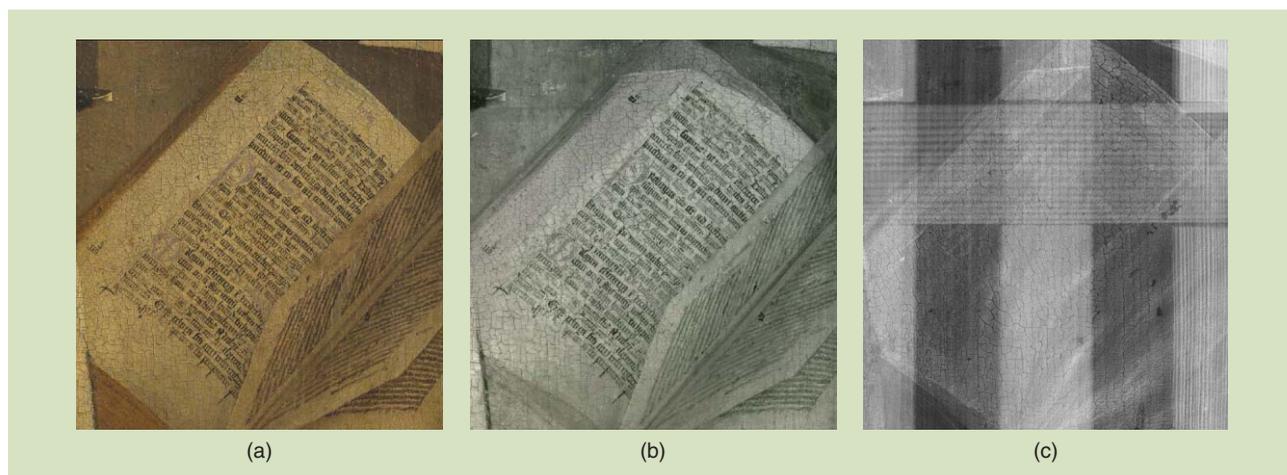
DATA SET: "CLOSER TO VAN EYCK"

Until 2012, digitized scans of old photographic negatives, acquired by Alfons Dierick [12] and kept in the archives of Ghent University, were the only available high-resolution data set of *The Ghent Altarpiece*. The development process of these negatives was mainly undocumented, which resulted in a data set where the images vary strongly in quality. Earlier reported results of digital image processing on *The Ghent Altarpiece*, such as crack detection, virtual crack inpainting [7], [24], and pearl analysis [23], were all based on images from that old data set.

We report the results on extremely high-resolution images that are publicly available in [31]. This data set is the result of an interdisciplinary research project that ran from April 2010 until June 2011, with the goal to assess the structural condition of *The Ghent Altarpiece* and determine whether a full restoration of van Eyck's polyptych was necessary. The surfaces of the altarpiece were documented with the following imaging modalities: digital macrophotography (with a pixel size of $7.2 \mu\text{m}$; full panels, 140 extreme close-ups, and some cleaning tests), infrared macrophotography (in the same resolution), infrared reflectography, and X-radiography. New acquisitions will be added to this data set in the scope of the current conservation-restoration campaign.

IMAGE PROCESSING IN SERVICE OF PAINTING RESTORATION

We address two potential applications of image processing to support restoration of paintings: 1) detecting possible areas of overpaint based on the analysis of crack patterns and 2) virtually inpainting losses.



[FIG2] Acquisitions of *The Ghent Altarpiece*: (a) macrophotography, (b) infrared macrophotography, and (c) X-ray radiography. Images used with permission from [31]. Image copyright Ghent, Kathedrale Kerkfabriek, Luksaweb.

CRACK DETECTION IN PAINTINGS

Being able to accurately detect cracks can be very relevant to painting conservation, since cracking is one of the most common forms of deterioration. Fluctuations in humidity, causing the wooden support to shrink or expand, is the main reason for crack formation. Because the way in which cracks develop and spread partly depends upon the choice of materials and methods used by the artist, assessing cracks is useful for judging authenticity [4]. Cracks can also assist conservators by providing clues to the causes of degradation of the paint surface. An in-depth study of the factors contributing to their formation can support preventive measures [1]. Furthermore, the analysis of crack patterns provides noninvasive means of identifying the structural components of paintings [4].

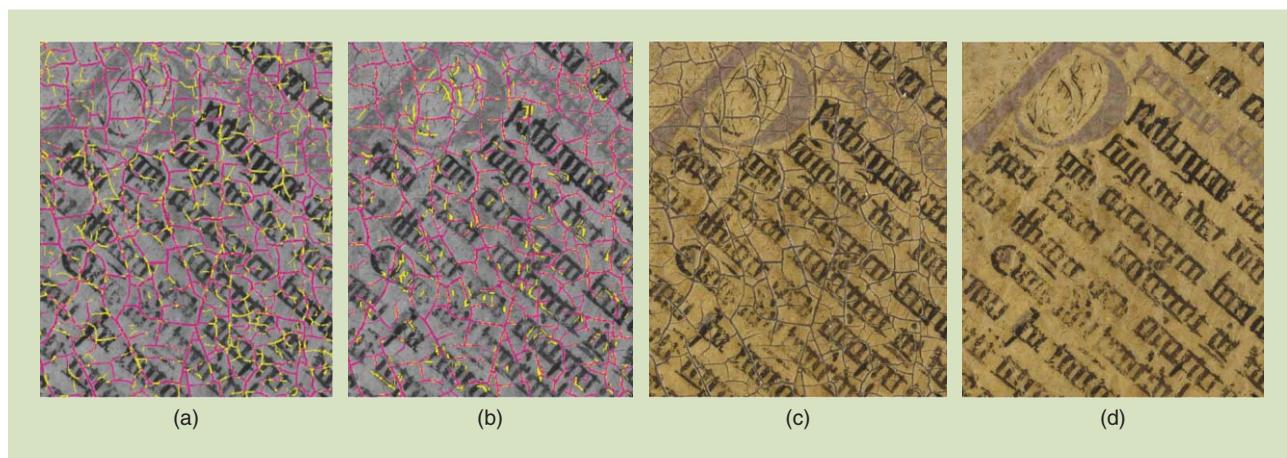
Visually, cracks can be categorized into bright cracks on a dark background or dark cracks on a bright background. One can further distinguish between different types of cracks such as ageing cracks, premature cracking (generally due to drying defects related to the painting materials or their application), or cracks formed only in the varnish layer when it becomes brittle through oxidation. The literature discusses mainly dark cracks; they are typically considered as having low luminance and being local (gray scale) intensity minima with elongated structure [14]. Different crack-detection techniques include simple thresholding, line detectors, and various morphological filters (see [1] for an overview). The method in [7] operates on a single image modality (visible image) and combines by means of a voting scheme three-crack detection techniques: oriented elongated filters, a multiscale extension of the morphological top-hat transformation, and a detection method based on dictionary learning [13].

THE BAYESIAN CONDITIONAL TENSOR FACTORIZATION METHOD FOR MULTIMODAL DATA

The newly acquired multimodal data set (see Figure 2 for an example) allows for new crack detection techniques that are able to make use of the information provided by each modality, thereby yielding a more reliable detection scheme. However, a pixel-perfect registration

is required prior to using all modalities together. The panels of *The Ghent Altarpiece* were already roughly registered for adjacent viewing in [31], but the spatial alignment of these preregistered images is not sufficient in the current context, as the images can be shifted by a few pixels or even exhibit local inconsistencies due to the different acquisition modalities. The nature of the different modalities and the stringent requirements for crack detection make direct registration a challenging task. However, since the cracks themselves are a more or less consistent component throughout all modalities, we used them for the registration process. Crude crack maps are first obtained by filtering the unregistered images with elongated filters and subsequent thresholding (more details of the exact procedure can be found in [7]). It should be noted at this point that the presence of false positives is not a nuisance as long as the locations of most of the cracks in each modality are identified. The crude crack maps obtained from the X-radiograph and the visual and infrared images are mutually registered using the algorithm described in [5] (using the infrared crack map as the reference). The resulting transformation is then applied to the original images.

Simply applying the methods described in [7], which were designed for a single image, requires choosing an additional set of parameters per modality, which would be too cumbersome. Here we adopt a semisupervised Bayesian approach that estimates for each pixel a posterior probability of belonging to the “crack” category given a large set of feature vectors extracted over all modalities. These feature vectors are obtained by processing each image modality with a number of different filters, commonly used in image processing, ranging from morphological filters to multiorientation filter banks, as described in [8]. The resulting feature vectors, hereafter denoted as *categorical predictors*, or *predictors* for short, are quantized into an experimentally chosen number of bins. Let X_1, \dots, X_p denote p predictors at a given pixel location, and let Y denote a hidden random variable, taking values $y \in \{0, 1\}$, where the label “1” denotes a crack pixel and “0” a noncrack pixel. The conditional probability $P(Y|X_1, \dots, X_p)$ is a $d_1 \times \dots \times d_p$ dimensional tensor, with d_j the number of quantization bins of the j th predictor X_j .



[FIG3] Crack detection and inpainting. (a) and (b) show a comparison between BCTF and the method in [7], where cracks detected by both methods are marked in red and the differences in yellow. (a) Yellow indicates cracks detected only by BCTF. (b) Yellow indicates cracks detected only by [7]. (c) and (d) show the original visible data and the result of our patch-based inpainting after BCTF-based crack detection.

Using the Bayesian conditional tensor factorization (BCTF) of [30], inspired by higher-order singular value decomposition [10], the conditional probability tensor can be decomposed as

$$P(Y = y | X_1 = x_1, \dots, X_p = x_p) = \sum_{j_1=1}^{k_1} \dots \sum_{j_p=1}^{k_p} \lambda_{j_1, j_2, \dots, j_p}(y) \prod_{m=1}^p \pi_{j_m}^{(m)}(x_m), \quad (1)$$

with the parameters subject to

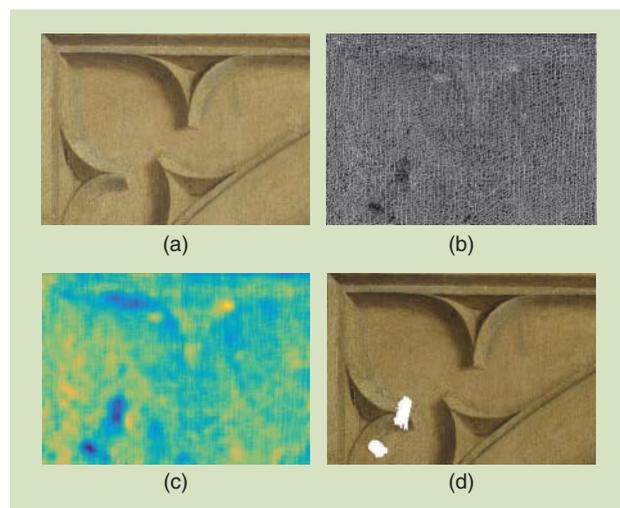
$$\sum_{j_m=1}^{k_m} \pi_{j_m}^{(m)}(x_m) = 1, \quad (2)$$

for every combination of (m, x_m) . The factorization coefficients $\lambda_{j_1, j_2, \dots, j_p}(y)$ can be seen as the latent class allocation probabilities and $\pi_{j_m}^{(m)}(x_m)$ as the response class probabilities, which control in a probabilistic manner how the levels of each predictor are clustered. The $k_m \in \{1, \dots, d_m\}$ value impacts the number of parameters used to characterize the m th predictor. In the special case where $k_m = 1$, (2) yields $\pi_1^{(m)}(x_m) = 1$, which means that $P(y | x_1, \dots, x_p)$ will not depend on x_m and the m th predictor can be excluded from the model. If $k_m = 1$ for most m s, the categorical predictor model becomes sparse. In practice, we do expect that only a few features have a significant impact on the classification results. More details on the exact posterior computation can be found in [30]. The resulting conditional probability tensor can be used as a lookup table where each entry contains a crack probability for a specific combination of predictor values. If this probability exceeds 0.5, we label the pixel as being part of a crack.

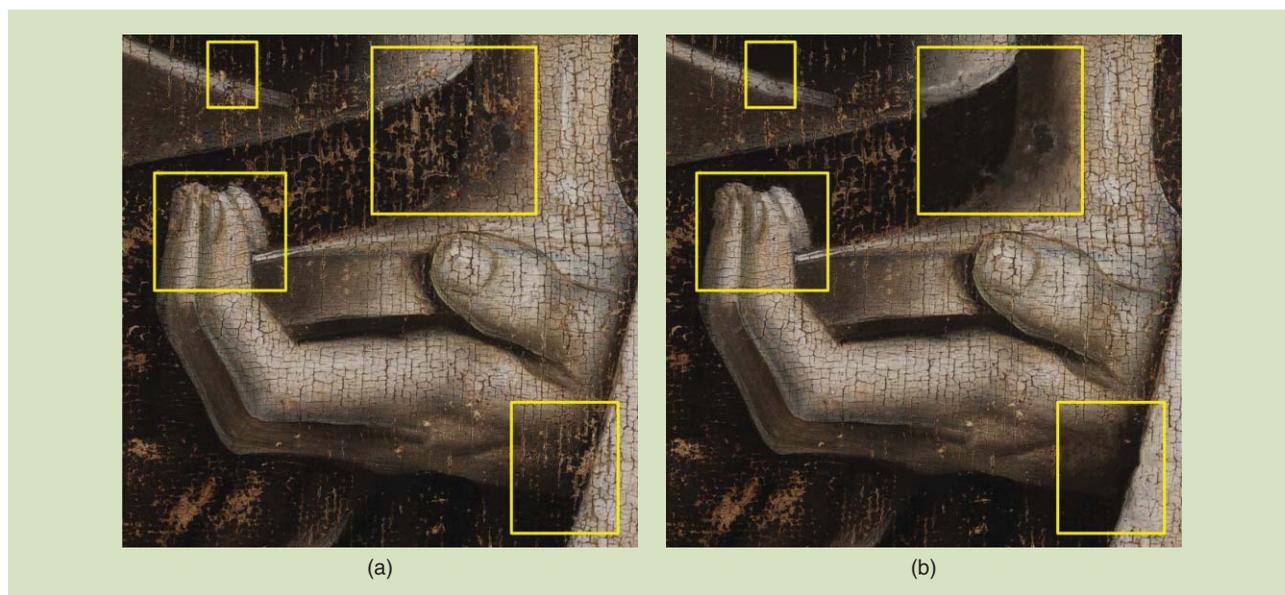
A comparison between the multimodal BCTF method and the crack detection method introduced earlier in [7] on the same part of the painting is depicted in Figure 3. It can be observed that the older method fails to detect some thin cracks, while it falsely labels some thin dark brushstrokes as cracks. It is clear that the multimodal BCTF method detects more cracks with fewer false positives.

IDENTIFYING OVERPAINT FROM CRACK PATTERNS

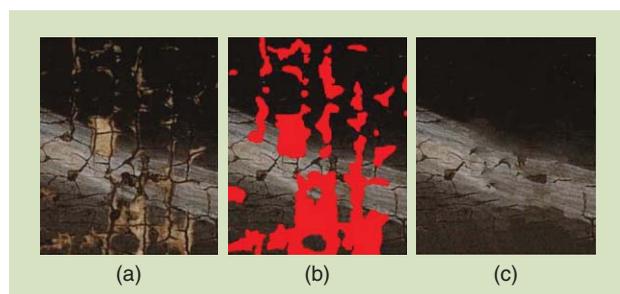
Some features of the detected crack patterns may have potential to guide the restorers to places of interest such as retouchings or heavily damaged areas. As an example, we applied the BCTF method, described previously, on the upper left corner of the *Joos Vijd* panel. Figure 4(a) shows that part of the painting and its corresponding crack map. A rather simple analysis consists of counting the number of crack pixels in a sliding window of 100×100 . In doing so we obtain a crack density map (see Figure 4) where we can identify low crack density zones as well as high crack density zones. The lowest (and highest) crack density zones are obtained automatically by hysteresis thresholding, where the first chosen threshold is chosen to be very close to the minimum (respectively maximum) value of the crack density. Painting conservators confirmed that the areas of lowest density marked in white in Figure 4(d) are old losses covered with retouching that also overlap on the surrounding original paint. Other zones of low crack density however,



[FIG4] A crack density analysis within the upper left corner of the *Joos Vijd* panel. (a) The high-resolution macrophotograph. (b) A detected crack map. (c) Crack density map (blue: low density; red: high density). (d) Detected zones of lowest crack density.



[FIG5] (a) A part of the cleaned *John the Evangelist* panel. The marked regions are inpainted in (b). Image copyright Ghent, Kathedrale Kerkfabriek; photo courtesy of KIK-IRPA, Brussels.



[FIG6] The results on a small part of the cleaned *John the Evangelist* panel. (a) The original image, (b) overlay with damaged regions marked in red, and (c) the result of the algorithm from [25]. Image copyright Ghent Kathedrale Kerkfabriek; photo courtesy of KIK-IRPA.

such as the ones seen in the upper left corner of the density map, correspond to thinner original paint that developed a different crack pattern over time. The interpretation of the crack maps, like any other diagnostic tool in art conservation, needs to be checked by conservators using other examination techniques and linking the evidence by their critical and material skills.

VIRTUAL INPAINTING

During the ongoing physical restoration of *The Ghent Altarpiece*, deteriorated retouching and overpaint are removed, revealing underlying losses in the original (see Figure 5). Digital image inpainting can virtually fill in these areas and provide a “simulation” for the impact of certain actions to be taken during the physical restoration process.

A recent overview of inpainting methods is given in [15], and applications to virtual restoration of paintings include [21] and [22]. Patch-based methods are capable of replicating both structure and texture by filling in the missing region patch-by-patch. In general, for each patch of the missing region (target patch), a well

matching replacement patch is found in the available part of the image (source region) and copied to the corresponding location. Preserving structures is achieved by defining the filling order [9], which gives priority to the target patches containing object boundaries and fewer missing pixels. The so-called global methods, like [17] and [25], allow the choice of multiple candidate patches (instead of choosing just one best match in a “greedy” manner) and define inpainting as a global optimization problem.

Figure 5 shows a part of the *John the Evangelist* panel that has been cleaned in the current restoration campaign of *The Ghent Altarpiece*. Removing overpaint revealed many losses that will be carefully inpainted by the conservators, using stable and reversible materials, to restore the visual coherence of the original image. In contrast to losses and abrasions, age cracks are not inpainted in actual conservation unless they severely interfere with the painted form. For this reason, the cracks are not inpainted intentionally in this experiment (unless they are inside a larger loss and assigned, therefore, to a missing region). Here we provide a virtual inpainting simulation of some parts obtained with the patch-based algorithm of [25] [see Figure 5 and the enlarged part in Figure 6(a)]. We have chosen to inpaint several figurative parts that contain structure and texture. The results show how challenging this problem is even for state-of-the-art inpainting methods. None of the methods tested so far produced a satisfying result that fully (albeit virtually) restored the painted form. Experienced conservators master a knowledge of the physical characteristics of the paint layers and of the painted forms that call upon complex visual perception and interpretation skills. Virtual inpaintings do not provide alternatives to their work on the original but test the potential of the methods that need to be further developed.

DOES CRACK FILLING HELP YOU READ THAT BOOK?

While cracks are rarely inpainted in the actual, physical restoration, virtual crack removal can be of interest in certain aspects.

For example, crack inpainting may improve the legibility of the text present in parts of the polyptych, which can be of great importance to art-historical and iconographical studies. Virtual inpainting of the book in the *Virgin Annunciate* panel was reported in [7] on the scans from the Dierick collection. Here we identify some limitations of state-of-the-art inpainting techniques for this type of problem and introduce an improved method and report the results on the new high-resolution scans from the data set in [31].

CHALLENGES IN THE BOOK OF VIRGIN ANNUNCIATE

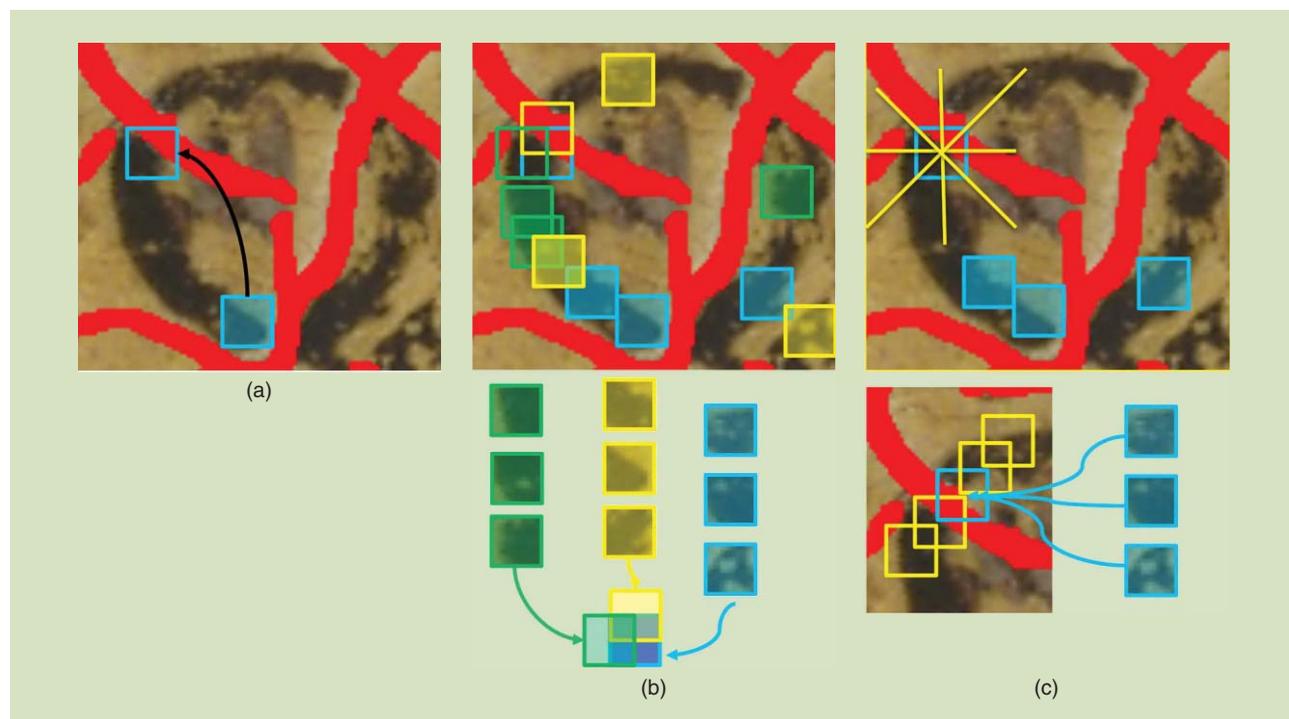
Since cracks typically appear in images as very thin and elongated regions, crack inpainting methods are often based on rather simple, pixel-wise operations, including median filtering [14], [26], interpolation [2], and controlled anisotropic diffusion [14]. In cases where high-resolution scans are available, such that the width of some cracks spans multiple pixels, patch-based inpainting methods [9], [17], [24] typically yield better results [7], [27].

The book in the panel *Virgin Annunciate* (Figure 2) is a very challenging case for virtual inpainting because the width of cracks varies greatly and some cracks are difficult to distinguish from parts of the letters. Moreover, as the cracks are typically surrounded by bright clouds of background matching color, due to the lifting and abrasion of the surrounding paint and thereby imposed light reflections, the immediate areas around the cracks are also unreliable. The crack inpainting method from [7] specifically tailored to this application already generally improved the legibility of the text and was shown to outperform some “general-purpose” patch-based inpainting methods like [9] and [17]. However, some

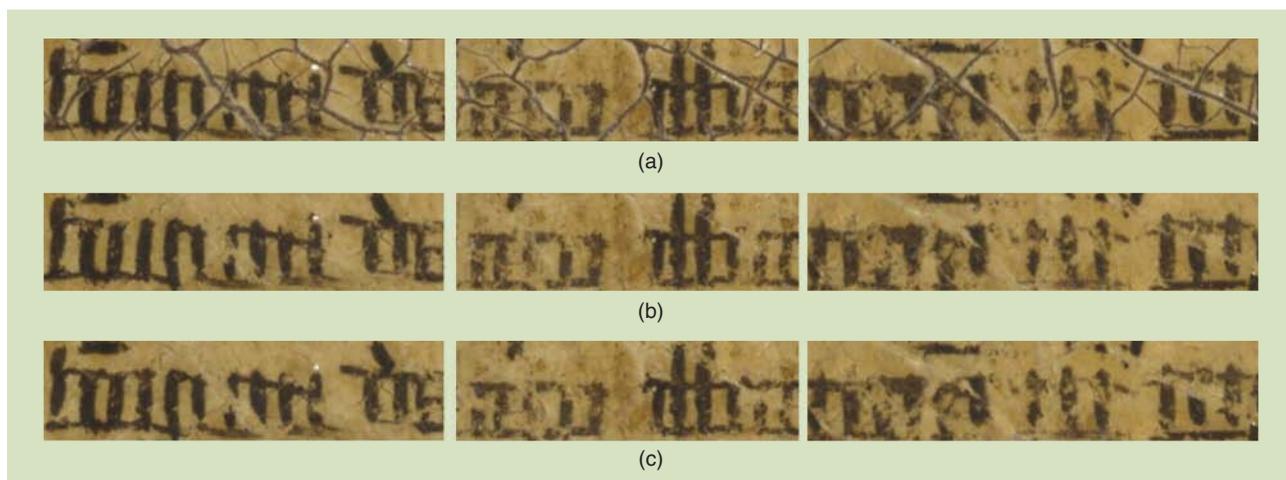
problems are still present, e.g., parts of the letters through which wide cracks are passing are occasionally deleted after virtual inpainting. To alleviate this problem, a better approach to handling continuation of image structures is needed. We discuss a possible solution with encouraging initial results next.

CRACK INPAINTING

In cases where painted structures, like the characters in Figure 3, are relatively small compared to the crack width, it is very difficult for the inpainting algorithm to infer the correct structure locally. Patch-based inpainting methods typically handle structure propagation by defining the right filling order [9], [15]. Once the filling order is determined, most of the methods choose plausible candidates for replacement patches based solely on the agreement with the undamaged part of a single target patch [see Figure 7(a)], and concentrate on defining effective distance metrics between the known portion of the target and the candidates [20]. Matching only against a small part within the target patch increases the risk of propagating wrong textures and wrong colors into the missing region. Global methods, such as [17] and [25], allow multiple candidates and optimize their mutual agreement in the overlap regions [Figure 7(b)], but even this cannot ensure agreement with surrounding undamaged structures: the optimization that takes care that neighboring replacement patches mutually agree cannot “undo” the damage done by selecting wrong candidates in the first place. Ideally, undamaged areas around the target patch should be taken into account in the candidate selection as well, ensuring that plausible candidates agree with true structures [see Figure 7(c)].



[FIG7] Patch-based inpainting and the proposed improvement. (a) A greedy approach chooses one replacement patch based on the known part of the damaged patch. (b) A global approach considers multiple candidates and their mutual agreement. (c) Our approach in addition adapts the candidate selection according to the locally detected structures.



[FIG8] (a) Parts of the original image, (b) the results of our method, and (c) the method of [7].

Guiding the selection of candidate patches by the agreement with undamaged areas is related to the idea of global visual coherence introduced in [29] for video completion. The approach of [29] is very effective for replicating larger missing areas and textures, but it does not treat continuation of curvilinear structures. Alternative solutions that propagate structures along user-specified lines [19], [28] showed excellent results in photo editing, but for our application the amount of the user intervention required by such methods would be prohibitive.

The main idea of our approach is to simultaneously detect directions of local structure propagation and adapt the candidate selection accordingly. We propose a fully automatic and low-complexity method for selecting the candidate replacement patches based on their agreement with the undamaged part of the target patch and with the neighboring undamaged areas, along directions where the structures are likely to propagate. Let ϕ_i denote an image patch centered at position i , $\mathcal{S}(\phi_i, \phi_j)$ a certain measure of similarity between ϕ_i and ϕ_j , and denote by $N_{i,k}$ a directional neighborhood of ϕ_i along direction k . We define prior preference $P_{i,j}$ for selecting a source patch ϕ_i as a candidate replacement for the damaged target patch ϕ_j as follows:

$$P_{i,j} = \mathcal{S}(\phi_i, \phi_j) + \max_k \sum_{l \in N_{j,k}} \mathcal{S}(\phi_i, \phi_l). \quad (3)$$

The first term measures, as usual, the similarity with the known part of the target patch. The novelty is in the second term, which takes care of the agreement with the wider context around the target patch. In particular, the more the candidate patch ϕ_i fits with the neighborhood of the damaged patch ϕ_j in any direction where structures of interest are likely to propagate, the more preference it will get in the selection process. Common measures of patch similarity are defined in terms of the sum of squared differences among the patches $\mathcal{D}(\phi_i, \phi_j) = \|\phi_i - \phi_j\|^2$, calculated over the known pixels. We used $\mathcal{S}(\phi_i, \phi_j) = -\mathcal{D}(\phi_i, \phi_j)$ for the candidate selection in (3). With this improved candidate selection process in combination with simple greedy inpainting (selection of one replacement patch at

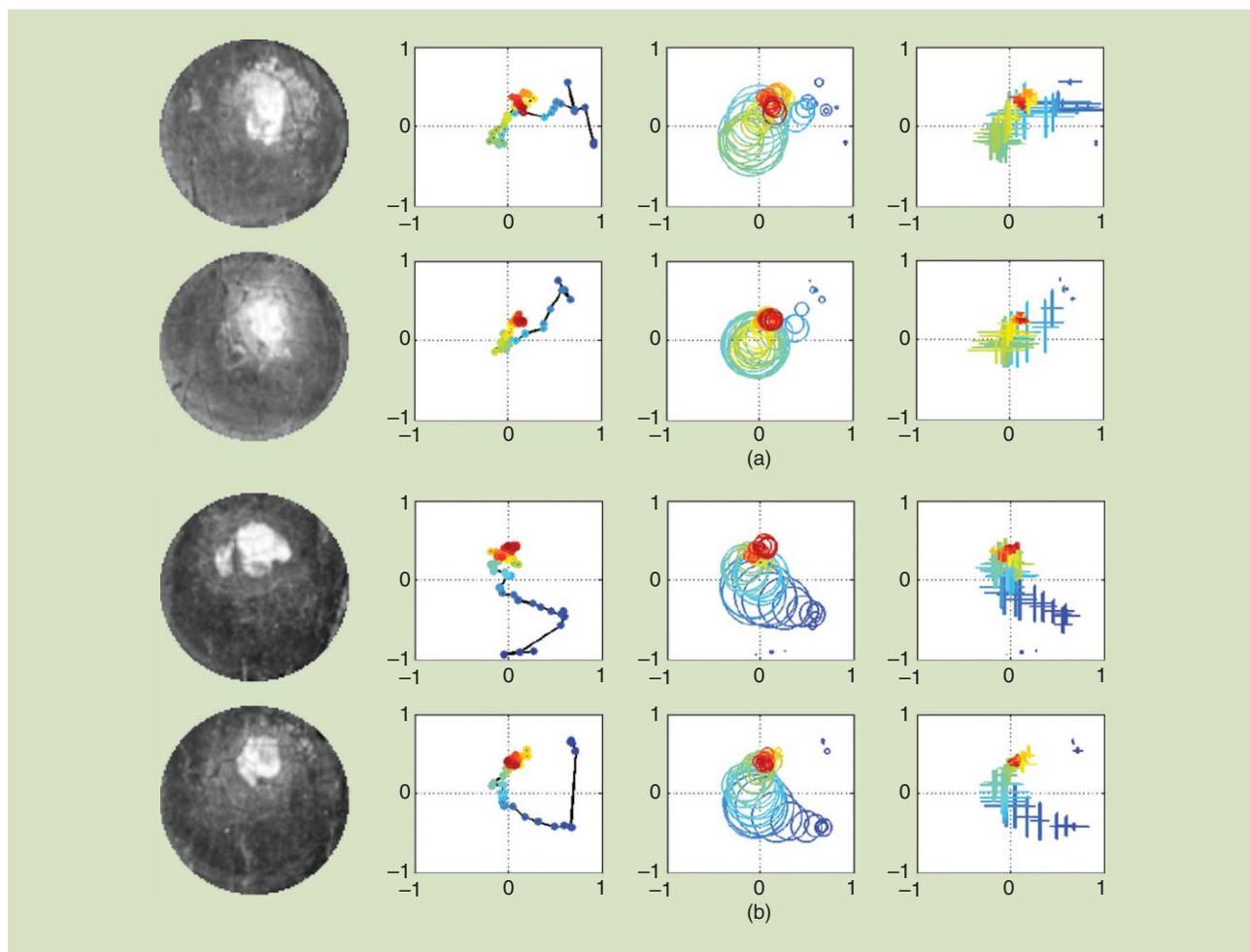
each position), we already obtain a clear improvement over the earlier method from [7], as is visible in Figure 8 (notice, in particular, that the effect of deleting parts of letters is less severe). One can also select multiple candidates, with several largest values of $P_{i,j}$ in (3) and subsequently solve the resulting “puzzle” using a global optimization method like in [17] and [25]. It would be interesting to explore also alternative solutions, like the statistics of patch offsets [16] or hierarchical, superresolution-based inpainting [18].

WHAT CAN PEARLS TELL US?

Painted pearls, which are abundant in *The Ghent Altarpiece*, provide a nice case study for the statistical analysis of the consistency of the painterly execution. Spatial histograms, or spatio-grams [3], were employed in [23] as digital signatures of painted pearls and showed potential to distinguish pearls painted by different artists. In particular, the pearls in the copy of the panel *Just Judges*, made by J. Van der Veken between 1939 and 1951 to replace the panel stolen in 1934, showed clearly different spatio-grams than those from other panels of the altarpiece. Similar conclusions were drawn when comparing the spatio-grams of the recent reconstructions by other artists. However, this earlier analysis in [23] was performed on the old scans of the altarpiece, with varying resolutions, which may have affected to some extent the numerical findings. It is important to verify the main conclusions of this earlier analysis in the light of the new high-quality photographic material. We also go a step further, extending the study to different panels of the altarpiece and making a hypothesis that the consistency of the painted pearls could provide an additional support for the division of hands between the painters or within the workshop, as well as for detecting possible areas of former restorations and overpainting campaigns.

DIGITAL PEARL SIGNATURES FROM SPATIOGRAMS

An image spatio-gram [3] is a generalized histogram, with second-order spatial moments. Suppose an image consists of N pixels and denote the spatial position of the n th pixel $n \in \{1, \dots, N\}$ by



[FIG9] (a) and (b) Two pairs of pearls from *The Ghent Altarpiece* and their corresponding spatiogram triple-plots (S1, S2, S3).

$\mathbf{p}_n = (x_n \ y_n)^T$ and its intensity by I_n . Let b denote a histogram bin, being a range of pixel intensities and let $\mathbf{1}_b(x)$ denote the indicator function (returning one if $x \in b$ and zero otherwise). The spatiogram triplet for bin b is then computed as follows:

$$\begin{aligned}
 c_b &= \eta \sum_{n=1}^N \mathbf{1}_b(I_n) \\
 \boldsymbol{\mu}_b &= \eta c_b^{-1} \sum_{n=1}^N \mathbf{p}_n \mathbf{1}_b(I_n) \\
 \boldsymbol{\Sigma}_b &= \eta c_b^{-1} \sum_{n=1}^N (\mathbf{p}_n - \boldsymbol{\mu}_b)(\mathbf{p}_n - \boldsymbol{\mu}_b)^T \mathbf{1}_b(I_n). \quad (4)
 \end{aligned}$$

The normalizing constant η is chosen such that $\sum_{b=1}^B c_b = 1$. For bins with $c_b = 0$, the values of $\boldsymbol{\mu}_b$ and $\boldsymbol{\Sigma}_b$ are also set to zero (not of interest). To enable comparison between images of different sizes, we normalize all spatial coordinates to the range $[-1, 1]$. For the purpose of visualization of the highly dimensional spatiogram data, [23] proposed spatiogram triple-plots (S1, S2, S3) illustrated in Figure 9:

- S1: connected centers of bins, $\boldsymbol{\mu}_b = (\bar{x}_b \ \bar{y}_b)$
- S2: $\boldsymbol{\mu}_b$ -positioned counts of bins (the radii of the circles are proportional to bin counts)

- S3: $\boldsymbol{\mu}_b$ -positioned variances in the x - and y -direction.

Figure 9 demonstrates that mutually similar pearls produce similar spatiogram triple-plots.

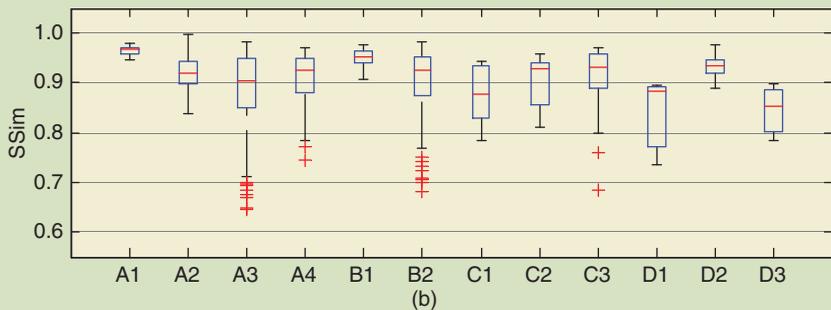
CONSISTENCY OF THE PEARLS IN THE ALTARPIECE

Here, we evaluate the consistency of painted pearls in *The Ghent Altarpiece* on the data set in [31]. We measure similarity between two painted pearls with the spatiogram similarity index SSim [6], where $0 \leq \text{SSim} \leq 1$. In particular, we select 12 sets of pearls from five different panels, as marked in Figure 10(a). The size of each set and the average radius of the pearls (in pixels) are given in parentheses: A1(4, 576), A2(12, 265), A3(24, 273), A4(12, 144), B1(7, 177), B2(20, 129), C1(5, 138), C2(6, 111), C3(8, 180), D1(3, 276), D2(6, 224), and D3(3, 239). We make a comparative analysis of the similarity of painted pearls within each set (within-set similarity) and between different sets (cross-set similarity).

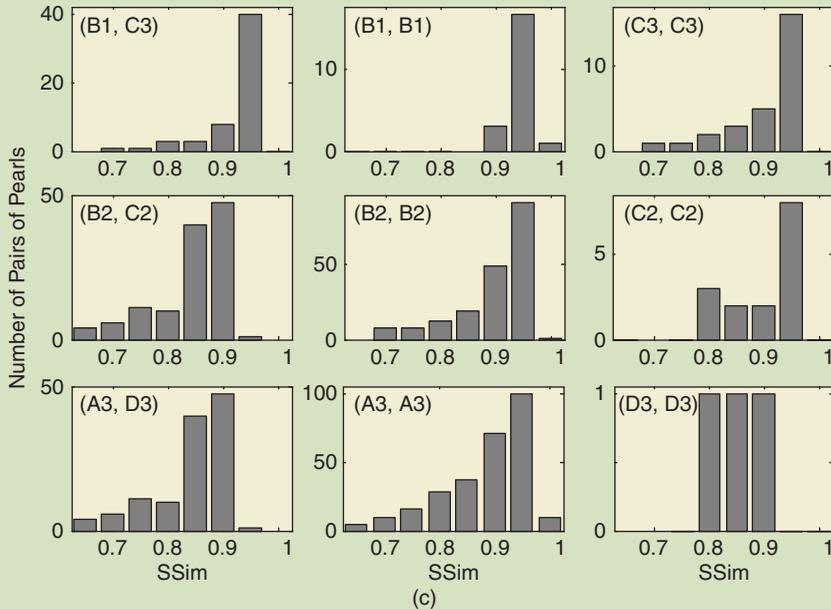
Figure 10(b) shows within-set similarity for the 12 pearl sets, computed from 64-bin spatiograms. Clearly, the largest pearls (A1) are the most similar, which agrees with the findings of [23], but now we can also see that the consistency within each set depends also on the panel and the position of the pearls. For



(a)



(b)



(c)

[FIG10] Exploring the consistency of the painted pearls in different panels. (a) The selected sets of pearls: A—the clothing decoration of *God the Father*; B—the hat of *Cumaean Sibyl*; C—the diadem and brooch of *The Virgin Annunciate*; D1—the decoration in *Prophet Zachary*; and D2, D3—the diadem and brooch in *The Archangel Gabriel*. (b) The corresponding SSim values. The central marks show the medians, the boxes indicate 25th and 75th percentiles, the extreme vertical lines (whiskers) are 1.5 times the interquartile range, and + marks denote the outliers. (c) Some cross-set and within-set SSim histograms.

example, C2 shows much higher within-set similarity than D1, even though D1 pearls are twice as large. This could be attributed to the position of D1 pearls (very high in the altarpiece, where they are less visible to viewers). Similarly, SSim values are less consistent for B2 than for B1 of the *Cumaean Sibyl*, in line with the fact that B2 pearls are smaller, as well as less densely/neatly arranged than the B1 pearls. Sets C1–C3 (from *The Virgin Annunciate*) and D2 and D3 (from *The Archangel Gabriel*) are on equally important places in the altarpiece: two outer panels in the middle row of the closed view; see Figure 1(b). Our analysis suggests that C3 pearls are more consistent than C1 and C2, which agrees with the fact that they are larger, more visible, and have a more central place in the panel. Furthermore, D2 pearls are almost twice as large as C1 (which appear in the same relative position in the other panel) and indeed give more consistent spatiograms. However, there are some interesting exceptions that deserve extra attention. D3 pearls seem less consistent than D2, even though they are larger and at a more visible place (closer to the viewer). It is still unclear whether this could be (partly) attributed to less careful execution, the state of varnish, possible retouching or overpaint, or simply to having a rather small set (three pearls).

It is also interesting to assess consistency of the painted pearls of similar sizes in different panels. Figure 10(c) shows cross-set SSim for three different combinations of pearl sets in comparison to within-set SSim for the same sets. The SSim histograms show a high similarity between B1 and C3 sets, and much less between B2 and C2. This can be attributed to the fact that B2 and C2 pearls are relatively small. Moreover, B2 is a large set of pearls decorating a hat, and executing them more consistently would not change the overall visual impression much. Also, the cross-set similarity between A3 and D3 is much smaller than the similarity within A3, which is not surprising since A3 pearls are in the central panel of *God the Father* in the open view and are brimming with jewels and reflections, while the D3 ones are in the closed view, figuring different type of lighting, possibly executed by different hands within the workshop and have quite likely undergone different conservation treatments in the past.

CONCLUSIONS

Signal processing shows promise in helping in the decision-making process that is involved in a painting's conservation and restoration. Our initial results show that analysis of crack patterns could indicate certain areas of overpaint, even though the processed crack maps still need to be interpreted by conservators using other examination techniques. State-of-the-art inpainting techniques still do not succeed in fully restoring the painted form at a level that would match the criteria of art conservators, but the interaction between the two communities provides already a crucial feedback for improving virtual inpainting techniques in this challenging application. The use of statistical analysis to assess the consistency of the painting style can be of interest for art-historical interpretation of the content. Our analysis of the consistency of the painted pearls in *The Ghent Altarpiece* points to some instances (specific painted objects) that might be of interest to art historians and conservators to examine in more detail.

AUTHORS

Aleksandra Pižurica (aleksandra.pizurica@telin.ugent.be) received the diploma degree in electrical engineering from the University of Novi Sad, Serbia, in 1994; the M.Sc. degree in telecommunications from the University of Belgrade, Serbia, in 1997; and the Ph.D. degree in engineering from Ghent University, Belgium, in 2002. Since 2009, she has been a professor in statistical image modeling at Ghent University and was elected as a principal investigator in the Research Department Multimedia Technology of iMinds. She is an associate editor of *IEEE Transactions on Image Processing*.

Ljiljana Platiša (ljiljana.platisa@telin.ugent.be) is a postdoctoral researcher in the Image Processing and Interpretation Research group within the Department of Telecommunications and Information Processing, Ghent University, Belgium. She received her master's degree in electrical engineering from the University of Novi Sad, Serbia, in 2001 and her Ph.D. degree in computer science engineering from Ghent University, Belgium, in 2014. Her work experience spans both academia and industry. Her research focuses on theoretical and psychophysical aspects of image quality, covering both technical and task-specific approaches.

Tijana Ružić (tijana.ruzic@telin.ugent.be) received her M.Sc. degree in electrical and computer engineering from the University of Novi Sad, Serbia, in 2008 and the Ph.D. degree in engineering from Ghent University, Belgium, in 2014. She is currently with the Image Processing and Interpretation Group, Department of Telecommunications and Information Processing, Ghent University, as a postdoctoral researcher. Her main research interests are patch representations and Markov random field models in image restoration.

Bruno Cornelis (bcorneli@etro.vub.ac.be) received the master's degree of industrial engineer in electronics from Erasmushogeschool Brussel, Belgium, in 2005 and the master's degree in electrical engineering from Vrije Universiteit Brussel (VUB), Belgium, in 2007. He obtained a Ph.D. degree in applied sciences and engineering from VUB in 2014. During his Ph.D. studies, he investigated the use of various image processing tools in support of art scholarship. His research interests include statistical data analysis and sparse representations in image processing. Currently he is a postdoctoral researcher with the Department of Mathematics at Duke University, Durham, North Carolina.

Ann Dooms (adooms@etro.vub.ac.be) holds a Ph.D. degree in mathematics and a professorship at the Department of Electronics and Informatics (ETRO) of Vrije Universiteit Brussel (VUB) and iMinds. In October 2007, she founded a research team at ETRO in multimedia forensics. Topics include the protection of multimedia through digital watermarking, perceptual quality assessment, compressive sensing, and digital painting analysis. She was elected as a member of the Jonge Academie and the IEEE Technical Committee in Information Forensics and Security. In 2014, she received the Second Prize in Science Communication of the Royal Flemish Academy of Belgium in Sciences and Arts.

Maximiliaan Martens (maximiliaan.martens@UGent.be) received his Ph.D. degree in art history at the University of California, Santa Barbara, in 1992. He is a professor of art history at the University of Ghent, Belgium, a member of the Royal

Flemish Academy of Science and Art of Belgium, and the director of the Ghent Interdisciplinary Centre for Art and Science. His research focuses on fine arts in The Netherlands, 14th–16th centuries, and more particularly on artistic production in an urban context and on the interdisciplinary study of the arts and sciences (imaging techniques, image processing, and analytical chemistry).

Hélène Dubois (helene.dubois@kikirpa.be) is a painting conservator and art historian. She trained at the Université Libre de Bruxelles and at the Hamilton Kerr Institute, University of Cambridge, United Kingdom. She worked at the Doerner-Institut in Munich, the J. Paul Getty Museum in Malibu, California, the Royal Museums of Fine Arts in Brussels, Belgium, and the Limburg Conservation Institute in Maastricht, The Netherlands. With the Royal Institute for Cultural Heritage (KIK-IRPA), Brussels, since 2008, she is the research coordinator of the conservation project of the altarpiece of the *The Adoration of the Mystic Lamb* and researching the material history of the altarpiece for a Ph.D. dissertation at Ghent University, Belgium.

Bart Devolder (bart.devolder@kikirpa.be) received his master's degree in conservation and restoration of paintings from the Artesis Hogeschool Antwerpen, Belgium. He held internships at the Akademia Sztuk Pięknych Krakow, Poland, Royal Institute for Cultural Heritage (KIK/IRPA), Brussels, and at the Musée du Louvre, Paris. He received a fellowship from the Straus Center for Conservation at the Harvard University Art Museums (2003–2004), and he was the Andrew W. Mellon Fellow in Painting Conservation at the National Gallery of Art, Washington D.C. (2004–2007). Before joining the KIK-IRPA team as the on-site coordinator/painting conservator for the restoration of *The Ghent Altarpiece*, he worked for the Kimbell Art Museum and Amon Carter Museum of American Art in Fort Worth, Texas, as associate conservator of paintings.

Marc De Mey (marc.demey@kvab.be) has a background in psychology and philosophy of science and is a member of the Royal Flemish Academy of Science and Art of Belgium and a former director of its institute for advanced studies, Vlaams Academisch Centrum. Until his emeritus status in 2005 at Ghent University, Belgium, he taught cognitive science in the Faculty of Letters. He has been a Fulbright scholar at the Harvard Center for Cognitive Studies and Peter Paul Rubens Professor at the University of California, Berkeley.

Ingrid Daubechies (ingrid@math.duke.edu) is a James B. Duke Professor of Mathematics at Duke University. She received her B.S. and Ph.D. degrees in physics from the Vrije Universiteit Brussel, Belgium, where she also started her research career. In the United States, she was previously with AT&T Bell Labs (1987–1994), Rutgers University (1991–1993), and Princeton University (1994–2010). She received the Benjamin Franklin Medal for Electrical Engineering and the James Kilby Medal of the IEEE Signal Processing Society, both in 2011. She is a Fellow of the IEEE.

REFERENCES

- [1] F. S. Abas, "Analysis of Craquelure patterns for content-based retrieval," Ph.D. dissertation, Univ. of Southampton, 2004.
- [2] M. Barni, F. Bartolini, and V. Cappellini, "Image processing for virtual restoration of artworks," *IEEE Multimedia*, vol. 7, no. 2, pp. 34–37, 2000.
- [3] S. T. Birchfield and S. Rangarajan, "Spatiograms versus histograms for region-based tracking," in *Proc. IEEE Conf. Computer Vision and Pattern Rec. (CVPR)*, 2003, pp. 1158–1163.

- [4] S. Bucklow, "A stylometric analysis of craquelure," *Comput. Human.*, vol. 31, 1998.
- [5] I. A. Carreras, C. O. S. Sorzano, R. Marabini, J. M. Carazo, C. O. de Solorzano, and J. Kybic, "Consistent and elastic registration of histological sections using vector-spline regularization," in *Computer Vision Approaches to Medical Image Analysis* (Vol. 4241: Lecture Notes in Computer Science, vol. 4241). Berlin, Germany: Springer, 2006, pp. 85–95.
- [6] C. O. Conaire, N. E. O'Connor, and A. F. Smeaton, "An improved spatiogram similarity measure for robust object localisation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, 2007, pp. 1-1069–1-1072.
- [7] B. Cornelis, T. Ružić, E. Gezels, A. Dooms, A. Pižurica, L. Platiša, J. Cornelis, M. Martens, M. De Mey, and I. Daubechies, "Crack detection and inpainting for virtual restoration of paintings: The case of *The Ghent Altarpiece*," *Signal Process.*, vol. 93, no. 3, pp. 605–619, 2013.
- [8] B. Cornelis, Y. Yang, J. Vogelstein, A. Dooms, I. Daubechies, and D. Dunson, "Bayesian crack detection in ultra high resolution multimodal images of paintings," in *Proc. 18th Int. Conf. Digital Signal Processing (DSP 2013)*, Santorini, Greece, 2013.
- [9] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Processing*, vol. 13, no. 9, pp. 1200–1212, Sept. 2004.
- [10] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, pp. 1253–1278, 2000.
- [11] M. De Mey, M. Martens, and C. Stroo, *Vision and Material: Interaction Between Art and Science in Jan van Eyck's Time*. Brussels, Belgium: KVAB Press, 2012.
- [12] A. L. Dierick, *The Ghent Altarpiece: Van Eyck's Masterpiece Revisited—In Colour and Life Size*. Ghent, Belgium, 1996.
- [13] M. Elad, M. Aharon, and A. M. Bruckstein, "K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [14] I. Giakoumis, N. Nikolaidis, and I. Pitas, "Digital image processing techniques for the detection and removal of cracks in digitized paintings," *IEEE Trans. Image Processing*, vol. 15, no. 1, pp. 178–188, 2006.
- [15] C. Guillemot and O. Le Meur, "Image inpainting: Overview and recent advances," *Signal Process. Mag.*, vol. 31, no. 1, pp. 127–143, Jan. 2014.
- [16] K. He and J. Sun, "Statistics of patch offsets for image completion," in *Proc. European Conf. Comput. Vision (ECCV)*, 2012, pp. 16–29.
- [17] N. Komodakis and G. Tziritas, "Image completion using efficient belief propagation via priority scheduling and dynamic pruning," *IEEE Trans. Image Processing*, vol. 16, no. 11, pp. 2649–2661, Nov. 2007.
- [18] O. Le Meur, M. Ebdelli, and C. Guillemot, "Hierarchical super-resolution-based inpainting," *IEEE Trans. Image Processing*, vol. 22, no. 10, pp. 3779–3790, 2013.
- [19] J. Lee, D.-K. Lee, and R.-H. Park, "Robust exemplar-based inpainting algorithm using region segmentation," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 553–561, 2012.
- [20] R. Martínez-Noriega, R. Roumy, and G. Blanchard, "Exemplar-based image inpainting: Fast priority and coherent nearest neighbor search," in *Proc. IEEE Int. Workshop on Machine Learning for Signal Processing*, 2012, pp. 1–6.
- [21] G. Papandreou, P. Maragos, and A. Kokaram, "Image inpainting with a wavelet domain hidden Markov tree model," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, 2008, pp. 773–776.
- [22] 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.
- [23] L. Platiša, B. Cornelis, T. Ružić, A. Pižurica, A. Dooms, M. Martens, M. De Mey, and I. Daubechies, "Spatiogram features to characterize pearls in paintings," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, 2011, pp. 801–804.
- [24] T. Ružić, B. Cornelis, L. Platiša, A. Pižurica, A. Dooms, W. Philips, M. Martens, M. De Mey, and I. Daubechies, "Virtual restoration of *The Ghent Altarpiece* using crack detection and inpainting," in *Proc. Advanced Concepts for Intelligent Vision Systems (ACIVS)*, 2011, pp. 417–428.
- [25] T. Ružić and A. Pižurica, "Context-aware patch-based image inpainting using Markov random field modelling," *IEEE Trans. Image Processing*, vol. 24, no. 1, pp. 444–456, 2015.
- [26] S. V. Solanki and A. R. Mahajan, "Cracks inspection and interpolation in digitized artistic picture using image processing approach," *Int. J. Recent Trends Eng.*, vol. 1, no. 2, pp. 97–99, May 2009.
- [27] G. S. Spagnolo and F. Somma, "Virtual restoration of cracks in digitized image of paintings," *J. Phys.: Conf. Ser.*, vol. 249, no. 1, p. 012059, 2010.
- [28] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum, "Image completion with structure propagation," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 861–868, 2005.
- [29] Y. Wexler, E. Shechtman, and M. Irani, "Space-time completion of video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 463–476, 2007.
- [30] Y. Yang and D. B. Dunson, "Bayesian conditional tensor factorizations for high-dimensional classification," Jan. 2013. arXiv:1301.4950v1. [Online]. Available <http://arxiv.org/abs/1301.4950>
- [31] Closer to van Eyck: Rediscovering *The Ghent Altarpiece* Web site. [Online]. Available: <http://closer tovanyeck.kikirpa.be/>

Kin-Man Lam, Carlos Óscar S. Sorzano,
Zhilin Zhang, and Patrizio Campisi

Undergraduate Students Compete in the IEEE Signal Processing Cup: Part 1

Signal processing is the key to success for the efficient storage, transmission, and manipulation of information. The liveliness of signal processing relies on having a large number of students who undertake this research and embark on this career path. Signal processing is part of the curriculum in many undergraduate engineering programs. Some of the students also do their capstone or final-year projects on signal processing. To increase students' interest in signal processing and to get them to better appreciate its applications in real life, the IEEE Signal Processing Society (SPS) has created an undergraduate competition, referred to as the *Signal Processing Cup (SP Cup)* [1], which provides undergraduate students with an opportunity to form teams and work together to solve a challenging and interesting real-world problem using signal processing techniques and methods. The first competition was held at the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2014. Based on its success, it has now been instated as an annual event at ICASSP conferences.

This column will offer a series of articles that introduce the SP Cup, outline how graduate students are supported in the competition, how they apply their signal processing knowledge, what they have learned from the competition, and relay their feedback and experience. It is articulated into three parts, two of which will be published in upcoming issues of *IEEE Signal Processing Magazine*. In this issue, Part 1 introduces the SP Cup and the competition topics for the first and second

editions. Parts 2 and 3 will describe the signal processing techniques used by the teams in the first and second competitions, respectively. We will describe how the students work together on their projects and present their feedback on key aspects of both competitions.

STUDENT TEAMS AND SUPPORT

To join the competition, undergraduate students are required to form a team. Each team must be composed of one faculty member (to supervise the team members), one graduate student (to assist the supervisor), and at least three but no more than ten undergraduate students. The teams can choose the programming language they prefer for the competition. As the competition is sponsored by MathWorks, each team that registers for the SP Cup is provided with complimentary software, including MATLAB and some selected toolboxes. The three top teams are selected to participate in the final competition at ICASSP, where the prizes are awarded. The teams are awarded prizes of US\$5,000, US\$2,500, and US\$1,000 for first, second, and third places, respectively.

The three selected teams are also supported by the SPS with travel grants to attend the final competition at ICASSP. Each team member is offered up to US\$1,200 for continental travel or US\$1,700 for intercontinental travel. A maximum of three members per team are eligible for travel support. Furthermore, all attending members are given a complimentary registration so that they can also attend the conference, meet other researchers and professors, and experience the signal processing research world.

Through these SP Cup competitions, more undergraduate students can put their signal processing knowledge into practice in a real-world project. They will

learn how to cooperate with other team members and develop their interest in signal processing research. The top teams also have the opportunity to expand their horizon by attending the largest signal processing conference—ICASSP. So far, two editions of the competition have been run, and both of the competition themes were proposed by the Bioimaging and Signal Processing Technical Committee of the IEEE SPS. In the first edition, students were challenged to enhance the resolution of a three-dimensional (3-D) model of macromolecular structures obtained by transmission electron microscopy and used in structural biology. Approximately 100 undergraduate students from all over the world registered for the competition and were grouped into 26 different teams. Finally, 12 teams submitted their work.

The theme for the second competition was “Heart Rate Monitoring During Physical Exercise Using Wrist-Type Photoplethysmographic (PPG) Signals.” About 270 undergraduate students split among 66 teams registered for the competition. Ultimately, 49 teams submitted their work. Three teams were selected to participate in the final competition at ICASSP 2015. The following sections will briefly introduce the two competition themes, and how they were designed for undergraduate students.

THE COMPETITION THEME OF THE FIRST EDITION

The first edition of the SP Cup offered students the opportunity to design an algorithm to improve the resolution of biological macromolecules as observed in a transmission electron microscope. The goal was to provide structural biologists with tools capable of better showing the atomic structure of proteins, which determine the physiological and pathological

Digital Object Identifier 10.1109/MSP.2015.2419291

Date of publication: 15 June 2015



[FIG1] Members from the three selected teams, the judging panel, sponsors, and organizers of SP Cup 2014 attended the final SP Cup competition at ICASSP 2014.



[FIG2] Members from the three selected teams, the judging panel, sponsors, and organizers of the SP Cup 2015 attended the final SP Cup competition at ICASSP 2015.

behavior of our cells. By being able to analyze the 3-D structure of protein targets, pharmaceutical companies can design highly specific drugs that strongly interact with the target molecule of interest.

Students dived into the realm of image restoration, deconvolution, deblurring, denoising, superresolution, and sparse coding to produce state-of-the-art algorithms that allowed the resolution of the reconstructed volumes to increase by about 0.5 \AA (an atom has a radius between 0.5 and 2 \AA). This increase in resolution is another step forward in the scientific quest of being able of analyzing the structure and function of each of the pieces that comprises life.

THE COMPETITION THEME OF THE SECOND EDITION

In the second edition of the SP Cup, students were asked to design algorithms to estimate heart rate using PPG signals recorded from subjects' wrists during physical exercise. These algorithms can potentially be used in smart watches and wristbands for health monitoring and fitness tracking.

The challenge in this task was to overcome strong interference caused by body movements in exercise to accurately estimate heart rate. A successful system

architecture to solve this challenging problem generally consists of three components: motion-artifact removal, power spectrum estimation, and selection of spectral peaks corresponding to heart rate. Signal processing techniques for the first two components are generally taught in undergraduate signal processing courses, such as adaptive filtering for noise removal and nonparametric power spectrum estimation. For the third component, students have a lot of flexibility to design selection algorithms with full considerations to overcome various practical situations. In this competition, students can separately work on each of the parts. However, since the three components interplay with each other, students also need to collaborate in close relationship to design a successful system as a whole.

This competition not only provided an opportunity for students to solve an important practical signal processing problem tapping into their knowledge taught in class, but it also prepared them for entering related industries especially in the field of wearable health care.

COMPETITION RESULTS

The First SP Cup competition was held at ICASSP 2014 on 8 May. The judging panel

found it difficult to rank the final placement. Based on the performances of the algorithms developed and the presentations made by the three selected teams (Figure 1), the final results were as follows:

- *First place:* EPOCH (Anik Khan, Forsad Al Hossain, Tawab Ullas, Md. Abu Rayhan, and Mohammad Ariful Haque) from Bangladesh University of Engineering and Technology
- *Second place:* NtUeLsA (Kai-Wen Liang, Yen-Chen Wu, Guan-Lin Chao, Kuan-Hao Huang, Shao-Hua Sun, Ming-Jen Yang, Po-Wen Hsiao, Ti-Fen Pan, Yi-Ching Chiu, Wei-Chih Tu, and Shao-Yi Chien) from National Taiwan University
- *Third place:* Uchihias (Emroz Khan, Shiekh Zia Uddin, Mukhlasur Rahman Tanvir, and Md. Kamrul Hasan) from Bangladesh University of Engineering and Technology.

The second SP Cup competition was held during ICASSP 2015 on 20 April (Figure 2). The final results were as follows:

- *First place:* Signal Processing Crew Darmstadt (Alaa Alameer, Bastian Alt, Christian Sledz, Hauke Radtki, Maximilian Huettenrauch, Patrick Wenzel, Tim Schaeck, and Michael Muma) from Technische Universitaet Darmstadt

■ **Second place:** Supersignal (Sayeed Shafayet Chowdhury, Rakib Hyder, Anik Khan, Md. Samzid Bin Hafiz, Zahid Hasan, and Mohammad Ariful Haque) from Bangladesh University of Engineering and Technology

■ **Third place:** SSU (Gyehyun Baik, Minkyu Jung, Hyunil Kang, Jungsub Lee, Baeksan Ohn, Sunho Kim, and Sungbin Im) from Soongsil University.

From the first to the second edition of the competition, the number of undergraduate students participating in the event has increased from about 100 to 270. We foresee that students will become increasingly aware of this competition, and more and more students will get involved.

FORTHCOMING COMPETITIONS

In Parts 2 and 3, we will describe the approaches developed by the students and interview some who participated. We will report their feedback and the secrets of their successes. The third edition of the SP Cup will be held at ICASSP 2016. The

theme of the 2016 competition will be announced in September of this year. We hope that this series of columns will help undergraduate students who are interested in the future competitions to gain a solid understanding of the SP Cup and help them prepare for it.

AUTHORS

Kin-Man Lam (enklam@polyu.edu.hk) is a professor with the Centre for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University. He is the IEEE Signal Processing Society director of membership services, an area editor of *IEEE Signal Processing Magazine*, and was a coorganizer of the first IEEE Signal Processing Cup in 2014.

Carlos Óscar S. Sorzano (coss@cnb.csic.es) is the technical director of the Instruct Image Processing Center–National Center of Biotechnology (CSIC) and a member of the Bioimaging and Signal Processing Technical Committee of the IEEE Signal Processing Society. He

was a coorganizer of the first IEEE Signal Processing Cup in 2014.

Zhilin Zhang (zhilinzhang@ieee.org) is a staff research engineer and manager with Samsung Research America, Dallas, Texas. He is a coorganizer of the 2015 IEEE Signal Processing Cup and a member of the Bioimaging and Signal Processing Technical Committee of the IEEE Signal Processing Society.

Patrizio Campisi (patrizio.campisi@uniroma3.it) is a full professor with the Department of Engineering at the Università degli Studi “Roma Tre,” Roma, Italy. He is the IEEE Signal Processing Society director of student services. His other IEEE Signal Processing Society activities currently include senior associate editor of *IEEE Signal Processing Letters* and general cochair of the IEEE International Workshop on Information Forensics and Security 2015.

REFERENCE

[1] SP Cup Web page. [Online]. Available: <http://www.signalprocessingsociety.org/community/sp-cup/>

SP

IEEE Signal Processing Society Student Membership – FREE to try in 2015

Are you an IEEE Student Member or pondering about signing up? Sign up SPS!

A Membership in the Signal Processing Society (SPS), the IEEE's first society, can help you lay the ground work for many years of success ahead

- **Discounts** on conference registration fees and eligibility to apply for **travel grants**
- **Networking and job opportunities** at the SPS Student Career Luncheon
- **High-impact IEEE Signal Processing Magazine** at your fingertip with opportunities to publish your voice in it

Already a SPS Member? Refer a friend and get rewarded

- IEEE **“Member-Get-a-Member”** reward up to **\$90** per year
- New IEEE student member can **try 2015 SPS Membership FREE**
- Continue SPS exclusive benefits such as **travel grants** and eligibility to enter **SPS student competition** to win a grand prize



Visit <http://ieee.org/join> and select Signal Processing Society; use code **SP15STUAD** for existing IEEE student members and **SP15STUNW** with new IEEE membership sign-up. More at <http://goo.gl/xXUfTF>. Free SPS 2015 exclude existing SPS members. Refer to <http://goo.gl/cxxYh7> for member referral award.

Digital Object Identifier 10.1109/MSP.2015.2434232

Image Analysis: The New Bottleneck in Plant Phenotyping

Plant phenotyping is the identification of effects on the phenotype $P = G \times E$ (i.e., the plant appearance and performance) as a result of genotype differences (i.e., differences in the genetic code) and the environmental conditions to which a plant has been exposed [1]–[3]. According to the Food and Agriculture Organization of the United Nations, large-scale experiments in plant phenotyping are a key factor in meeting the agricultural needs of the future to feed the world and provide biomass for energy, while using less water, land, and fertilizer under a constantly evolving environment due to climate change. Working on model plants (such as *Arabidopsis*), combined with remarkable advances in genotyping, has revolutionized our understanding of biology but has accelerated the need for precision and automation in phenotyping, favoring approaches that provide quantifiable phenotypic information that could be better used to link and find associations in the genotype [4]. While early on, the collection of phenotypes was manual, currently noninvasive, imaging-based methods are increasingly being utilized [5], [6]. However, the rate at which phenotypes are extracted in the field or in the lab is not matching the speed of genotyping and is creating a bottleneck [1].

While the bottleneck was previously the equipment (the hardware), it is now the analysis (the software). There is a need to develop accurate, robust, and automated analysis algorithms that can extract phenotypic information from

experiments on the small (cell) or large scale (field), in two or three dimensions, in the lab but more importantly in the field on real crops. These algorithms should be coupled with affordable platforms and should deal with an immense amount of data produced in these experiments. Experts (from biology as well as data analysis) now agree that the analysis of imaging data is currently the weakest, or even the missing, link due to the major challenges in computer vision and image processing we are currently facing.

COMPUTER VISION AND IMAGE PROCESSING CHALLENGES

Noninvasive plant investigations are done on different scales and modalities using a variety of sensors [2], [5]. This includes optical imaging, hyperspectral imaging to reveal rich pixel information on plant properties, and even magnetic resonance imaging (MRI) and positron emission tomography (PET). Spatial scales vary from the microscopic subcellular level to large outdoor fields. Typical problems in measuring a plant's visible properties comprise measuring size, shape, and other structural traits of whole plants, their organs, or plant populations.

Plants are not static, but self-changing systems with complexity in shape and appearance increasing over time. They emerge below image resolution and grow exponentially in time until, for a single leaf, growth levels off typically at several cm^2 size—i.e., several orders of magnitude change. Relevant timescales for cellular processes may be seconds or minutes, for growing leaves in the range of hours, and the status of whole plants changes over days or even months, in which the surrounding environmental (as well as measurement) conditions may also change.

Algorithms must deal with the aforementioned complexity, and the following sections describe unique challenges by illustrating typical applications. Clearly, the list of applications can never be complete, but we present some of the major themes.

CELLS AND ORGANS: DETECTION, TRACKING, AND STRUCTURAL BREAKS

One of the earliest forms of phenotyping where imaging-based setups were used is in the context of microscopy [2]. Plant tissue samples are excised and imaged under a microscope to reveal the cellular structure [cf. Figure 1(a)]. From an image processing perspective, the automated delineation of cell walls to establish cell morphology and cell count is typically needed.

However, more interesting problems arise from the use of recent techniques such as confocal microscopy, optical projection tomography, and optical coherence microscopy, which permit the noninvasive quantification of cellular morphometry at a variety of scales and depths. These techniques enable the observation of plant tissue dynamics on a short (and long) timescale, therefore tracking problems arise. These become particularly challenging when cell genesis needs to be observed and quantified, since cell division and expansion impose high spatiotemporal fidelity requirements. From a computer vision perspective, this problem, which also occurs in other biomedical applications, entails the inference of time and location of when and where such events occur within the scene, a task radically different from the typical tracking of objects entering or leaving the scene.

Over the last decade, several controlled setups [see Figure 1(c)–(e)] have emerged

that image top-down views of small rosette plants, e.g., Arabidopsis or young tobacco, acquiring either one plant per image or several plants at once [see Figure 1(f)–(g)], housed in so-called growth chambers, where environmental conditions are controlled. Even in this very restricted imaging scenario, fully automatic segmentation of single plants can be a challenge due to, e.g., background clutter from moss growing on the soil, plant-to-plant overlap, heavy contrast changes due to (self-)shadowing, leaf color changes due to stress (e.g., drought), different light conditions and pathogen infections, and plant shape or size variation due to genotypic differences (cultivars or mutants) and treatments.

Segmenting single leaves is a typical multi-instance segmentation task [see

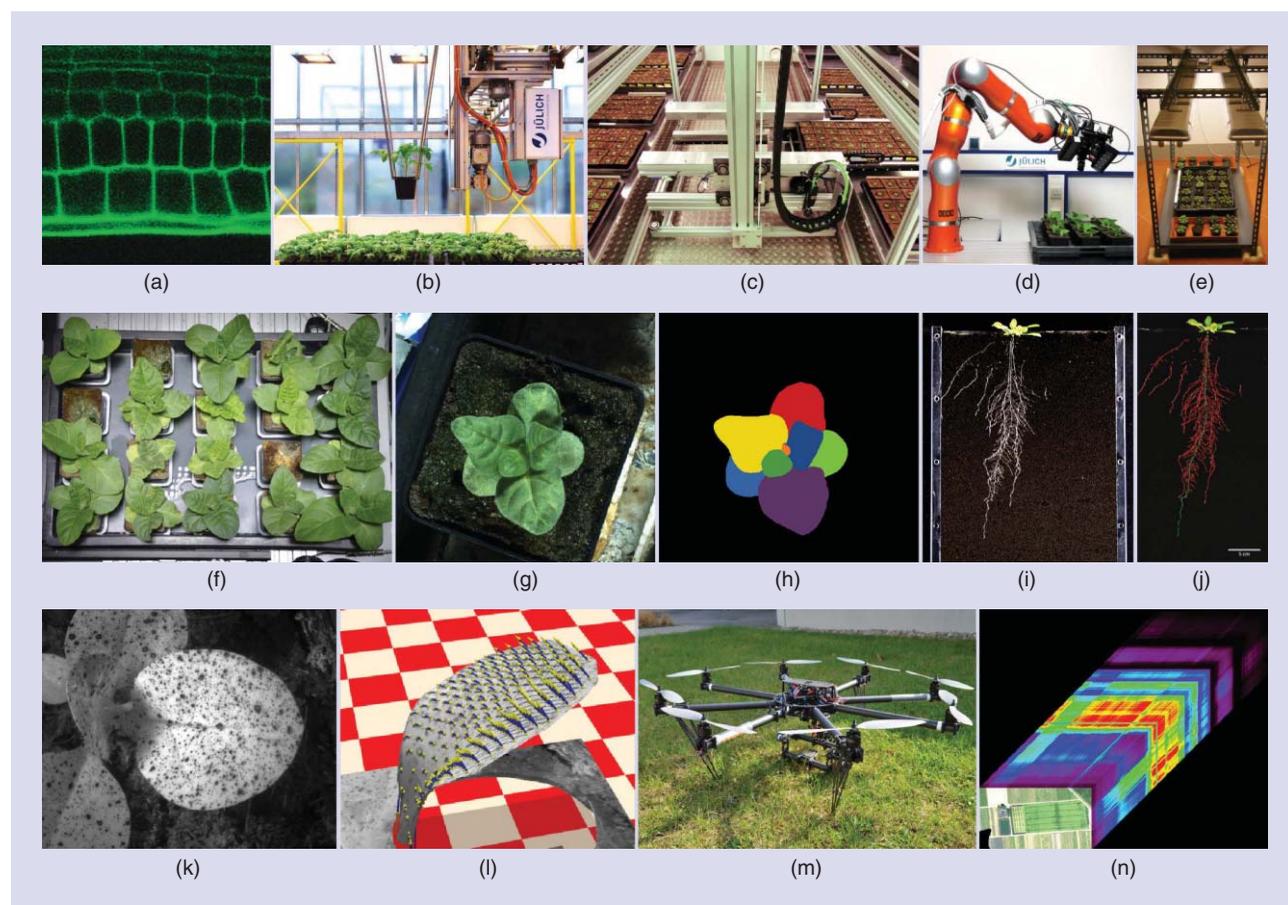
Figure 1(g) and (h)]; however, even though all the objects share a wide range of features (e.g., they are mostly green with similar brightness distributions), they show rich variations. Leaves differ in size over several orders of magnitude, introducing a structural break due to resolution limitations, and algorithms need to deal with leaves emerging in the scene.

In addition, leaves vary in shape, and while they do share a certain basic shape, they overlap, bend, and vary in pose. Even for the same species, leaves may differ substantially, as leaf shape, size, color, and overall appearance of a plant depend on the genotype (e.g., there are thousands of mutants available for Arabidopsis alone), environmental factors (drought, low or high light, and temperature), and the age

of each leaf. Readily apparent approaches based on learning shape from a labeled data set reveal their limitations when having to deal with such shape diversity and different acquisition conditions. While counting and segmenting leaves from such images can be simple for a human, no automated algorithmic solution is yet available that comes close to human performance.

WHOLE PLANTS: ANATOMICALLY CORRECT 3-D GEOMETRIC MODELING

For larger plants, reconstruction from a single image and viewpoint is not sufficient. Most approaches aim at obtaining an as complete three-dimensional (3-D) shape reconstruction as possible, geometrically



[FIG1] An example of plant phenotyping setups and images. (a) At the smallest scale, microscopy can image cells. (b)–(d) Movable imaging setups [15] or a (e) single overview camera setup can be used to image (f) many plants or (g) single plants, where (h) leaf segmentation is a sought-after outcome yielding growth measurements. Roots can be imaged in (i) rhizotrons [15], requiring (j) delineation. (k)–(l) Optical flow tracking [15] can measure finer leaf level growth. Airborne vehicles, e.g., (m) drones, can provide information on fields, e.g., (n) hyperspectral images [15]. [(a) is adapted from [7] and reproduced by permission of Elsevier. (b), (c), and (m) are courtesy of Alexander Putz, (i) and (j) are courtesy of Kerstin Nagel, and (n) is courtesy of Uwe Rascher.]

modeling the overall above-ground part of a plant, i.e., the shoot. However, details of parts are also investigated, such as grains on an ear (e.g., of corn), berries on wine grapes, flower development, etc. Imaging becomes more and more automated using conveyor-belt or robotized systems [see Figure 1(b)–(d)], allowing high throughput with thousands of plants. Automation of image analysis is then a must.

A variety of 3-D measuring strategies is currently being investigated, e.g., correspondence-based triangulation methods, silhouette-based carving, time-of-flight cameras, or light detection and ranging laser scanning (see [8] for a comprehensive overview). Setups are usually tailored to a particular species and conditions. This is, for example, due to size and image resolution constraints, or self-occlusion and self-similarity hampering triangulation.

A major challenge for all 3-D measuring methods is plant motion during acquisition. Time delays due to scanning or sequential image acquisition lead to notable geometric distortions, especially for outdoor measurements with wind. The data then cannot be described by a static model and all current approaches doing so fail one way or another.

From the 3-D data, quantitative information about plant traits need to be extracted. Simple summary traits, such as covered volume or plant height, could be estimated from images alone without 3-D reconstruction. But organ-wise traits, e.g., accurate leaf size or branching angle, require interpretation of 3-D data and plant part models. Simple models are used today (e.g., fitting two-dimensional surfaces to patches and merging them), but for most species new anatomically correct models are required.

WHOLE PLANTS BELOW GROUND: CLUTTERED IMAGES OF ROOTS

It is not possible to look through soil with the naked eye. Thus, classical root system analysis is invasive, meaning that plants are dug out and the roots washed and imaged. Usual image analysis then applies threshold-based segmentation, connected component labeling, and skeletonization, followed by estimation of traits such as overall graph length, branching angles,

and others. All solutions available to date have only limited effectiveness when root systems are heavily entangled. Obviously, no time-series analysis can be performed when plants are dug out.

In soil, roots can be imaged noninvasively using so-called rhizotrons [9], i.e., flat pots with large vertical windows, such that parts of the roots visibly grow along the window [see Figure 1(i) and (j)]. In dark soil and at high spatial resolution, segmentation of bright roots may be done with solutions developed, e.g., for angiograms in medicine; but under realistic conditions this is difficult: even with high-resolution cameras (in the 30 megapixel range) fine roots may be only few pixels wide, blurred and with poor contrast to the surrounding soil. Many current segmentation solutions are slow or even break down when applied to such large images. Thus, computational efficiency is an issue. In addition, windows can get scratched by frequent use and soil contains all sorts of clutter. To date, reliable segmentation of such images can only be done semiautomatically, requiring user assistance. Even learning-based methodologies yield unimpressive results, which point to the need for finding (or learning) better feature representations.

Using penetrating radiation or modalities such as MRI, PET, and X-ray computed tomography, roots can be imaged in soil in 3-D, where different imaging techniques yield complementary contrast information and metabolic function (e.g., with PET). Challenges are similar to medical applications including proper (co)registration of time series of deforming objects of potentially different modalities, disentangling objects, measuring geometric traits, etc. However, artifacts and structures are different.

ADDING DYNAMICS: TRACKING, FLOW, AND GROWTH ESTIMATION

For many plant traits, temporal dynamics are of high relevance. Growth analyses on the local tissue level are typically performed on image sequences with frame rates in the range of one per minute. A long-established technique restricts the leaf of interest to a plane by pulling it flat and images it using a single camera.

Growth is then calculated as divergence of an estimated optical flow field. Unfortunately, with this simple engineering solution, gene expression analyses have shown that “tension-stress genes” are turned on during such experiments, and thus the observed growth may be influenced on the molecular level. For non-fixed leaves moving in 3-D, calculating scene flow from multicamera “light-field” image sequences has been investigated [see Figure 1(k) and (l)]. This allows precise translation and rotation field estimation. Local growth can also be estimated from divergence, however, signal-to-noise-ratio is relatively poor. To date, no reliable local growth measuring technique without fixating leaves is available.

When aiming for growth analysis (in terms of summary growth over an organ), segmentation or reconstruction techniques as described earlier are needed. For simple plant architectures, e.g., young tobacco with up to eight leaves, leaf-wise tracking in temporally sufficiently high-resolved data sets has been demonstrated [10]. No reliable method for leaf-wise tracking has been reported in the literature so far for when time intervals become larger, or plant complexity is higher.

THE GREENHOUSE, FIELD, AND FARM: MORE VARIABILITY

While experiments in the laboratory do advance our knowledge of biological systems and their functioning, ultimately phenotyping must translate the knowledge to the society and stakeholders, such as breeders and farmers [3]. Phenotyping investigations must then be conducted under “real” (or realistic) conditions in the greenhouse or field, on crops that carry agricultural importance, such as corn, wheat, rice, barley, etc. [11].

Starting with the greenhouse, automated systems that are able to water and image plants, either move the plant to the imaging station or move the imaging apparatuses to the plants. Independent of setting, any positioning differences, either of the camera or the plant, radically complicate the process of establishing temporal correspondences between consecutive measurements. Taking the imaging apparatus

outside and in the field introduces additional challenges. Several approaches exist that mount sensors on specialized carriers: human-controlled tractors or other ground vehicles, or in the air with unmanned aerial vehicles [Figure 1(m)] operated either remotely or in an automated fashion. Image data differ tremendously in resolution, detail, motion blur, or clutter, severely affecting subsequent analysis tasks, thus, more robust algorithms are necessary. Computational efficiency is an issue, as the amount of imaging data produced is enormous [cf. Figure 1(n)], and analysis tasks can be significantly complex. Efforts in directly using analysis results for cultivation practices are the central theme in precision agriculture [3], which aims at tailoring treatment at the individual plant level. Thus, computer vision becomes crucial in supporting the whole process and evidently there is now the additional challenge of identifying low-complexity approaches to robust vision.

AFFORDABILITY: COPING WITH RESTRICTIONS

Currently, most versatile solutions are too expensive, and many labs instead develop highly customized (hardware and image analysis) solutions tailored to their experimental setting that are capable of addressing only specific phenotyping problems. Even when they are affordable,

this variability in methods and setups creates standardization problems.

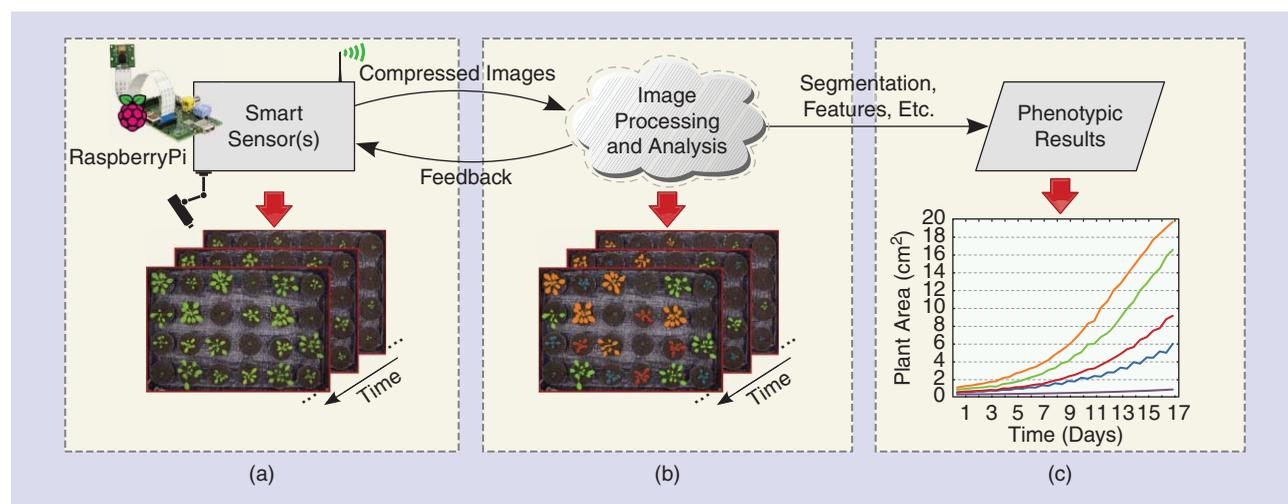
The use of off-the-shelf commercial equipment (such as commercial cameras [12] or the Kinect [5]) could facilitate standardization across experiments, lower the entry barrier, offer affordable solutions, and help many labs adopt the image-based approach to plant phenotyping.

Our recent project [16] aims to provide a universal turnkey and modular platform based on a distributed sensing and analysis framework [13], as shown in Figure 2. This distributed approach presents several key advantages. Affordable and easy-to-install sensors can be deployed in laboratories (growth chambers), the greenhouse, or the field to cover wide areas, before resorting to more costly and complex solutions based on robotics and automation. It is easy to become accustomed to a cloud-based storage and analysis application that is always up to date. It relieves users from maintaining a computing infrastructure and, importantly, it also permits consistency in experiments among different labs by standardizing equipment and analysis.

This centralized design, particularly when combined with an open architecture, can benefit the entire community, providing a modular and expandable architecture (by changing or adding new

camera sensors), favoring software reuse (e.g., user-contributed algorithms can be adopted by other labs), and knowledge sharing (e.g., a common repository of acquired data and meta-data, and also the analysis application itself learning on the user's feedback).

Affordability and remote processing, however, pose technical challenges. The choice of optics and the fixed field of view restrict the quality (in resolution and sharpness) of the acquired images and the plants this setup can image (e.g., it may not be suitable for not coplanar plants). An affordable sensor will have limited computational power and knowledge access, thus, it requires low-complexity algorithms to perform some of the tasks outlined in previous sections, and as such remote processing is necessary. Then the transmission of (possibly) large volumes of image data necessitates compression to meet bandwidth constraints. While this loss of information will affect the accuracy of the analysis algorithm, recent advances in application-aware compression can tune compression parameters to meet analysis accuracy needs [13], [14]. From a software engineering perspective, backward compatibility of the analysis framework and of the computational backbone has to be ensured, such that experimental protocols and results obtained previously remain valid.



[FIG2] (a) Affordable camera sensors (e.g., based on the Raspberry Pi [17]) acquire time-lapse sequences of the scene, including one or multiple plants. (b) Images are compressed and transmitted to the cloud, where high computational power and a broad knowledge base enable sophisticated computer vision tasks (e.g., leaf segmentation and tracking, optical flow analysis). Additionally, information is fed back to the sensor. Relying on Web-based graphical user interfaces, (c) phenotyping results are presented to the user for interpretation.

A TIMELY AND UNIQUE CHALLENGE

A quantitative description of plant phenotypes is a key ingredient for a knowledge-based bioeconomy, and this not only literally helps in the efforts to feed the world but is also essential for fiber and fuel production, the so-called Green Revolution 2.0. In fact, comparing the “Top 10 list of Emerging Technologies” in 2012 according to the world economic forum, the top 1, 2, 3, and 5 technologies are directly addressed by plant phenotyping research [18]. Recently, we have even witnessed direct investments in helping the translation of agricultural technology in farming. For example, Farm2050 [19] includes the information extraction powerhouse Google and drone company 3-D Robotics among its partners.

There is not only growing interest from the application side, both scientifically and commercially, but exciting computer vision and image processing problems exist that differ from other biomedical applications. While medicine focuses on the status of a single species (i.e., humans) in a diagnostic capacity, plant phenotyping addresses a large number of different plant species with hundreds to thousands of genotypes (cultivars) per species, usually in group-wise experiments. It addresses the development over time in addition to static snapshots and under a wide range of environmental conditions, using various imaging setups (as opposed to medical imaging where predefined protocols are in place and equipment variability is relatively limited). Thus, even within a single application, diverse conditions need to be

addressed, to ascertain a robust image-based measurement of the phenotypic trait. Plant phenotyping at a high throughput requires reliable image processing algorithms that could batch process many data accurately, and an integration with genetic databases and other frameworks.

The previous sections outlined a series of challenges (e.g., dealing with structural breaks in tracking/detection), for which our community can get involved. In this article, although we focus on extracting information from images, data mining and combing the information from genotyping, environmental, and phenotyping sources are by themselves a big undertaking as well. Jointly, we must make the effort to solve these problems and push the envelope further, and by including the resources in Table 1, we hope to help facilitate this. We need to cooperate with different disciplines to integrate expertise across the spectrum and provide biologically or agronomically meaningful and technically robust solutions [3], [7] to help resolve this bottleneck.

[TABLE 1] GET INVOLVED: A COLLECTION OF ONLINE RESOURCES.

NAME	DESCRIPTION
SPECIALIZED SCIENTIFIC EVENTS	
CVPPP2014 [20] AND CVPPP2015 [21]	THE FIRST AND SECOND WORKSHOP FOR COMPUTER VISION PROBLEMS IN PLANT PHENOTYPING, ALSO HOSTING CHALLENGES SUCH AS THE LEAF SEGMENTATION CHALLENGE (LSC)
IAMPS [22]	INTERNATIONAL WORKSHOP ON IMAGE ANALYSIS METHODS FOR THE PLANT SCIENCES
PHENODAYS [23]	INTERNATIONAL SYMPOSIUM INVOLVING SEED INDUSTRY, BREEDING INSTITUTES, AND ACADEMIC BREEDING GROUPS
IPPS [24]	INTERNATIONAL PLANT PHENOTYPING SYMPOSIUM
ICPA [25]	INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE
IMAGE DATABASES	
LSC CHALLENGE [26]	IMAGES AND LEAF-BASED SEGMENTATION MASKS AS PART OF THE FIRST LSC CHALLENGE
MAIZEGDB [27]	IMAGES OF MAIZE
CWFID [28]	THE CROP/WEED FIELD IMAGE DATA SET (CWFID) CONTAINS IMAGES WITH CROP/WEED DELINEATIONS FOR A CLASSIFICATION TASK IN PRECISION AGRICULTURE
PHENOPSIS DB [29]	ARABIDOPSIS THALIANA PHENOTYPING DATABASE
CONSORTIA AND ORGANIZATIONS	
IPLANT COLLABORATIVE [30]	CONNECT SCIENTISTS TO PUBLIC DATA SETS, MANAGE AND STORE THEIR DATA AND EXPERIMENTS, ACCESS HIGH-PERFORMANCE COMPUTING, ETC.
IPPN [31]	INTERNATIONAL PLANT PHENOTYPING NETWORK
EPPN [32]	EUROPEAN PLANT PHENOTYPING NETWORK
EPSO [33]	EUROPEAN PLANT SCIENCE ORGANISATION
FESPB [34]	FEDERATION OF EUROPEAN SOCIETIES OF PLANT BIOLOGY
IEEE RAS [35]	AGRICULTURAL ROBOTICS AND AUTOMATION
ISPA [36]	INTERNATIONAL SOCIETY OF PRECISION AGRICULTURE
E-AGRICULTURE [37]	ICT FOR SUSTAINABLE AGRICULTURE
BSA [38]	BOTANICAL SOCIETY OF AMERICA, LISTING FURTHER PLANT SOCIETIES AND ORGANIZATIONS [39]
SOFTWARE DATABASES	
PLANT IMAGE ANALYSIS [40]	THIS DATABASE CURRENTLY PROVIDES A COLLECTION OF APPROXIMATELY 120 ANALYSIS TOOLS

ACKNOWLEDGMENTS

This work was partially supported by a Marie Curie Action: “Reintegration Grant” (grant number 256534) of the EU’s Seventh Framework Programme. We also thank those who shared images and provided insight and feedback on this article.

AUTHORS

Massimo Minervini (m.minervini@imtlucca.it) is a research associate with IMT Institute of Advanced Studies in Lucca, Italy.

Hanno Scharf (h.scharf@fz-juelich.de) is a senior scientist and the head of Quantitative Image Analysis at Forschungszentrum Jülich, IBG-2, Germany.

Sotirios A. Tsafaris (s.tsafaris@imtlucca.it) is an assistant professor and director of the Pattern Recognition and Image Analysis unit at IMT Institute of Advanced Studies in Lucca, Italy, and an adjunct professor with the Department of Electrical Engineering and Computer Science at Northwestern University, Illinois.

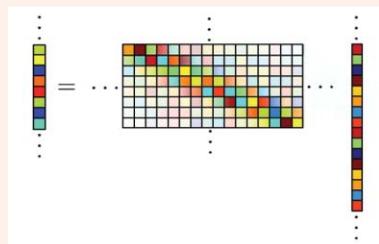
REFERENCES

- [1] D. Houle, D. R. Govindaraju, and S. Omholt, "Phenomics: The next challenge," *Nature Rev. Genet.*, vol. 11, no. 12, pp. 855–866, 2010.
- [2] S. Dhondt, N. Wuyts, and D. Inzé, "Cell to whole-plant phenotyping: The best is yet to come," *Trends Plant Sci.*, vol. 18, no. 8, pp. 433–444, 2013.
- [3] F. Fiorani and U. Schurr, "Future scenarios for plant phenotyping," *Annu. Rev. Plant Biol.*, vol. 64, no. 1, pp. 267–291, 2013.
- [4] R. C. O'Malley and J. R. Ecker, "Linking genotype to phenotype using the arabidopsis unimutant collection," *Plant J.*, vol. 61, no. 6, pp. 928–940, 2010.
- [5] L. Li, Q. Zhang, and D. Huang, "A review of imaging techniques for plant phenotyping," *Sensors*, vol. 14, no. 11, pp. 20078–20111, 2014.
- [6] E. P. Spalding and N. D. Miller, "Image analysis is driving a renaissance in growth measurement," *Curr. Opin. Plant Biol.*, vol. 16, no. 1, pp. 100–104, 2013.
- [7] T. P. Pridmore, A. P. French, and M. P. Pound, "What lies beneath: Underlying assumptions in bio-image analysis," *Trends Plant Sci.*, vol. 17, no. 12, pp. 688–692, 2012.
- [8] M. P. Pound, A. P. French, E. H. Murchie, and T. P. Pridmore, "Automated recovery of 3D models of plant shoots from multiple colour images," *Plant Physiol.*, 2014.
- [9] K. Nagel, A. Putz, F. Gilmer, K. Heinz, A. Fischbach, J. Pfeifer, M. Faget, S. Blossfeld, M. Ernst, C. Dimaki, B. Kastenholz, A.-K. Kleinert, A. Galinski, H. Scharr, F. Fiorani, and U. Schurr, "GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons," *Funct. Plant Biol.*, vol. 39, no. 11, pp. 891–904, 2012.
- [10] B. Dellen, H. Scharr, and C. Torras, "Growth signatures of rosette plants from time-lapse video," *IEEE/ACM Trans. Computat. Biol. Bioinform.*, 2015, DOI: DOI: 10.1109/TCBB.2015.2404810
- [11] C.-R. Shyu, J. Gree, D. P. K. Lun, T. Kazic, M. Schaeffer, and E. Coe, "Image analysis for mapping immeasurable phenotypes in maize," *IEEE Signal Processing Mag.*, vol. 24, no. 3, pp. 115–118, 2007.
- [12] S. A. Tsafaris and C. Noutsos, "Plant phenotyping with low cost digital cameras and image analytics," in *Information Technologies in Environmental Engineering*. New York: Springer, 2009, pp. 238–251.
- [13] M. Minervini and S. A. Tsafaris, "Application-aware image compression for low cost and distributed plant phenotyping," in *Proc. Int. Conf. Digital Signal Processing (DSP)*, 2013, pp. 1–6.
- [14] M. Minervini, C. Rusu, and S. A. Tsafaris, "Unsupervised and supervised approaches to color space transformation for image coding," in *Proc. Int. Conf. Image Processing (ICIP)*, 2014, pp. 5576–5580.
- [15] [Online]. Available: http://www.fz-juelich.de/ibg/ibg-2/EN/methods_jppc/methods_node.htm
- [16] [Online]. Available: <http://www.phenotiki.com>
- [17] [Online]. Available: <http://www.raspberrypi.org>
- [18] [Online]. Available: <http://reports.weforum.org/global-agenda-council-2012/councils/emerging-technologies/>
- [19] [Online]. Available: <http://www.farm2050.com>
- [20] [Online]. Available: <http://www.plant-phenotyping.org/CVPPP2014>
- [21] [Online]. Available: <http://www.plant-phenotyping.org/CVPPP2015>
- [22] [Online]. Available: <http://www.phytsystems.ulg.ac.be/iamps15>
- [23] [Online]. Available: <http://www.phenodays.com>
- [24] [Online]. Available: http://www.plant-phenotyping.org/chennai_2014
- [25] [Online]. Available: <https://www.ispag.org/ICPA>
- [26] [Online]. Available: <http://www.plant-phenotyping.org/CVPPP2014-dataset>
- [27] [Online]. Available: <http://www.maizegdb.org/cgi-bin/imagebrowser.cgi>
- [28] [Online]. Available: <https://github.com/cwfid/dataset>
- [29] [Online]. Available: <http://bioweb.supagro.inra.fr/phenopsis/>
- [30] [Online]. Available: <http://www.iplantcollaborative.org>
- [31] [Online]. Available: <http://www.plant-phenotyping.org>
- [32] [Online]. Available: <http://www.plant-phenotyping-network.eu>
- [33] [Online]. Available: <http://www.epsoweb.org>
- [34] [Online]. Available: <http://www.fespb.org>
- [35] [Online]. Available: <http://www.ieee-ras.org/agricultural-robotics>
- [36] [Online]. Available: <https://www.ispag.org>
- [37] [Online]. Available: <http://www.e-agriculture.org>
- [38] [Online]. Available: <http://botany.org>
- [39] [Online]. Available: <http://botany.org/outreach/international/>
- [40] [Online]. Available: <http://www.plant-image-analysis.org>

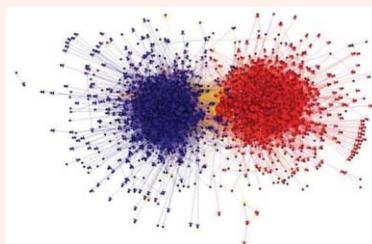
SP

SigView.org Multimedia Tutorials

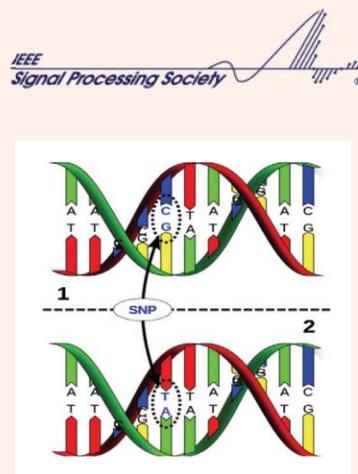
- ❖ Check out tutorials by leading signal processing experts
- ❖ Enable IEEE SPS members to create, host, and share multimedia tutorials from existing slides deck and media



Fundamentals of Compressive Sensing by Mark Daveport



DSP on Graphs by José Moura



Big Data and Machine Learning in Cancer Genomics by Ali Bashashati

- ❖ Missed ICASSP? Check out ICASSP Plenary talks since 2010 on SigView
- ❖ Stay tuned for recent talks from the School of ICASSP 2015, IEEE SPS Distinguished Lectures, and more

Digital Object Identifier 10.1109/MSP.2015.2434233

Understanding Game Theory via Wireless Power Control

In this lecture note, we introduce the basic concepts of game theory (GT), a branch of mathematics traditionally studied and applied in the areas of economics, political science, and biology, which has emerged in the last 15 years as an effective framework for communications, networking, and signal processing (SP). The real catalyst has been all of the blooming issues related to distributed networks in which the nodes can be modeled as players in a game competing for system resources. Some relevant notions of GT are introduced by elaborating on a simple application in the context of wireless communications, notably the power control in an interference channel (IC) with two transmitters and two receivers.

RELEVANCE

Recently, the mathematical tools of GT [1] have attracted a significant interest by the wireless communications and SP engineering communities [2, Part II] due to the need for designing autonomous, distributed, and flexible systems in which the available resources are allocated through low-complexity and scalable procedures. Games are appealing, owing to some characteristics that are not common in classical optimization: as an example, GT can handle interactive situations in which each player can only have a partial control over the optimization variables while using its own performance metric. It is true that commonalities can be found with other disciplines, such as multiobjective optimization [3], convex optimization [4], and learning theory [5],

but GT possesses many distinguishing features that make it essential for the standard current toolbox of communications as well as SP engineers.

PREREQUISITES

The readers require basic knowledge in linear algebra, wireless communications, and signal processing theory.

WHAT IS A GAME?

To take advantage of GT and its associated theoretical tools, the first step is to model the problem at hand as a game. In doing so, three ingredients must be identified:

- *players* who represent the main actors in the problem, having conflicting interests and affecting the performance of everyone else in the game
- *a set of strategies* available to each player that determines what each player can do
- *a utility function* for each player that measures its degree of satisfaction as a function of the combination of all player's choices.

This description may encompass a large number of situations: to mention a few examples, players in a game can be base stations (BSs) allocating the resources in a cellular network to increase the system throughput, or watermarking devices choosing algorithms to face potential attackers.

The objective of the modeling effort is to describe the game using its strategic-form representation: a triplet $\langle \mathcal{K}, \{S_k\}_{k \in \mathcal{K}}, \{u_k\}_{k \in \mathcal{K}} \rangle$, where

$$\mathcal{K} = \{1, \dots, K\}$$

is the set of players, where K is the number of players; S_k is the set of strategies for each player k ; and $u_k(\mathbf{s})$ is the utility function (also known as *reward* or *payoff*)

associated to player k for a combination of choices $\mathbf{s} = [s_1, \dots, s_K] = [s_k, \mathbf{s}_{\setminus k}]$, where $\mathbf{s}_{\setminus k} = [s_1, \dots, s_{k-1}, s_{k+1}, \dots, s_K]$ denotes the strategies taken by all other players except player k (the *opponents*).

In general, the game outcome $u_k(\mathbf{s})$ for player k depends on all players' choices through \mathbf{s} , which stems out from the interaction of the players with possibly conflicting interests. This brings forth a couple of distinguishing features of GT:

- each player k can have a different performance metric; this feature is captured by a per-player specific function $u_k(\mathbf{s})$, which accounts for the player's nature
- each player k has partial control ($s_k \in S_k$ only) over the optimization variables.

The first property is strictly tied with multiobjective optimization [3], although a clear difference exists in the scope of the optimization variables, as in multiobjective optimization we have full control over all variables. The second property is tightly related to the framework of distributed optimization [4], with which it shares many intersections, although there are specific differences: one of the most important is that, while in distributed optimization the agents follow some common given rules, in GT the players act as independent decision makers.

THE NEAR-FAR EFFECT GAME

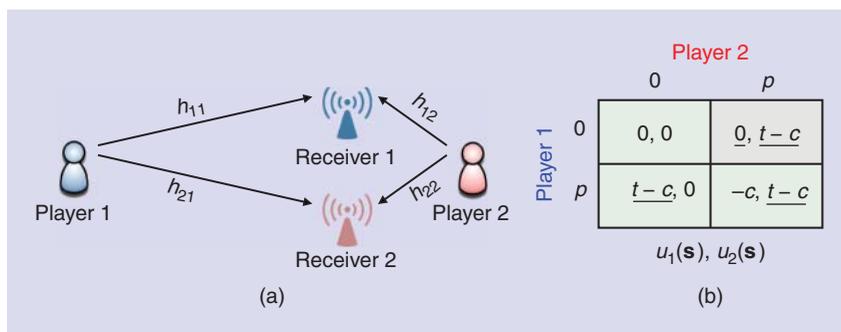
To picture the meaning of the strategic-form representation, let us consider an example taken from a very common wireless communications scenario: the IC, represented in Figure 1(a), in which the two transmitters interfere with each other in the attempt to reach their own receiver. This simple scheme encompasses many scenarios: it can be used to model

- a multicellular system with red and blue nodes belonging to two different cells
- a heterogeneous network, where the red and blue nodes belong to a macrocell and a small cell, respectively
- a cognitive radio system, where the red and blue nodes are primary users (PUs) and secondary users (SUs), respectively
- a device-to-device system, where the receivers are also network nodes.

Using GT, we can model the problem at hand in a suitable manner and provide the theoretical tools to solve it. In this case, solving means devising the optimal transmission strategy to be selected by the two wireless terminals of Figure 1(a). In particular, we assume that the two nodes are allowed either to transmit at a certain power level p or to stay idle. This situation can be modeled as a game, with $K = 2$ players and strategy sets $\mathcal{S}_k = \{0, p\}$ for $k \in \mathcal{K} = \{1, 2\}$. For simplicity, we also assume that the two terminals choose their strategies simultaneously (i.e., without being informed of the other's choice) once and for all (i.e., they cannot make any changes after observing the outcome of the game)—in GT parlance, we call this a *static game*. Finally, since players and strategy sets are both countable, the game is termed *finite*.

As depicted in Figure 1(a), player 1 (the far terminal) is located much farther away from both receivers than player 2 (the near terminal). To describe this situation in a mathematical fashion, we introduce the power gains $h_{jk} \in \mathbb{R}^+$ experienced by terminal k 's signal when propagating to receiver j . For simplicity, let us assume $h_{jk} = h_k$ for $j = 1, 2$ with $h_1/h_2 \ll 1$ (we will better quantify this ratio at the end of this section), thereby giving rise to the so-called near-far problem.

We now need to define a utility function, and to do so we consider that each terminal achieves a degree of satisfaction that depends both on the success of its transmission and on the energy spent to transmit at power s_k . Mathematically, this translates into a (dimensionless) utility $u_k(\mathbf{s}) = t_k(\mathbf{s}) - c_k(\mathbf{s})$, where $t_k(\mathbf{s})$ accounts for the outcome of the transmission and $c_k(\mathbf{s})$ measures the cost



[FIG1] A representation of the near-far effect (NFE) game. (a) The topology of the network. (b) The payoff matrix.

associated with using s_k . We assume that the cost scales linearly with the transmit power, and that it is independent of the other terminal's strategy: $c_k(\mathbf{s}) = cs_k/p$. Measuring $t_k(\mathbf{s})$ is more complicated, as it has to capture the interaction between the players as a function of the selected strategies s_1 and s_2 . In practice, successful reception of a signal in a multiple-access scenario (such as the one considered here), be it in the time, frequency, space, or code domain, depends on the signal-to-interference-plus-noise ratio (SINR) γ_k , which measures the ratio of the useful received signal power to the amount of undesired power collected at the receiver. Under the assumption of additive white Gaussian noise (AWGN) with power σ^2 , we get

$$\gamma_k(\mathbf{s}) = \frac{\Gamma h_k s_k}{\sigma^2 + h_{\setminus k} s_{\setminus k}} = \mu_k(s_{\setminus k}) s_k \geq 0, \tag{1}$$

where $\setminus k = 2$ if $k = 1$, $\setminus k = 1$ if $k = 2$, and $\Gamma \geq 1$ is the processing gain, which depends on the multiple access technology and the receiver processing. For the time being, let us assume the transmission to be successful if and only if $\gamma_k \geq \gamma_{req}$, where the minimum SINR γ_{req} depends on some system parameters. So, when $\gamma_k < \gamma_{req}$, the transmitted message cannot be decoded at receiver k , and $t_k(\mathbf{s}) = 0$. On the contrary, when $\gamma_k \geq \gamma_{req}$, receiver k can correctly receive the information associated to user k 's signal, and $t_k(\mathbf{s}) = t$, where t is a dimensionless parameter that accounts for the throughput achieved at destination. To properly capture the cost-benefit analysis that regulates any practical wireless system, it makes sense to assume $t \gg c$.

A profitable way to investigate finite static games in their strategic form, such as our NFE game, is through the so-called payoff matrix [Figure 1(b)] in which player 1's strategies are identified by the rows, player 2's strategies by the columns, and the entries of the matrix (the pair of numbers in the box) represent the utilities ($u_1(\mathbf{s}), u_2(\mathbf{s})$) achieved by the players. Under the assumptions that $p = \sigma^2 \gamma_{req} / (h_1 \Gamma)$ and $(h_1/h_2) < 1/(1 + \gamma_{req} \Gamma)$ (see [6] for more details), it is easy to fill out each box of the matrix based on the hypotheses listed above.

Once the game is in its strategic form, we have to solve it, i.e., to predict its outcome. In the NFE game, we assume that both players

- are rational
- only control their own strategies
- know each other's payoff.

The first assumption means that each player is a utility-maximizer decision maker. The second hypothesis casts this problem as a noncooperative game in which the players compete to unilaterally maximize $u_k(\mathbf{s})$. Finally, the third hypothesis involves the concept of complete information that each player has about the game. By inspecting the payoff matrix in Figure 1(b), it is apparent that player 2's best strategy is represented by $s_2^* = p$ whatever s_1 is, since $t - c > 0$ under the assumption $t \gg c$. For this reason, the strategy $s_2 = 0$ is said to be strictly dominated by $s_2 = p$, as $u_2([s_1, 0]) < u_2([s_1, p]) \forall s_1 \in \mathcal{S}_1$. This is known to player 1 as well, who rationally chooses to play $s_1^* = 0$. As a conclusion, the predictable outcome of the NFE game is $\mathbf{s}^* = [s_1^*, s_2^*] = [0, p]$, as

lecture NOTES continued

		Player 2	
		0	p
Player 1	0	0, 0	$0, \underline{t-c}$
	p	$\underline{t-c}, 0$	$-c, -c$
		$u_1(\mathbf{s}), u_2(\mathbf{s})$	

[FIG2] The payoff matrix of the IC game.

highlighted by the shaded box in Figure 1(b). In GT parlance, this game has been solved by applying the iterated elimination of dominated strategies, or iterated dominance for short [1, Ch. 1].

THE IC GAME

Let us now slightly modify the scenario represented in Figure 1(a). Assume, for example, that player 1 is moved closer to its receiver, such that the distance between player 1 and both receivers becomes the same as the distance between player 2 and both receivers. For simplicity, let us also suppose $h_{jk} = h$ for $j, k = 1, 2$. By using $p = \sigma^2 \gamma_{\text{req}} / (h\Gamma)$, following the same considerations taken for the NFE game, it is easy to obtain the payoff matrix reported in Figure 2. As an exercise, we can verify that no strictly dominated strategies exist, and thus we cannot apply the iterated dominance procedure used to solve the NFE game.

To get out of this impasse, we introduce the concept of best response (BR) $b_k(\mathbf{s}_{\setminus k})$, which is mathematically defined as

$$b_k(\mathbf{s}_{\setminus k}) = \underset{s_k \in S_k}{\operatorname{argmax}} u_k([s_k, \mathbf{s}_{\setminus k}]), \quad (2)$$

i.e., the best that we can get out of the game once we know the opponents' moves $\mathbf{s}_{\setminus k}$. Since player 1 chooses rows, we can compute its BR by examining the columns that can be possibly selected by player 2. When $s_2 = 0$, $b_1(s_2 = 0) = p$. Conversely, when $s_2 = p$, $b_1(s_2 = p) = 0$. The same can be obtained for player 2, and we end up with the players' BRs, underlining the relevant payoffs in Figure 2. We find two boxes (shaded background) containing the BRs of both players, representing two stable states, where *stable* here means that such states are attained by some multiple agents with conflicting interests that compete through self-optimization, and they eventually reach a point where none of them has any incentive from which to unilaterally deviate.

A point that possesses such properties is termed a *Nash equilibrium (NE)* of the game, which is defined as a strategy profile $\mathbf{s}^* = [s_k^*, s_{\setminus k}^*]$ such that, for all $k \in \mathcal{K}$,

$$u_k([s_k^*, \mathbf{s}_{\setminus k}^*]) \geq u_k([s_k, \mathbf{s}_{\setminus k}^*]) \quad \forall s_k \in S_k, \quad (3)$$

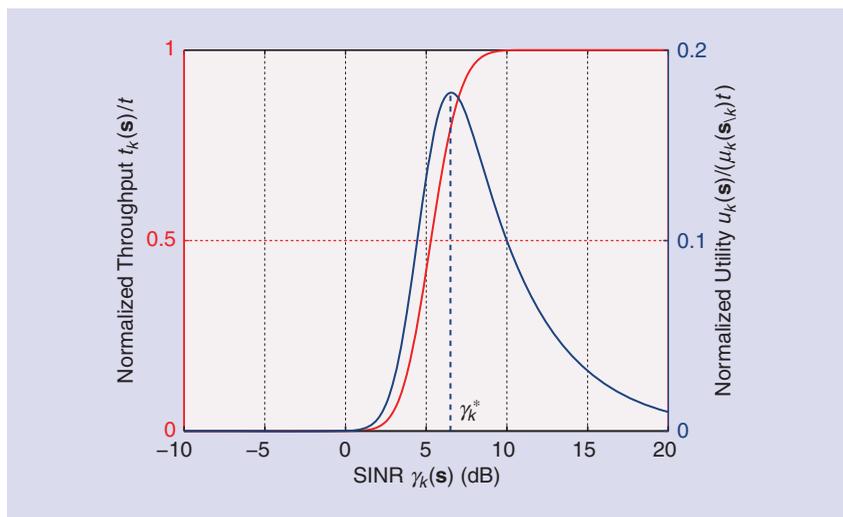
or, equivalently, $s_k^* \in b_k(\mathbf{s}_{\setminus k}^*)$. As an exercise, check that $\mathbf{s}^* = [0, p]$ is the unique NE of the NFE game.

The notion of NE encompasses many interpretations of GT (not discussed here for brevity) that interested readers can find in many textbooks (e.g., [7, Ch. 1]). Modeling the players as self-optimizing decision makers finds a suitable application especially in the context of SP in which the devices can be programmed to do so. Since each player has only a partial control of the game, the concept of NE is tightly coupled with the application of distributed algorithms and machine-learning techniques [7, Part II].

Due to space constraints in this lecture note, we will not discuss theorems on equilibrium existence [1, Ch. 1], which establish the existence of the NE in particular classes of games, and on equilibrium uniqueness [2, Ch. 3]. When uniqueness cannot be ensured, like in the case of the IC game, we face the problem of equilibrium selection. One solution to this issue is the concept of correlated equilibrium (CE) [1, Ch. 2], a generalization of the NE, where an arbitrator helps the players to correlate their strategies, so as to favor a decision process in the interplay, e.g., letting them adopt $\mathbf{s}^* = [p, 0]$.

INTRODUCING CONTINUOUS POWERS

The solutions of NFE and IC games, in which at most one terminal can successfully transmit, directly stem out of choosing a binary strategy set $S_k = \{0, p\}$ for both players. Let us see what happens if any power level in the continuous interval $[0, p]$ can be selected. This amounts to setting $S_k = \{s_k \in \mathbb{R} : 0 \leq s_k \leq p\}$. Within this setting, the power control problem can be studied as a continuous game [7, Ch. 2]. In our attempt to get closer to a realistic scenario, let us also modify the utilities to better model how real data networks work in practice. A good approximation for the effective throughput in a packet-oriented transmission is $t_k(\mathbf{s}) = t(1 - \exp\{-\gamma_k(\mathbf{s})\})^L$ [8], whose behavior is depicted (red line, left axis) in Figure 3. In our expression, L denotes the number of information bits per packet (here, $L = 20$), and t is the



[FIG3] The throughput (red) and utility (blue) as functions of the SINR (continuous NFE game).

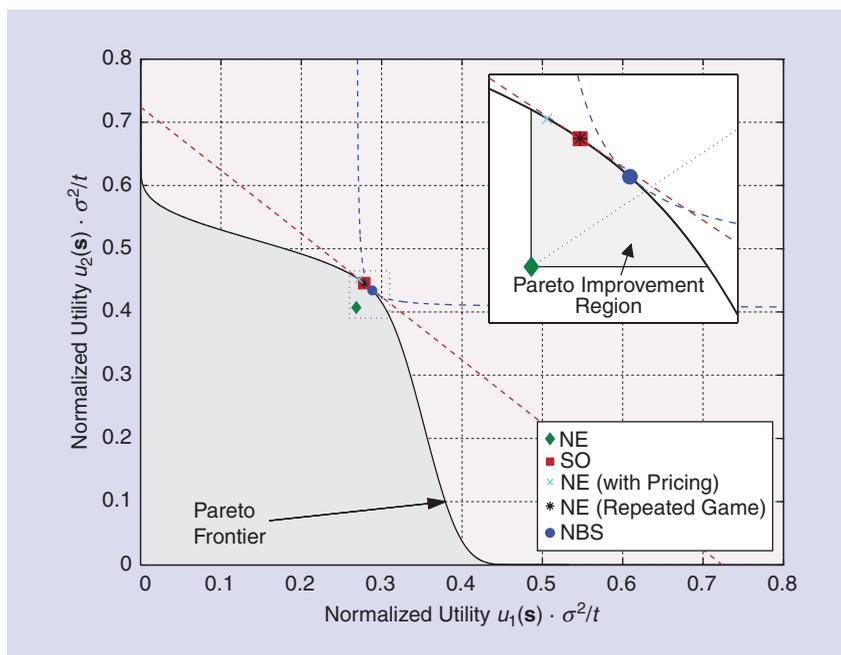
communication rate (in bits/second). To properly capture the tradeoff between obtaining a satisfactory throughput and saving transmit power—similar to what is considered for the NFE and IC games—we will adopt a “green” approach based on improving each player’s energy efficiency [8]. This can be done by defining player k ’s utility as the ratio between throughput and power expenditure, thus accounting for the number of bits correctly delivered per joule of energy consumed:

$$u_k(s) = t_k(s)/s_k \quad [\text{b/J}], \quad (4)$$

whose normalized behavior is reported (blue line, right axis) in Figure 3, with $\mu_k(s_{\setminus k})$ defined as in (1).

Using straightforward manipulation, player k ’s BR (2) turns out to be $b_k(s_{\setminus k}) = \min\{p, \gamma_k^*/\mu_k(s_{\setminus k})\}$, where γ_k^* is the “optimal” SINR such that $\partial t_k(s)/\partial \gamma_k(s) |_{\gamma_k(s)=\gamma_k^*} = (1/\gamma_k^*)t_k(s) |_{\gamma_k(s)=\gamma_k^*}$ [8]. For instance, when $L = 20$, we get $\gamma_k^* \cong 4.5 = 6.5$ dB for $k = 1, 2$ (see Figure 3). Based on such BR, the continuous IC game presents a unique NE, represented by the fixed point $s_k^* = b_k(s_{\setminus k}^*)$ for $k = 1, 2$ [8].

How can the NE be “visualized”? Let us consider a particular realization of the network sketched in Figure 1(a), with the following parameters: $h_{11}=0.75, h_{21}=0.25, h_{12}=0.5, h_{22}=1; \Gamma = 4; p/\sigma^2 = 5; \text{ and } L = 20$. (The MATLAB code for all examples presented in this lecture note is available for download in [9].) The solution of this game is given by the NE $s^*/\sigma^2 = [2.99, 1.97]$, yielding normalized utilities $(\sigma^2/t)u_1(s^*) = 0.269$ and $(\sigma^2/t)u_2(s^*) = 0.407$. We display in Figure 4 the (normalized) utilities at the NE (green diamond) on the bidimensional normalized utility plane, given by all achievable utility pairs $(u_1(s), u_2(s))$ (shaded region), for any strategy profile $s \in \mathcal{S}_1 \times \mathcal{S}_2$ (the utility plan can be found via a numerical search using [9]). Note that $s_1^* > s_2^*$, and $u_1(s^*) < u_2(s^*)$: this is due to the better channel conditions experienced by player 2 (both in the direct and the interference links), that make it achieve the optimal SINR γ_2^* with a lower power consumption than player 1. However, unlike the finite version of the



[FIG4] The normalized utility plan (continuous-power NFE game).

NFE game, where $s_1^* = 0$ (and thus $t_1(s^*) = 0$), now player 1 can successfully connect to its receiver, getting a throughput $t_1(s^*) \cong 0.8 t$ [the same as $t_2(s^*)$], at the cost of a slightly higher power consumption ($s_1^*/s_2^* \cong 1.52$), and thus with a lower energy efficiency.

IS THE NE EFFICIENT?

A natural question that arises regards the actual efficiency, or, the performance, of the NE: Is the NE efficient? To address this question, we first need to agree upon our performance metric. In GT, a convenient way to assess how desirable a solution is involves the concept of efficiency, evaluated in terms of Pareto optimality. A profile \bar{s} is Pareto-optimal (PO) if there exists no other s such that 1) $u_k(s) \geq u_k(\bar{s})$ for all $k \in \mathcal{K}$, and 2) $u_k(s) > u_k(\bar{s})$ for some $k \in \mathcal{K}$. In our continuous IC game, the performance achieved by the PO profile set is represented by the contour of the shaded area in Figure 4, which is called the *Pareto frontier*. Clearly, if we increase $u_1(s)$ (i.e., if we move rightward along x), then $u_2(s)$ decreases, and the same happens if we increase $u_2(s)$ (by moving upward along y). Still, Pareto optimality does not qualify as our performance metric.

We have to further introduce the notion of social welfare (SW), that is often used as a convenient measure for the efficiency of a strategy vector [7, Ch. 2]. Formally, the social-optimal (SO) profile \check{s} is the PO profile that maximizes the SW, defined as the weighted sum-utility $\sum_{k \in \mathcal{K}} w_k u_k(s)$, where the weights $\{w_k\}_{k \in \mathcal{K}}, \sum_{k \in \mathcal{K}} w_k = 1$, allow us to account for different classes of service: as an example, unequal weights can be useful to model PUs (higher w_k ’s) and SUs (lower w_k ’s) in a cognitive network. In our two-player game, we can identify \check{s} as the tangent point between the Pareto frontier and a line with slope $-w_1/w_2$. As an example, if we consider $w_1 = w_2 = 0.5$ (i.e., if the two players have the same priority), $\check{s}/\sigma^2 = [2.20, 1.55]$, thus yielding $(\sigma^2/t)u_1(\check{s}) = 0.278$ and $(\sigma^2/t)u_2(\check{s}) = 0.446$ (Figure 4).

We have zoomed in on a section of Figure 4 (see inset) to show that the NE s^* is socially inefficient since its performance is distinct from (more specifically, poorer than) that achieved by \check{s} . In general, \check{s} cannot be achieved by distributed algorithms, rather, it is the result of a global optimization and, in our case, it turns out to be unbalanced toward player 2 (the one with better channel conditions): this is

lecture NOTES continued

reminiscent of the waterfilling policy [4] that allocates most resources to the users who can achieve higher throughputs. More importantly, the magnification in Figure 4 shows that there are a multitude of profiles that provide utilities lying in the Pareto improvement region in which $u_k(s) \geq u_k(s^*)$ for all $k \in \mathcal{K}$ (the shaded region in the inset). Consequently, the next question is: How can we improve the efficiency of the NE? In this lecture note, we focus on three popular methods:

- modifying the utility functions
- letting the players interact more than once
- letting the players cooperate.

PRICING THE STRATEGIES

The simplest method to improve the efficiency of the NE while maintaining the game structure is by modifying the utility function. This can be done, for instance, by introducing some form of externality. This approach is, in spirit, close to mechanism design [1, Ch. 7]. For the power control games studied so far, one might think of charging the players for the powers they consume by introducing a pricing factor α (in $\text{b/J} \cdot \text{W}^{-1}$): $\tilde{u}_k(s) = u_k(s) - \alpha s_k = t_k(s)/s_k - \alpha s_k$ [10]. The rationale behind this approach is the following: if each transmitter is discouraged from being aggressive (due to power taxation), the multiple access interference (MAI) experienced by the other is reduced, and both SINRs at the equilibrium stay as close as possible to γ_k^* (provided that the AWGN power σ^2 is not dominant—a condition that always holds in multiple-access systems).

To evaluate the benefits of this method, we compute the NE \tilde{s}^* of the modified game using the BR approach (see [9]). The blue cross marker in Figure 4 represents the performance $((\sigma^2/t) u_1(\tilde{s}^*), (\sigma^2/t) u_2(\tilde{s}^*))$ of the original utility function (4) (which is the one we are actually interested in), using $\alpha = 0.12 t/\sigma^4$, which yields $\tilde{s}^*/\sigma^2 = [2.17, 1.57]$. As can be seen, the performance of \tilde{s}^* is very close to the Pareto frontier, further favoring player 2's performance compared to the SW. We have thus improved the efficiency of the game solution while maintaining the noncooperative nature of the interplay

(with all its desirable properties). The main drawback is that this improvement can only be achieved after a proper tuning of α (which highly depends on system parameters). As an exercise, one could evaluate the performance of \tilde{s}^* as a function of α (see [9]).

REPEATING THE GAME

The inefficiency of the NE is mainly due to the selfish behavior (in the sense of self-optimization) of the players. An effective method to induce cooperation—while maintaining the noncooperative nature of the interaction—is forcing the players to interact more than once. A typical example of this approach is a repeated game, in which a static game is repeated N times. For instance, assume that the two transmitters of Figure 1(a) interact a number N of times, each time selecting their optimal transmit powers $s_k(n)$, where n is the time index [11]. When introducing the notion of time, each strategy set \mathcal{S}_k becomes a complete plan of actions that depends on the unfolding of the game through time.

Similarly, the utility functions must account for 1) the partial utilities $u_k(s(n)) = t_k(s(n))/s_k(n)$ received at each stage n of the game, with $s(n)$ denoting the profile selected by the players at time n , and 2) how much past utilities should be weighted (i.e., decay) compared to present utilities. A simple example is the exponential decay, where the utility at time n is weighted by the factor δ^n , $0 \leq \delta \leq 1$, and the total utility after N repetitions of the game is $u_k^N(s) = \sum_{n=0}^N \delta^n u_k(s(n))$. By letting $N \rightarrow \infty$, we further define a normalized utility $u_k^\delta(s) = (1 - \delta) \sum_{n=0}^{+\infty} \delta^n u_k(s(n))$ [11]. The parameter δ is the so-called discount factor, and its meaning is borrowed from microeconomics: a payoff received at the present time n is larger by a factor $1/\delta$ than the payoff of the next stage, and smaller by a factor δ than that of the previous one. This means that, if players are patient (in the SP and communications context, delay tolerant), δ is typically close to 1. Conversely, if players are impatient (i.e., delay sensitive), δ is typically close to 0.

Extending the concept of NE to repeated games [1, Ch. 5], we can show that the optimal strategy $s^\delta = \{s^\delta(n)\}_{n=0}^{+\infty}$

is for both players to select $s_k^\delta(n) = \check{s}_k$ if $(s_k^\delta(n-1) = \check{s}_k \text{ and } s_{\setminus k}^\delta(n-1) = \check{s}_{\setminus k})$, and $s_k^\delta(n) = s_k^*$ otherwise, with $s_k^\delta(0) = \check{s}_k$, and \check{s} and s^* being the SO and NE points, respectively, provided that $\delta \geq \underline{\delta}$ (i.e., if they are delay-tolerant enough), where $\underline{\delta}$ is a function of the network parameters [11]. In other words, in the repeated IC game, cooperation is enforced by letting the players interact an indefinite number of times: this is successful due to threatening future punishments for the player(s) who defect.

The effectiveness of this approach is apparent in Figure 4, where the performance of s^δ , represented by the black asterisk, coincides with the SW, under the assumption $\delta \geq \underline{\delta}$. As a drawback, players must have knowledge of \check{s} : this only occurs if each player k knows all channel gains $\{h_{jk}\}_{j,k \in \mathcal{K}}$, which might not be viable for all scenarios (e.g., in a cognitive network). Repeated games are a subclass of dynamic games, which are often used in SP problems to account for time evolution (see [7, Ch. 3] for more details).

INTRODUCING COOPERATION AMONG THE PLAYERS

In the techniques considered so far, we have focused on improving the efficiency of the solution without considering any issues of fairness. In the aforementioned example, the SW is obtained by favoring player 2 to the detriment of player 1's performance, as is apparent in the inset of Figure 4, where the SW is far away from the projection of the NE over the Pareto frontier—obtained by intersecting it with the “fair” line with slope 1 and passing through $(u_1(s^*), u_2(s^*))$. We can balance efficiency and fairness by explicitly introducing cooperation among the players, assuming some explicit exchange of information. The fundamental difference of a cooperative approach is that, while in the games assumed so far, cooperation can only be induced as the result of matching it with self-optimization (i.e., unilateral deviations are not beneficial anyway), now that the players are willing to cooperate, as they know that they can mutually benefit from reaching an agreement. In GT parlance, this is called a *bargaining problem* [2, Ch. 7], whose analytical tools are tightly related to

SP techniques, such as consensus algorithms [12].

Consider again the continuous IC game and assume that the players can collaborate to select a satisfactory profile $\dot{s} \in \mathcal{S}_1 \times \mathcal{S}_2$. In case they fail to reach an agreement, each player k gets $u_k(s^*)$, where s^* is the NE of the noncooperative game studied before. On the contrary, the players now strive to attain the Nash bargaining solution (NBS), i.e., the (unique) PO profile that satisfies

$$\dot{s} = \operatorname{argmax}_{s \in \mathcal{S}} \prod_{k=1}^2 (u_k(s) - u_k(s^*)),$$

where the subset $\mathcal{S} \subseteq \mathcal{S}_1 \times \mathcal{S}_2$ is such that $u_k(s) \geq u_k(s^*)$ for all $k \in \mathcal{K}$ and $s \in \mathcal{S}$. Interestingly, the NBS has close analogies with proportional fair allocation mechanisms, as discussed in [2, Ch. 7]. As is apparent, the NBS tries to increase as much as possible the utilities of the players with respect to the NE in a fair manner.

The graphical interpretation of the NBS is shown in Figure 4: the NBS \dot{s} corresponds to the profile such that $(u_1(\dot{s}), u_2(\dot{s}))$ is the point of tangency between the Pareto frontier and the hyperbola with vertex in $x = u_1(s^*)$ and $y = u_2(s^*)$. Hence, the point of tangency lies by definition in the Pareto improvement region, as illustrated in Figure 4 using a blue dot. In our usual network configuration, $\dot{s}/\sigma^2 = [2.26, 1.52]$, yielding in this case $(\sigma^2/t)u_1(\dot{s}) = 0.288$ and $(\sigma^2/t)u_2(\dot{s}) = 0.434$. The performance of \dot{s} lies in between the SW and the maximum-fairness projection of the NE performance, thus trading off efficiency and fairness. The reason why the NBS is unbalanced toward player 2 lies again in its better channel condition, which makes it stronger in negotiation [2, Ch. 7].

For more than two players, we can also consider a more general cooperative framework, the coalitional GT [13], which provides the theoretical tools to investigate situations in which subsets of players can bind agreements to work together, aiming at improving their joint utility. This approach is particularly useful in many areas of SP, such as spectrum sensing for cognitive systems (see [2, Ch. 13] for further details).

CONCLUDING REMARKS

In this lecture note, we introduced the very basic notions of GT by using a power

control problem for a wireless interference channel as the *leitmotiv*: by further detailing and adding features to this “toy example,” we presented the concepts of players, strategies, utilities, NE, and Pareto and social optimality, among others. Interested readers who want to deepen their knowledge of GT are encouraged to review specific textbooks, such as general ones (e.g., [1]), and those specifically tailored to an SP audience (e.g., [2] and [7]).

REPRODUCIBLE RESEARCH

We have provided supplementary material in [9] that is available for download. The material includes MATLAB code that can reproduce all the simulation results.

AUTHORS

Giacomo Bacci (gbacci@mbigroup.it) is a product manager for satellite and terrestrial communications at MBI srl, Pisa, Italy. After receiving the Ph.D. degree in information engineering from the University of Pisa, Italy, in 2008, he was a postdoctoral research fellow at the University of Pisa in 2008–2014, and also a visiting postdoctoral research associate at Princeton University, New Jersey, in 2012–2014. He is the recipient of the FP7 Marie Curie International Outgoing Fellowship 2011, the Waveform Diversity and Design conference 2007 Best Student Paper Award, the IEEE Wireless Communications and Networking Conference 2013 Best Paper Award, and the 2014 International Union of Radio Science Young Scientist Award.

Luca Sanguinetti (luca.sanguinetti@unipi.it) is an assistant professor at the University of Pisa, Italy. In 2007–2008, he was a postdoctoral associate at Princeton University, New Jersey. From 2013 to 2015, he was also with Centrale Supélec, France. His research topics are in the broad area of wireless communications and signal processing. He is currently serving as an associate editor of *IEEE Transactions on Wireless Communications*. He is the recipient of the FP7 Marie Curie Intra-European Fellowship 2013 and the IEEE Wireless Communications and Networking Conference 2013 and 2014 Best Paper Awards.

Marco Luise (marco.luise@unipi.it) is a full professor of telecommunications at the

University of Pisa, Italy. Formerly an editor of *IEEE Transactions on Communications*, he is a cofounder of the *International Journal of Navigation and Observation*, a division editor of the *Journal of Communication and Networks*, and the coordinator of the European Network of Excellence in Wireless Communications (NEWCOM#). He has chaired a number of scientific conferences, including the International Union of Radio Science Symposium ISSSE'98, European Signal Processing Conference 2006, European Wireless 2010, and IEEE International Conference on Acoustics, Speech, and Signal Processing 2014, and has authored more than 300 publications in the broad area of signal processing, wireless communications, and positioning. He is a Fellow of the IEEE.

REFERENCES

- [1] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.
- [2] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [3] E. Björnson, E. Jorswieck, M. Debbah, and B. Ottersten, “Multi-objective signal processing optimization: The way to balance conflicting metrics in 5G systems,” *IEEE Signal Processing Mag.*, vol. 31, no. 6, pp. 14–23, Nov. 2014.
- [4] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, “Convex optimization, game theory, and variational inequality theory,” *IEEE Signal Processing Mag.*, vol. 27, no. 3, pp. 35–49, May 2010.
- [5] J. B. Predd, S. R. Kulkarni, and H. V. Poor, “Distributed learning in wireless sensor networks,” *IEEE Signal Processing Mag.*, vol. 23, no. 4, pp. 56–69, July 2006.
- [6] G. Bacci and M. Luise, “How to use a strategic game to optimize the performance of CDMA wireless network synchronization,” in *Mechanisms and Games for Dynamic Spectrum Allocation*, T. Alpcan, H. Boche, M. L. Honig, and H. V. Poor, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2013, pp. 518–551.
- [7] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks: Fundamentals and Applications*. Waltham, MA: Academic Press, 2011.
- [8] F. Meshkati, H. V. Poor, and S. C. Schwartz, “Energy-efficient resource allocation in wireless networks,” *IEEE Signal Processing Mag.*, vol. 24, no. 3, pp. 58–68, May 2007.
- [9] G. Bacci, L. Sanguinetti, and M. Luise, MATLAB code. [Online]. Available: <https://github.com/lucasanguinetti/in-game-theory>
- [10] C. Saraydar, N. B. Mandayam, and D. Goodman, “Efficient power control via pricing in wireless data networks,” *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.
- [11] M. Le Treust and S. Lasaulce, “A repeated game formulation of energy-efficient decentralized power control,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2860–2869, Sept. 2010.
- [12] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
- [13] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Başar, “Coalitional game theory for communication networks,” *IEEE Signal Processing Mag.*, vol. 26, no. 5, pp. 77–97, Sept. 2009.

SP

Signal Processing and Automation in Anesthesia

To anyone who has seen or been to an operating room (OR) or an intensive care unit (ICU), it is obvious that monitoring technology plays an increasingly important role in such clinical settings. This proliferation of monitors and devices means that clinicians in the OR or the ICU must absorb, integrate, and interpret a large amount of data to make clinical decisions. This often results in an information overload that may be detrimental to the effectiveness of decision making under critical conditions. In other areas of activity, such as in the aviation industry, advances in technologies such as signal processing, feature extraction, smart alarm systems, human-computer interfaces, and automation have helped the human operator deal with critical events by increasing situation awareness. In the OR, and particularly in the ICU, where the early detection of deterioration of critically ill patients could significantly reduce preventable morbidity, mortality, and cost, the introduction of such methodologies may be essential to increase patient safety. In addition to a brief description of the state-of-the-art approaches, this column reviews some of the efforts my research group at the University of British Columbia (BC) and BC Children's Hospital has performed over the last decade to address those issues, starting with the OR, concentrating on anesthesia.

AN OVERVIEW OF ANESTHESIA

The goals of anesthesia are to allow the surgeon to operate in optimal conditions

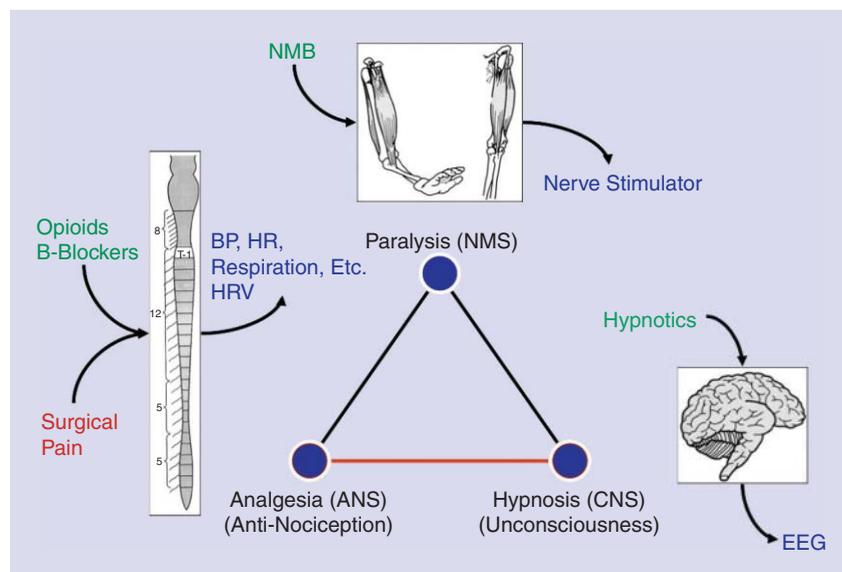
Digital Object Identifier 10.1109/MSP.2015.2405752

Date of publication: 15 June 2015

while occulting the patient from the effects of the surgical procedure and maintaining homeostasis and hemodynamic stability as much as possible. For this, the anesthesiologist administers a number of drugs to the patients (Figure 1): 1) hypnotics that act primarily on the brain to induce unconsciousness (also known as a hypnotic state) in the patient to prevent intra-operative awareness and memorization; 2) analgesics to suppress nociceptive reactions in the presence of painful stimulus caused by the surgical act; and 3) neuromuscular blocking (NMB) agents (mostly curare derivatives) to induce paralysis to suppress reflex muscle activity (occasionally administered, particularly in abdominal procedures).

The role of the anesthesiologist is to carefully dose the amounts to avoid underdosing, which can lead to intraoperative memorization, hypertension, tachycardia, and possible postoperative posttraumatic stress, as well as overdosing, which can lead to prolonged electrocortical silence, hypotension, bradycardia, and, due to the toxicity of the drugs involved, may lead to serious or even fatal intra- or postoperative consequences for the patient. Recent studies have linked too-deep anesthesia to postoperative delirium, long-term cognitive losses, and even increased one-year mortality rates.

There are two broad classes of anesthetic agents: inhaled agents and intravenous agents. Common inhaled anesthetics,



[FIG1] Modern balanced anesthesia uses a combination of 1) a hypnotic drug that acts on the central nervous system (CNS) to induce unconsciousness, 2) analgesics and β -blockers that act on the autonomic nervous system (ANS) for antinociception, and 3) NMB agents that act on the neuromuscular system (NMS) to induce a state of muscular relaxation akin to paralysis. Depth of hypnosis (DOH) can be monitored via electroencephalography (EEG); blood pressure (BP), heart rate (HR), and HR variability (HRV) can be used to assess the autonomic system; and an ulnar nerve stimulator can be used to assess the level of muscular relaxation.

often used in combination with nitrous oxide, are:

- desflurane
- isoflurane
- sevoflurane.

Inhaled anesthetics have a combined hypnotic and analgesic effect but also have a strong hypotensive action. An advantage of inhaled anesthetics is that measuring the difference between inhaled and exhaled concentrations allows an accurate estimation of plasma or brain drug uptake. Modern total intravenous anesthesia (TIVA) usually uses propofol as the hypnotic agent and remifentanyl as the analgesic agent. Propofol is characterized by fast redistribution and metabolism, provides rapid emergence, and has good anti-emetic properties. Remifentanyl is characterized by a very rapid onset and brevity of action, thus minimizing undesirable opioid-induced side effects. That, combined with the high specificity of both agents, makes them ideal for feedback control of anesthesia. Hence, the vast majority of studies of closed-loop control of anesthesia have been performed using intravenous anesthesia. Finally, NMB agents such as rocuronium and atracurium are commonly used as a bolus at the beginning of the intervention to facilitate intubation. NMB boluses are sometimes repeated during the intervention to provide optimal conditions for the surgeon if necessary.

We next divide the anesthesia procedure into three distinct stages: induction, maintenance, and emergence.

INDUCTION

Induction, or the phase during which the patient is being rendered unconscious, although quite short is critical in many ways. As soon as the patient loses consciousness, e.g., 2–3 minutes after a bolus of propofol is administered, they will usually stop breathing and need to be rapidly intubated to allow artificial ventilation. To facilitate insertion of the endotracheal tube, the bolus of propofol is usually preceded by a bolus of opioid such as remifentanyl. Furthermore, as soon as the patient loses consciousness, to blunt any reflex during intubation, NMB is generally administered. Overdosing the patient at induction may

lead to severe hypotension, which will need to be corrected with vasopressors, and may place elderly or fragile patients into too-deep an hypnotic state. This may lead to prolonged periods of electrocortical silence, thought to have harmful long-term effects. Minimizing the amount of overshoot at induction is thus critical.

MAINTENANCE

During the maintenance phase, it is necessary to maintain an adequate DOH and to blunt nociceptive reactions. When using inhaled anesthetics, the measurement of the end-tidal vapor concentration provides the anesthesiologist with a reliable feedback quantity. The situation is more complex with TIVA, as no measurement of arterial concentration of propofol or remifentanyl is available. In the absence of brain monitoring, the anesthesiologist will use hemodynamic parameters such as HR and BP for guidance, or watch for patient movement. The development of TIVA has been made easier through the development of pharmacokinetic (PK) model-driven infusion devices. These devices reach a desired plasma (or effect site) theoretical concentration by using a computer-controlled infusion pump driven by the PK parameters of the drug. The resulting so-called target controlled infusion (TCI) [1] anesthesia is used extensively in most of the developed world except in the United States, where it is not approved by the U.S. Food and Drug Administration.

EMERGENCE

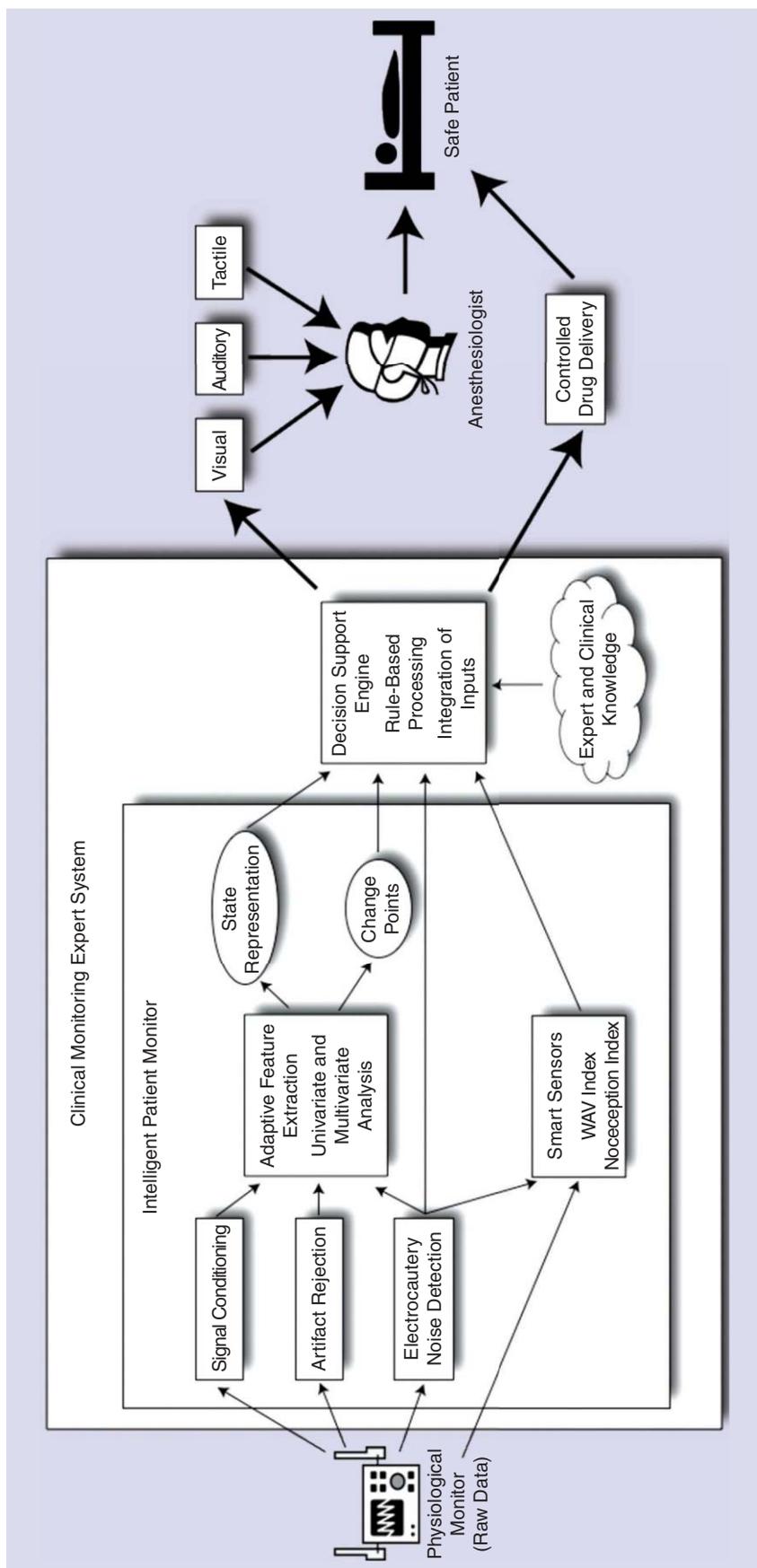
The emergence from anesthesia is simply achieved by turning off delivery of the hypnotic and analgesic agents used during the surgery. This is usually done during skin closure so that the patient wakes up more quickly at the end of the surgery. An additional bolus of a long acting opioid may be given for postoperative pain management. Extubation takes place as soon as the patient shows clinical signs of wakefulness.

TOOLS FOR REAL-TIME DECISION SUPPORT IN CRITICAL CARE

The task of an anesthesiologist is often compared with that of an airplane pilot in that both function as part of high-

performance teams with high stakes in complex and potentially unpredictable environments, with serious mishaps potentially resulting in loss of life. Both must display a high degree of skill and accuracy and use sophisticated technology. As a result, both use checklists, safety protocols, in-depth error analysis, and have to deal with a constant flow of data from numerous monitors. Finally, both have to be constantly prepared to deal with potential disaster. Whereas technology such as automation has significantly increased safety in aviation, it has yet to penetrate the anesthesia world. A number of research groups have been developing sophisticated decision-support systems for anesthesia and critical care. Figure 2 shows the Intelligent Anesthesia Navigator (IAN) proposed by our research team. IAN was developed to improve the system function of the “bedside cockpit” of the OR by applying advanced signal processing solutions that intelligently analyze and interpret the large amounts of physiological sensor data and turn it into clinically useful information. These solutions were extended from algorithms for feature extraction to integration of information with expert knowledge, to the design of novel human interaction interfaces and, finally, to enhanced automation. Indeed we see the successful introduction of intelligent monitoring and automated control as the only practical solution to the problem of information overload produced by the increasing number of sensors developed for physiological monitoring.

A major component of such a system is a smart alarm, or more exactly a smart alert system that cuts down on the number of false positives, a serious nuisance in an OR. However, the first step is to ensure that artifacts are detected and, if possible, rejected. In an OR, the electrical surgical unit (ESU) passes a high-frequency current from an active electrode to a grounding plate via the patient's body, resulting in a high-frequency noise swamping both EEG and ECG measurements. Electrocautery noise is thus a significant obstacle to the real-time implementation of clinical monitoring expert systems. A solution described in [2] consists of projecting the



[FIG2] The IAN consists of a number of algorithms and devices to help clinicians provide safe anesthesia and sedation. In a first step, data collected from physiological monitors are checked for artifact before being analyzed and interpreted to provide clinically relevant information to the clinician. Once interpreted, the results can be sent to a clinical expert system to provide alerts and alarms to the clinician using a number of modalities. Furthermore, some data can be used as feedback variables by an automated drug delivery system providing the exact amount of drugs required to maintain the patient in the anesthetic state chosen by the clinician.

ECG into a wavelet basis in which the QRS complex is sparse whereas the additive noise is not. Then, by assessing the l_0 norm of groups of wavelets coefficients it allows discrimination of the ECG signal from the electrocautery noise. Whenever possible, redundant measurements, when available, should be used to maintain physiological monitoring. For instance, HR can be obtained not only from the ECG but also from the photoplethysmograph obtained from pulse oximetry. As the latter is less affected by the ESU, a simple fusion algorithm based on a hybrid median filter has been shown to successfully reject ESU noise and produce a robust estimate of HR [3]. Once data quality is ensured, another important task of such a clinical decision support system is to transform those data into information that is clinically relevant. Although all monitors come with threshold-based alarms, those have been not only useless but actually harmful as false alarm fatigue leads the clinicians to ignore them, thus resulting in monitors beeping while nobody pays attention to them. Building smarter alarms, or rather alerts that are clinically relevant is thus of great interest to clinicians. This has led to the development of number of trend detection and change point detection algorithms for use in the OR [4] and the ICU [5]. Such a system, coupled with an expert system for real-time clinical decision support, has recently been shown to vastly improve the detection of critical events during anesthesia in a human patient simulator [6]. Finally, to cut down on auditory pollution we developed a novel vibrotactile belt to transmit those alerts to the clinician in a way that is nonintrusive for the personnel in the OR and

demonstrated its usability and efficacy in a real clinical environment [7].

SENSING FOR CLOSED-LOOP CONTROL OF ANESTHESIA

DEPTH OF HYPNOSIS

The effects of anesthetic drugs on the EEG have been known since the early 1940s when neurophysiologists observed that the EEG of anesthetized patients contained slower waves with higher amplitudes. However, raw EEGs are difficult to interpret in real time and thus a number of techniques have been used to extract univariate features from the EEG to quantify the hypnotic component of anesthesia. Two such features of historical interest are the median frequency (MEF) and the spectral edge frequency (SEF), i.e., the frequency up to which 95% of the EEG power is present. However, this is not until the advent of the bispectral index (BIS) monitor that EEG has become more common place. The BIS monitor is based on the observation that with increasing anesthetic depth, EEG frequencies tend to synchronize. This led to the use of the bispectrum to characterize phase coupling of different frequencies. The BIS monitor combines a number of bispectra, bicoherence indices and power spectral values to derive a [0–100] index known as *DOH*. An index of 100 represents the awake state while it decreases with increasing concentration of anesthetics. General anesthesia is obtained for an index between 60 and 40. Lower values represent deep hypnotic states and usually are not desirable. Introduced in the mid-1990s, the BIS monitor largely dominates the market for DOH monitors.

The second most common DOH monitor is the M-Entropy monitor (GE Healthcare, Helsinki, Finland), introduced in 2003, which provides two indices, the state entropy (SE), a measure of the irregularity of frontal EEG activity within the frequency range of 0.8–32 Hz; and the response entropy (RE), a measure of the irregularity of frontal EEG activity within the frequency range of 0.8–47 Hz. While SE is a surrogate of the BIS, the difference between RE and SE is thought of as an

indication of nociception because it may contain some facial electromyography.

Although it provides anesthesiologists with a reliable index of hypnosis, the BIS introduces a large and variable delay, is inherently nonlinear, tending to evolve in stepwise manners during transient phases of anesthesia, and is essentially a black box, hard to characterize for control design purposes. On the other hand, the M-Entropy, which responds much faster, is a simpler algorithm but tends to provide a very noisy index unless heavily filtered and thus delayed.

The more recent NeuroSense monitor (NeuroWave Systems Inc., Cleveland Heights, Ohio) that addresses those concerns was developed specifically for use in closed-loop control of anesthesia. It derives a bilateral index based on wavelet decomposition of a frontal EEG, with emphasis on the γ -band (30–70 Hz) activity [8]. It has been shown to relate well with the BIS in steady state, but possesses much faster, delay-free, and constant dynamics over its entire range [9], characteristics that make it ideal for use in closed-loop control of the DOH.

NOCICEPTION

Nociception describes the stress response of the ANS to a noxious stimulus. This stress response must be controlled by the anesthesiologist by administering analgesics drugs. Sensing for nociception or analgesia has proved to be much more difficult to develop than for DOH. In the absence of specific monitors, anesthesiologists assess the patient's nociceptive reactions by keeping a close eye on HR and BP, both of which tend to increase sharply in case of a sympathetic response to an improperly blunted noxious stimulus. To date, the only commercially available system is GE's Surgical Stress Index (SSI) (GE Healthcare, Helsinki, Finland) computed from finger photoplethysmographic waveform amplitudes and pulse-to-pulse intervals. The ANI (MetroDoloris SAS, Loos, France) algorithm analyzes the tachogram with wavelets, and tracks the time-varying power in the HF band. The ANI index has been shown to respond to the administration of anesthetic drugs and

to nociceptive stimuli. A related technique is based on wavelet-based cardio-respiratory coherence, resulting in a normalized index that has been shown to respond to both nociceptive and antinociceptive events, [10]. Both techniques are currently undergoing clinical trials for validation.

An interesting and indirect method for assessing nociceptive reactions is based on the observation that a sudden electrocortical activation, e.g., demonstrated by a large and rapid increase in the DOH, is a reflection of an inadequate analgesic state or level of antinociception. This principle has been used in the most clinically tested closed-loop controller for anesthesia described in the next section.

NEUROMUSCULAR BLOCKADE

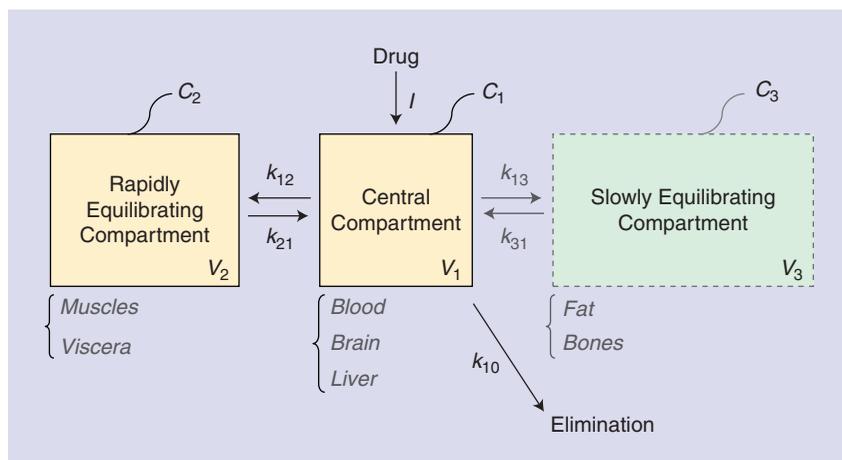
The monitoring of neuromuscular relaxation is performed by measuring, mostly by acceleromyography, evoked muscle response following supramaximal stimulation of the ulnar nerve. The standard stimulation consists of a train of four short pulses, and the level of paralysis is characterized by the ratio of the amplitude of the fourth response normalized by that of the first one, the so-called T4/T1 ratio. Because of the virtual on/off nature of the measurement, it is not an ideal one for use in closed-loop control. Fortunately, fine control is not a clinical requirement. Nevertheless, new techniques are being developed based on phonomyography, which consists of recording low-frequency sounds created during muscle contraction. The advantage of phonomyography is that it can be used at sites that are more clinically relevant than the adductor pollicis.

CLOSED-LOOP CONTROL OF INTRAVENOUS ANESTHESIA

MODELING FOR ANESTHESIA

Modeling of the distribution and effect of anesthetic drugs has traditionally been done using PK models for the former, and pharmacodynamic (PD) models for the latter. Typically, PK models are based on maxillary compartmental models, while PD models consist of a simple

life SCIENCES continued



[FIG3] A three-compartment PK model typically used for the hypnotic drug propofol.

compartment followed by a sigmoidal nonlinearity. For propofol, a three-compartment model (as shown in Figure 3) is used, yielding the transfer function between the infusion rate I_p and the plasma concentration C_p :

$$C_p(s) = \frac{1}{V_1} \frac{(s_1 + z_1)(s + z_2)}{(s + p_1)(s + p_2)(s + p_3)} I_p(s).$$

The PD model is usually described by a transfer function between the plasma concentration C_p and the effect site concentration C_e

$$C_e(s) = \frac{k_{e0}}{s + k_{e0}} C_p(s)$$

followed by a Hill equation relating C_e to the effect:

$$E(C_e) = \frac{C_e^\gamma}{EC_{50}^\gamma + C_e^\gamma},$$

where EC_{50} denotes the effect site concentration corresponding to a 50% clinical effect and γ is the cooperativity coefficient. For remifentanyl, most PK models involve only two compartments, resulting in a simpler transfer function.

Propofol and remifentanyl are known to interact with each other in a synergistic fashion in their hypnotic/analgesic effect [11]. This observation constitutes the basic assumption on which the balanced anesthesia concept is based. Studies of the PD interactions between these two drugs

have been limited to steady characterization through the use of response surfaces, which can then be parameterized, e.g., as

$$E(v_p, v_r) = \frac{(v_p + v + r + \alpha v_p v_r)^\gamma}{(v_p + v + r + \alpha v_p v_r)^\gamma + 1}$$

where v_p and v_r are, respectively, the effect site concentrations of propofol and remifentanyl normalized by their EC_{50} , and $\alpha > 0$ characterizes the synergy between the two drugs. Note that the interaction is equivalent to the use of a new fictitious drug $v = v_p + v_r + \alpha v_p v_r$.

While pharmacokineticians and dynamicists strive to improve the accuracy of PKPD models by introducing a number of covariates in an attempt to reduce the uncertainty of TCI systems, a number of studies have shown that to develop a clinically satisfactory closed-loop control systems, simpler models such as first-order plus delay have the same level of predictive power [12].

ANESTHETIC DRUG CONCENTRATION ESTIMATION

Whereas in inhaled anaesthetics, the end tidal anesthetic concentration can be used as a proxy of the plasma concentration, there is as yet no such possibility with intravenous anesthetics. Thus, anesthesiologists come to rely on open-loop prediction [see Figure 4(a)], in which the drug concentration is estimated by solving the patient PK described in terms of a multi-compartmental model. In the absence of

feedback, this approach inevitably suffers from interpatient variability. Any discrepancy between the real PK of an individual patient and its model counterpart results in a mismatch between the true versus estimated drug concentrations. However, a closed-loop approach, exploiting the clinical-effect measurement, i.e., the depth of anesthesia index wavelet-based anesthetic value CNS [see Figure 4(b)], has been shown to improve the accuracy of the estimation of intravenously administered propofol concentrations at the plasma and effect sites [13]. To strike the right compromise between robustness in face of interpatient variability and rejection of unknown surgical disturbances, the estimator was designed using μ -synthesis theory.

CLOSED-LOOP CONTROL OF DEPTH OF HYPNOSIS IN INTRAVENOUS ANESTHESIA

1) *Control Paradigm:* After ensuring a fast and safe induction, the anesthesiologist needs to maintain the patient in an adequate state of hypnosis, analgesia, and paralysis according to the requirements of the surgical procedure. The anesthetic and opioid titration needs to be constantly adjusted to avoid both under and overdosing of the patient. The idea of an automated system that would regulate drug dosing to maintain the adequacy of the anesthetic regimen is thus natural. Closed-loop anesthesia would not replace the anesthesiologist, but would allow them to concentrate on higher level tasks.

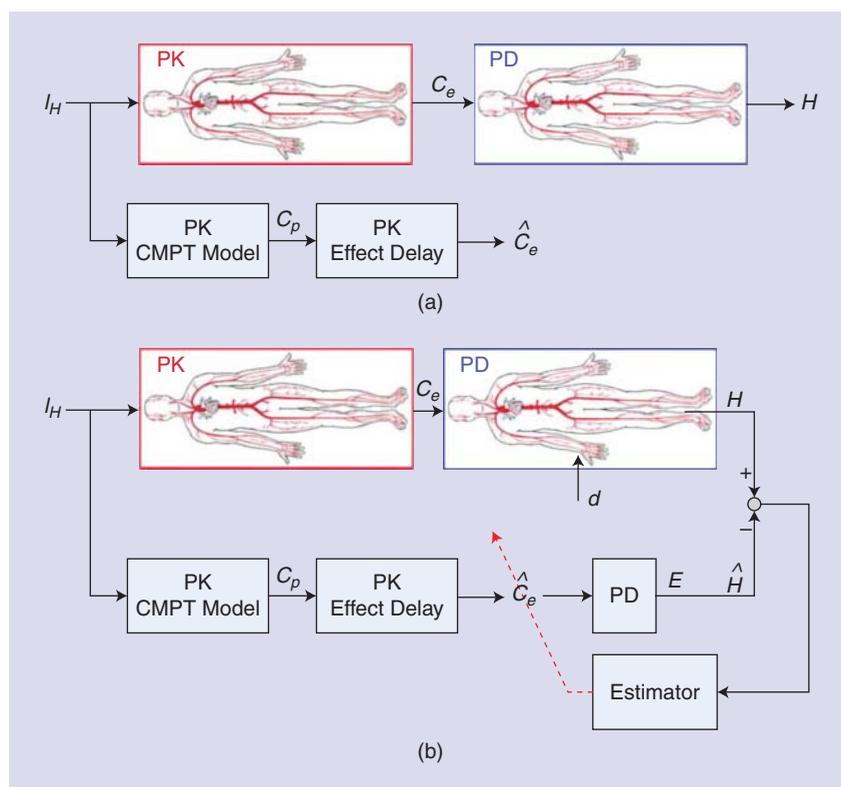
A closed-loop controller for anesthesia should induce the patient rapidly but with minimal overshoot and then maintain the patient in an adequate state of anesthesia and analgesia at least as well as an expert anesthesiologist. Translating this in control specifications is difficult, but for the DOH index, it could be translated into a rise time at induction of 3–4 minutes, with overshoot less than 10–15% and a damping ratio of at least 0.7. During maintenance, the DOH index should stay within ten points of the target about 85–90% of the time. Disturbance rejection should be such that in case of arousal

(which in control engineering terms can be thought of as an output disturbance), the patient response is rapidly suppressed, say within 2 minutes and without inducing oscillations. The clinical outcome should be improved hemodynamic stability, faster emergence, and possibly reduced drug consumption. The main challenge is the inherent variability, both interpatient and intraoperative, thus robust stability and robust performance are paramount to ensure patient safety when designing such a closed-loop control system.

2) *Research Systems*: The first efforts to automate anesthesia go back to the work of Mayo and Bickford in the early 1950s with their attempts to develop EEG-based automatic delivery of volatile agents; see [14] for a review of the progress from 1949 to 1980.

The advent of the BIS monitor in the mid-1990s dramatically changed the situation resulting in a significant increase in the number of studies, both simulated and clinical, on closed-loop control of depth of anesthesia.

The pioneering systems for BIS-guided closed-loop control of propofol infusion were heuristic systems developed by anesthesiologists without the involvement of control experts. Absalom [15] used a PID controller tuned in a very ad-hoc manner to adjust the propofol effect site concentration setpoint in a target site infusion system. This was clinically tested on 20 patients. Not surprisingly, the performance varied significantly from patient to patient, the system displaying instability for some of them. All patients were induced under TCI mode, with an effect site concentration chosen by the clinician who then switched to closed-loop control for the maintenance phase. A system with a similar structure is described in [16]. That controller is actually not a PID controller, but a rule-based system that is somewhat similar to a PD controller. After significant tuning, this system was tested against manual control in a randomized controlled trial involving 164 patients (83 in closed loop). That system was shown to outperform manual control in terms of BIS variability, and resulted in similar hemodynamic stability. In [17], the authors present what they term a *dual*



[FIG4] The open-loop prediction versus closed-loop estimation of anesthetic drug concentrations. (a) Open-loop prediction. (b) Closed-loop estimation. (Figure used courtesy of [13].)

loop that manipulates both the propofol and remifentanyl infusion rates based on the BIS alone. The basic idea is that sudden increases in the BIS index are due to a nociceptive reaction and reflect an inadequate analgesic state. The controller that manipulates the remifentanyl is a combination of a proportional action and a number of heuristic rules. Randomized clinical trials involving 167 patients on a wide variety of procedures showed that the system provides better control of the BIS than manual control, similar hemodynamic stability accompanied with increased remifentanyl consumption. This system, like its predecessor induces the patient in TCI mode, with manually set targets for both propofol and remifentanyl effect site concentrations.

Because these pioneering systems were designed heuristically, their theoretical properties are difficult, if not impossible to assess. A rigorous approach to robust PID tuning for anesthesia is described in [18] where a PID controller is robustly tuned for a population of

44 adults. Results of a feasibility study in adults showed that this simple controller provided adequate anesthesia. This led to the development of a similar system for pediatric use, which resulted in the DOH within ten points of the target 89% of the time in a clinical study on 108 patients reported in [19] and [20]. Those results indicate that a robust PID controller manages relatively well with the significant interpatient uncertainty in a pediatric population.

THE FUTURE

Despite the number of studies that have over the years demonstrated the feasibility and safety of real-time systems for monitoring and closed-loop control of anesthesia, the regulatory hurdles abound before it can be approved for routine clinical use. First and foremost, the clinical benefits of those systems have to be clearly demonstrated, and this will require extensive multicentric clinical studies involving tens of thousands of patients. The development of fallback and safety systems, as well as

the proof that they preserve patient safety, will be paramount to the broad clinical application of anesthesia automation systems. Many of the systems that have been used clinically so far have been developed without control engineers and consist of very heuristic controllers. It is thus important for signal processing, control, and software engineering experts to be at the heart of the design of those systems to ensure they are based on sound control theory, as regulatory approval will likely require guarantees in terms of stability, robustness and performance.

It is not entirely clear yet what level of control complexity is required to achieve clinically acceptable performance. Novel monitors of nociception are in the works and will hopefully allow true multivariable control of hypnosis and analgesia. It is also important to realize that the systems discussed here only consider a subset of what an anesthesiologist needs to do to ensure patient safety. Other systems, e.g., for hemodynamics and fluid management, are being considered and should undergo clinical trials in the coming years.

ACKNOWLEDGMENTS

I would like to thank the Natural Sciences and Engineering Research Council of Canada and Canadian Institutes of Health Research for supporting this work over the years through a number of grants as well as my various collaborators: clinicians, postdoctoral fellows, clinical fellows, graduate students, research engineers, and research associates who have inspired and contributed to this work over the last

ten years. This work was performed while I was an invited professor at the Université d'Avignon et des Pays de Vaucluse and then a 2015 STIAS Fellow.

AUTHOR

Guy A. Dumont (guyd@ece.ubc.ca) is a professor with the Department of Electrical and Computer Engineering at the University of British Columbia, Canada. He currently is a fellow of the Stellenbosch Institute of Advanced Study at the Wallenberg Research Centre at Stellenbosch University, South Africa.

REFERENCES

- [1] A. R. Absalom and M. M. R. F. Struys, *An Overview of TCI & TIVA*, 2nd ed. New York: Academic Press, 2007.
- [2] C. Brouse, G. A. Dumont, F. Herrmann, and J. M. Ansermino, "A wavelet approach to detecting electrocautery noise in the ECG," *IEEE Eng. Med. Biol. Mag.*, vol. 25, no. 4, pp. 76–82, July 2006.
- [3] P. Yang, G. A. Dumont, and J. M. Ansermino, "Sensor fusion using a hybrid median filter for artifact removal in intraoperative heart rate monitoring," *J. Clin. Monitor. Comput.*, vol. 23, no. 2, p. 7583, 2009.
- [4] P. Yang, G. A. Dumont, and J. M. Ansermino, "Adaptive change detection in heart rate monitoring in anesthetized children," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2211–2219, 2006.
- [5] S. Charbonnier and S. Gentil, "A trend-based alarm system to improve patient monitoring in intensive care units," *Control Eng. Pract.*, vol. 15, no. 9, pp. 1039–1050, 2007.
- [6] M. Gorges, P. Winton, V. Koval, J. Lim, J. Stinson, P. Choi, S. Schwarz, G. A. Dumont, and J. M. Ansermino, "Evaluation of an expert system for detecting critical events during anesthesia in a human patient simulator: A prospective randomized controlled study," *Anesthes. Analges.*, vol. 117, no. 2, pp. 380–391, 2013.
- [7] M. Dosani, K. Hunc, G. A. Dumont, D. Dunsmuir, P. Barralon, S. Schwarz, J. Lim, and J. M. Ansermino, "A vibro-tactile display for clinical monitoring: real-time evaluation," *Anesth. Analges.*, vol. 115, no. 3, pp. 588–594, 2012.
- [8] T. Zikov, S. Bibian, G. A. Dumont, M. Huzmezan, and C. Ries, "Quantifying cortical activity during general anesthesia using wavelet analysis," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 4, pp. 617–632, Apr. 2006.

[9] S. Bibian, G. A. Dumont, and T. Zikov, "Dynamic behavior of BIS, m-entropy and NeuroSENSE brain function monitors," *J. Clin. Monitor. Comput.*, vol. 25, no. 1, pp. 81–87, Nov. 2011.

[10] C. J. Brouse, W. Karlen, G. A. Dumont, D. Myers, E. Cooke, J. Stinson, J. Lim, and J. M. Ansermino, "Monitoring nociception during general anesthesia with cardiorespiratory coherence," *J. Clin. Monitor. Comput.*, vol. 27, no. 5, pp. 551–560, 2013.

[11] S. E. Kern, G. Xie, J. L. White, and T. D. Egan, "Opioid-hypnotic synergy: A response surface analysis of propofol-remifentanyl pharmacodynamic interaction in volunteers," *Anesthesiology*, vol. 100, no. 6, pp. 1374–1381, 2004.

[12] J.-O. Hahn, G. A. Dumont, and J. M. Ansermino, "A direct dynamic dose-response model of propofol for individualized anesthesia care," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 571–578, 2012.

[13] J.-O. Hahn, G. A. Dumont, and J. M. Ansermino, "Closed-loop control of anesthesia using clinical effect feedback," *IEEE Trans. Biomed. Eng. Lett.*, vol. 58, no. 1, pp. 3–6, 2011.

[14] R. Chilcoat, "A review of the control of depth of anaesthesia," *Trans. Inst. M.C.*, vol. 2, no. 1, pp. 38–45, 1980.

[15] A. R. Absalom, N. Sutcliffe, and G. N. Kenny, "Closed-loop control of anesthesia using bispectral index," *Anesthesiology*, vol. 96, no. 1, pp. 67–73, Jan. 2002.

[16] N. Liu, T. Chazot, A. Genty, A. Landais, A. Restoux, K. McGee, P. A. Lalo, B. Trillat, L. Barvais, and M. Fischler, "Titration of propofol for anesthetic induction and maintenance guided by the bispectral index: closed-loop versus manual control," *Anesthesiology*, vol. 104, no. 4, pp. 686–695, Apr. 2006.

[17] N. Liu, T. Chazot, S. Hamada, A. Landais, N. Boichut, C. Dussaussoy, B. Trillat, L. Beydon, E. Samain, D. I. Sessler, and M. Fischler, "Closed-loop coadministration of propofol and remifentanyl guided by bispectral index: A randomized multicenter study," *Anesthes. Analges.*, vol. 112, no. 3, pp. 546–557, Mar. 2011.

[18] G. A. Dumont, A. Martínez, and J. M. Ansermino, "Robust control of depth of anesthesia," *Int. J. Adapt. Control Signal Process.*, vol. 23, no. 5, pp. 435–454, May 2009.

[19] K. van Heusden, K. G. A. Dumont, K. Soltesz, C. Petersen, A. Umedaly, N. West, and J. M. Ansermino, "Clinical evaluation of robust PID control of propofol anesthesia in children," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 2, pp. 491–501, 2014.

[20] N. West, G. A. Dumont, K. van Heusden, C. Petersen, S. Khosravi, K. Soltesz, A. Umedaly, E. Reimer, and J. M. Ansermino, "Robust closed-loop control of induction and maintenance of propofol anesthesia in children," *Pediatric Anesthes.*, vol. 23, no. 8, pp. 712–719, 2013.

SP

IEEE Journal of Selected Topics in Signal Processing (JSTSP): Recent Special Issues



- ❖ June 2015 – Big Data
- ❖ April 2015 – Interactive Media Processing for Immersive Communications
- ❖ March 2015 – Signal Processing for Situational Awareness From Networked Sensors and Social Media
- ❖ February 2015 – Visual Signal Processing for Wireless Networks

Visit IEEE Xplore for more information about these special issues



Chun-Su Park

Fast, Accurate, and Guaranteed Stable Sliding Discrete Fourier Transform

Discrete orthogonal transforms such as the discrete Fourier transform (DFT), discrete Hartley transform (DHT), and Walsh–Hadamard transform (WHT) play important roles in the fields of digital signal processing, filtering, and communications. In recent years, there has been a growing interest in the sliding transform process where the transform window is shifted one sample at a time and the transform process is repeated.

Jacobsen and Lyons [1] introduced the sliding DFT (SDFT) algorithm, which computes the DFT bins using a recursive scheme. The SDFT can reduce the computational load of the DFT drastically; however, it suffers from potential instability. Several algorithms were proposed to guarantee the stability of the SDFT [2]–[4]. Basically, these algorithms guarantee the stability at the cost of computational complexity or computational accuracy.

This article introduces a new guaranteed stable SDFT (gSDFT) algorithm of which the computational requirement is the lowest among the existing stable SDFT algorithms. Moreover, the output of the proposed gSDFT is mathematically equivalent to that of the DFT regardless of the window size and the time index. This work is motivated by our previous study on the hopping DFT (HDFT) [5]. Let us start with a brief description of the existing stable SDFT algorithms.

EXISTING STABLE SDFT ALGORITHMS

In the SDFT scenario, the transform is computed on a fixed-length window of the signal. Consider a complex input signal

$x(n)$, $n = 0, 1, 2, \dots$, which will be divided into overlapping windows of size M . Let k be the frequency-domain index in the range $0 \leq k < M$. Then, at time index n , the k th bin of an M -point DFT is computed as

$$X_n(k) = \sum_{m=0}^{M-1} x(\hat{n} + m) W_M^{-km}, \quad (1)$$

where $\hat{n} = n - M + 1$ and $W_M = e^{j2\pi/M}$. Accordingly, by the circular shift property, the SDFT is formulated as [1]

$$\begin{aligned} X_n(k) &= \sum_{m=0}^{M-1} x(\hat{n} + m) W_M^{-km} \\ &= \sum_{m=0}^{M-1} x(\hat{n} + m - 1) W_M^{-k(m-1)} \\ &\quad + x(\hat{n} + M - 1) W_M^{-k(M-1)} \\ &\quad - x(\hat{n} - 1) W_M^k \\ &= W_M^k \sum_{m=0}^{M-1} x(\hat{n} + m - 1) W_M^{-km} \\ &\quad + x(\hat{n} + M - 1) W_M^k \\ &\quad - x(\hat{n} - 1) W_M^k \\ &= W_M^k (X_{n-1}(k) + x(n) - x(n-M)), \end{aligned} \quad (2)$$

where the periodicity property of the complex twiddle factor is exploited ($W_M^{k+M} = W_M^k$).

The SDFT has a marginally stable transfer function because its pole lies on the unit circle in the z -domain. In practice, the complex twiddle factor in (2) is represented by a floating-point format with finite precision. The numerical rounding of the complex twiddle factor might move the pole outside the unit circle and result in an unstable system. To address this issue, the stable SDFT algorithm that uses a simple recursive updating scheme was proposed in [2]. This algorithm is realized using a periodically time-varying system designed such that the numerical errors introduced by

finite precision arithmetic exponentially decay to zero over time.

Another stable SDFT, called rSDFT, was proposed in [3]. This algorithm forces the pole to be at a radius of r inside the unit circle by utilizing the damping factor r , thereby guaranteeing stability. The DFT approximation using the damping factor r is expressed as

$$\tilde{X}_n(k) = \sum_{m=0}^{M-1} x(\hat{n} + m) r^{M-m} W_M^{-km}, \quad (3)$$

where $0 \ll r < 1$ and the tilde symbol indicates an approximation of the DFT. Analogously, the SDFT recurrence in (2) is rewritten as

$$\begin{aligned} \tilde{X}_n(k) &= r W_M^k (\tilde{X}_{n-1}(k) \\ &\quad + x(n) - x(n-M) r^M). \end{aligned} \quad (4)$$

Note that, the output bin values of the rSDFT are different from those of the DFT in (1). Moreover, the errors accumulate in the resulting outputs.

The modulated SDFT (mSDFT) was recently introduced in [4]. The mSDFT first generates a modulated sequence by multiplying the input signal by the modulation sequence W_M^{-km} . Then, using the modulated sequence, the mSDFT formulates the recurrence of the DFT bins with $k = 0$ as

$$\begin{aligned} X_n(0) &= X_{n-1}(0) + W_M^{-km} x(n) \\ &\quad - W_M^{-k(m-M)} x(n-M) \\ &= X_{n-1}(0) + W_M^{-km} (x(n) \\ &\quad - x(n-M)), \end{aligned} \quad (5)$$

where $W_M^{-k(m-M)} = W_M^{-km}$. Finally, the desired $X_n(k)$ can be obtained by

$$X_n(k) = W_M^{k(m+1)} X_n(0). \quad (6)$$

By excluding the complex twiddle factor from the feedback of the resonator, the

sp **TIPS&TRICKS** continued

mSDFT has the pole located exactly on the unit circle and is unconditionally stable. However, it was reported in [4] that the computational requirement of the mSDFT is more than double that of the SDFT.

GUARANTEED STABLE SDFT

As mentioned previously, a stable SDFT filter can be designed by excluding the imprecise twiddle factor from the feedback loop. We first investigate the special relationship between the DFT outputs, which can be expressed without using the twiddle factor.

Basically, the SDFT computes the DFT bins by exploiting the recurrence relationship between successive DFT bins. We can derive the general formula between the DFT bins with L -hop distance by extending the relationship between the successive bins. Let $d(n) = x(n) - x(n - M)$ and, then, the resultant formula is obtained by substituting $X_n(k)$ into $X_{n-1}(k)$ L times in (2):

$$\begin{aligned} X_n(k) &= W_M^k(X_{n-1}(k) + d(n)) \\ &= W_M^{2k}(X_{n-2}(k) + d(n-1)) \\ &\quad + W_M^{-k}d(n) \\ &\quad \vdots \\ &= W_M^{Lk}(X_{n-L}(k) + d(n-L+1)) \\ &\quad + W_M^k d(n-L+2) + \dots \\ &\quad + W_M^{(L-1)k}d(n). \end{aligned} \quad (7)$$

Let us define $D_n^L(k)$ as the k th bin of the L -point updating vector transform (UVT), which is represented by

$$D_n^L(k) = \sum_{m=0}^{L-1} d(n-m)W_M^{(m-L+1)k}, \quad (8)$$

where $0 \leq k < M$. Then, using this notation, (7) is simplified as

$$X_n(k) = W_M^{Lk}(X_{n-L}(k) + D_n^L(k)). \quad (9)$$

This leads to the result that the DFT outputs at time index n can be directly computed from those at time index $(n - L)$ by exploiting $D_n^L(k)$. In the next section, we describe how to efficiently implement the UVT in (8).

Let us focus on the twiddle factor W_M^{Lk} , which is multiplied to the delayed output in the recursive computation (9). We can observe that the value of the twiddle factor is varied depending on the time hop L .

We consider the case $L = M/4$. The other cases, $L = M/2$ and $L = 3M/4$, can be treated in a similar manner. When L is equal to $M/4$, we have

$$W_M^{Mk/4} = (\cos(\pi/2) + j \sin(\pi/2))^k = j^k. \quad (10)$$

Accordingly, (9) can be further simplified as

$$X_n(k) = j^k(X_{n-M/4}(k) + D_n^{M/4}(k)). \quad (11)$$

Then, by the periodicity property of j , we obtain the following:

$$\begin{aligned} &\begin{bmatrix} X_n(4i+0) \\ X_n(4i+1) \\ X_n(4i+2) \\ X_n(4i+3) \end{bmatrix} \\ &= \begin{bmatrix} X_{n-M/4}(4i+0) + D_n^{M/4}(4i+0) \\ j(X_{n-M/4}(4i+1) + D_n^{M/4}(4i+1)) \\ - (X_{n-M/4}(4i+2) + D_n^{M/4}(4i+2)) \\ -j(X_{n-M/4}(4i+3) + D_n^{M/4}(4i+3)) \end{bmatrix}, \end{aligned} \quad (12)$$

where $i = 0, 1, \dots, M/4 - 1$. This forms the basis of an efficient scheme for computing the DFT outputs of the shifted window. Given the DFT outputs of the previous window at time index $(n - M/4)$, the outputs at time index n can be directly computed without performing the multiplication by the twiddle factor. For example, when $k = 4i$, the following relationship holds between $X_n(4i)$ and $X_{n-M/4}(4i)$:

$$\begin{aligned} &\text{Re}(X_n(4i)) + j \text{Im}(X_n(4i)) \\ &= \text{Re}(X_{n-M/4}(4i)) + \text{Re}(D_n^{M/4}(4i)) \\ &\quad + j[\text{Im}(X_{n-M/4}(4i)) \\ &\quad + \text{Im}(D_n^{M/4}(4i))], \end{aligned} \quad (13)$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the real and imaginary parts of a complex number, respectively. Then, each part of $X_n(4i)$ is obtained as

$$\begin{cases} \text{Re}(X_n(4i)) = \text{Re}(X_{n-M/4}(4i)) \\ \quad + \text{Re}(D_n^{M/4}(4i)), \\ \text{Im}(X_n(4i)) = \text{Im}(X_{n-M/4}(4i)) \\ \quad + \text{Im}(D_n^{M/4}(4i)). \end{cases} \quad (14)$$

Similarly, the efficient computation for the other cases, $k = 4i + 1, 4i + 2, 4i + 3$, can be derived. These results demonstrate that only two real additions are required for computing $X_n(k)$ using $X_{n-M/4}(k)$ and $D_n^{M/4}(k)$. Further, since the multiplication by the imprecise twiddle factor is excluded

from the recursive calculation, the numerical errors do not accumulate. Therefore, the proposed gSDFT algorithm is stable.

SLIDING UPDATING VECTOR TRANSFORM

The fast implementation of the UVT in (8) is a crucial part because it mainly determines the computational cost of the proposed gSDFT. Since the computation of the UVT is similar to that of the DFT, the traditional fast Fourier transform (FFT) algorithms can be used for its fast implementation. Here we use the radix-2 decimation-in-time (DIT) algorithm, which divides a UVT of size L into two interleaved UVTs of size $L/2$ at each recursive stage. Based on the DIT approach, we express $D_n^L(k)$ using decimated sequences as follows:

$$\begin{aligned} D_n^L(k) &= \sum_{m=0}^{L-1} d(n-m)W_M^{(m-L+1)k} \\ &= \sum_{p=0}^{L/2-1} d(n-2p)W_M^{(2p-L+1)k} \\ &\quad + \sum_{p=0}^{L/2-1} d(n-2p-1)W_M^{(2p-L+2)k} \\ &= W_M^{-k}\check{D}_n^{L/2}(k) + \check{D}_n^{L/2}(k), \end{aligned} \quad (15)$$

where

$$\begin{aligned} \check{D}_n^{L/2}(k) &= \sum_{p=0}^{L/2-1} d(n-2p)W_{M/2}^{(p-L/2+1)k} \\ \text{and} \\ \check{D}_n^{L/2}(k) &= \sum_{p=0}^{L/2-1} d(n-2p-1)W_{M/2}^{(p-L/2+1)k}. \end{aligned} \quad (16)$$

Equation (15) implies that we can obtain $D_n^L(k)$ using two UVT bins of the decimated sequences, $\{d(n), d(n-2), \dots, d(n-L+2)\}$ and $\{d(n-1), d(n-3), \dots, d(n-L+1)\}$. The decimation process is repeated until the resulting sequences are reduced to one-point sequences.

In the SDFT scenario, the UVT needs to be repeatedly computed at each time index. Thus, the computational requirement of the UVT can be reduced further if the UVT outputs of the current window can be obtained by exploiting the intermediate calculations of the previous window. From its definition, we simply derive that the UVT calculation at successive time indexes has the following relationship:

$$\begin{aligned} \check{D}_n^{L/2}(k) &= \sum_{p=0}^{L/2-1} d((n-1)-2p) W_{M/2}^{(p-L/2+1)k} \\ &= \check{D}_{n-1}^{L/2}(k). \end{aligned} \quad (17)$$

The above result presents that $\check{D}_n^{L/2}(k)$ is exactly the same as $\check{D}_{n-1}^{L/2}(k)$. This is an important fact to be considered because the intermediate calculations of the previous window can be reused for the current window without losing the accuracy of calculations. On the basis of these findings, we propose a sliding UVT (SUVT) algorithm.

At time index n , the proposed SUVT algorithm first computes $\check{D}_n^{L/2}(k)$ using the DIT approach. Then, according to (15), the UVT output $D_n^L(k)$ is obtained using $\check{D}_n^{L/2}(k)$ and $\check{D}_n^{M/4}(k)$, where $\check{D}_n^{M/4}(k)$ has already been obtained at the previous window position. The price to be paid for this strategy is the additional memory to maintain the necessary intermediate calculations of the previous window. Figure 1 shows an example of the proposed SUVT algorithm. It is noted that the proposed gSDFT algorithm has the same precision as the traditional FFT because the twiddle factors used in the SUVT are identical to those used in the traditional butterfly-based FFT algorithm [6].

OVERALL ALGORITHM

The proposed gSDFT algorithm repeatedly produces the DFT outputs by shifting the fixed-length window one sample at a time. For simplicity of explanation, we consider only the case $L = M/4$. At time index n , the overall algorithm for window size M proceeds as follows.

- Compute $d(n)$ using the input samples. Only one complex addition is needed in this step.
- Compute $\check{D}_n^{M/8}(k)$ using $\{d(n), d(n-2), \dots, d(n-M/4)\}$. Note that $\{d(n-2), d(n-4), \dots, d(n-M/4)\}$ have already been obtained during the computation process in the previous window positions. Since the decimation is performed $(\log_2 M - 3)$ times, the computation of $\check{D}_n^{M/8}(k)$ requires $(M/4)(\log_2 M - 3)$ complex multiplications and $(M/2)(\log_2 M - 3)$ complex additions.
- Compute the order- L UVT $D_n^{M/4}(k)$ using $\check{D}_n^{M/8}(k)$ and $\check{D}_n^{L/2}(k)$, where

$\check{D}_n^{M/8}(k)$ is precalculated in the previous window position. The resultant $D_n^{M/4}(k)$ is the input for the following step. This step requires $(M/2)$ complex multiplications and M complex additions as shown in (15). Note that additional memory of size $M/2$ is required to store $\check{D}_n^{M/8}(k)$ of the previous window.

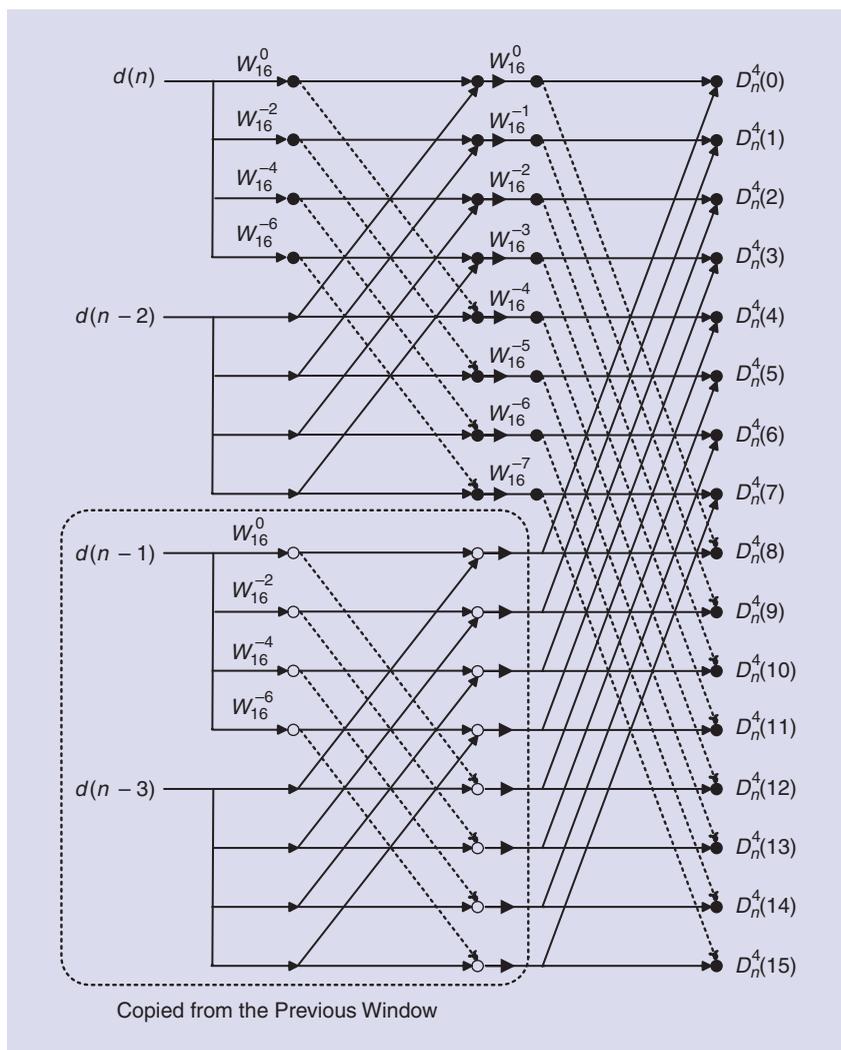
- According to (12), calculate $X_n(k)$ using $X_{n-M/4}(k)$ and $D_n^{M/4}(k)$. This step needs only $2M$ real additions.

A computational complexity comparison with varying M is presented in Table 1 where R_M and R_A denote the numbers of real multiplications and additions, respectively. In particular, when $M = 16$, the gSDFT algorithm

reduces the number of multiplications by 50% and 75% as compared to the rSDFT and mSDFT, respectively. When $M < 2^{13}$, the number of multiplications of the gSDFT is fewer than that of the mSDFT which is the state-of-the-art method. Therefore, for applications where the numerical errors should not be accumulated and the number of multiplications is the most important consideration, we would suggest the use of the gSDFT algorithm when M is less than 2^{13} . Otherwise, the mSDFT algorithm can be used for the applications.

The key characteristics of the proposed gSDFT algorithm are as follows:

- When M is less than 2^{13} , the number of multiplications of the proposed



[FIG1] The SUVT calculation for $M = 16$ and $L = 4$, where solid and dotted lines indicate the plus and minus signs, respectively. Only the filled circles need to be calculated.

[TABLE 1] THE COMPUTATIONAL REQUIREMENT FOR A COMPLEX INPUT SIGNAL.

ALGORITHM	OPERATION	WINDOW SIZE						
		4	16	64	256	1,024	4,096	M
FFT [6]	R_M	16	128	768	4,096	20,480	98,304	$2M \log_2 M$
	R_A	24	192	1152	6,144	30,720	147,456	$3M \log_2 M$
rSDFT	R_M	24	96	384	1,536	6,144	24,576	$6M$
	R_A	24	96	384	1,536	6,144	24,576	$6M$
mSDFT	R_M	48	192	768	3,072	12,288	49,152	$12M$
	R_A	40	160	640	2,560	10,240	40,960	$10M$
gSDFT	R_M	0	48	320	1,792	9,216	45,056	$M(\log_2 M - 1)$
	R_A	10	42	610	3,202	15,874	75,778	$(M/2)(3 \log_2 M + 1) + 2$

gSDFT is smaller than that of the state-of-the-art method.

- The recurrence presented in (12) is unconditionally stable and the numerical errors do not accumulate.
- The proposed algorithm produces an output signal that is mathematically equivalent to the DFT output in (1) at all time indexes.
- The SUVT can be computed independent of the recursive calculation of the SDFT. This separable feature is highly beneficial for hardware implementation.

SIMULATION RESULTS

We investigated the efficiency of the proposed gSDFT algorithm using a complex test signal, which was zero-mean Gaussian noise with a standard deviation equal to one. The simulation was performed in 64-bit double-precision arithmetic and M was set to 16. In our simulation, the numerical errors of the rSDFT [3], mSDFT [4], and gSDFT were, respectively, generated by recursively computing (4), (5), and (12) using the test signal. We repeated these recursive computations 10^6 times to accumulate the numerical errors

and then calculated the errors at each time index. All algorithms were implemented using a highly efficient ANSI-C code and the performance was evaluated on an Intel i5 3.4 GHz CPU with 8 GB RAM.

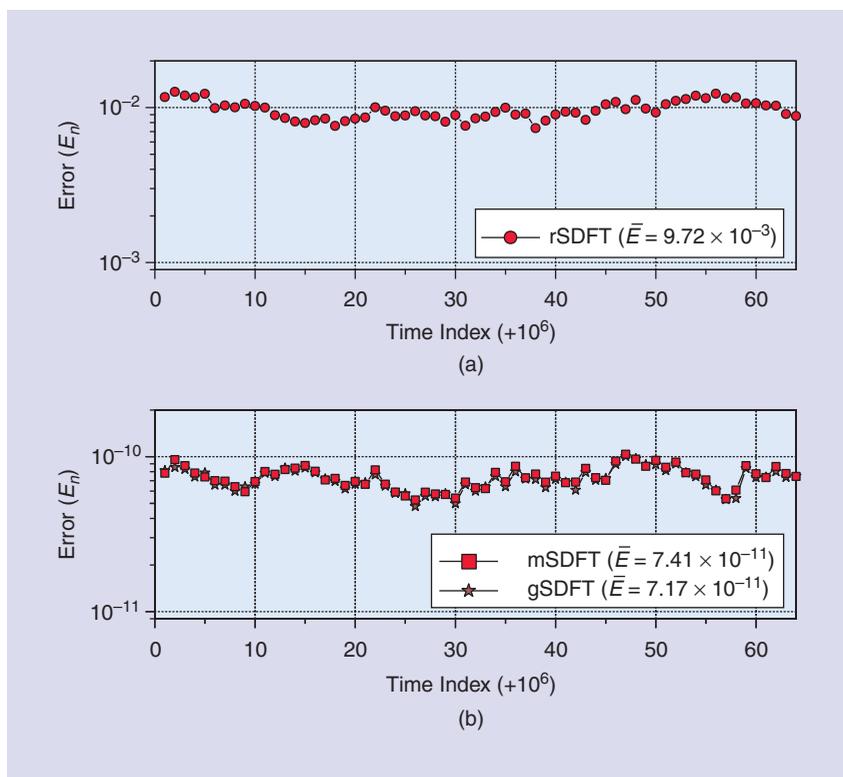
Figure 2 shows the measured results of all algorithms. The error E_n at time index n is calculated as

$$E_n = \sum_{k=0}^{M-1} |X_n^{\text{DFT}}(k) - X_n^{\text{Algorithm}}(k)| \quad (18)$$

where $X_n^{\text{DFT}}(k)$ represents the k th bin of the standard DFT in (1) and $X_n^{\text{Algorithm}}(k)$ is the k th bin of each algorithm. Figure 2 shows that the mSDFT and gSDFT algorithms significantly reduce the numerical errors as compared to the rSDFT, where the damping factor r of the rSDFT is set to 0.99999. We show in Figure 2 the average numerical error \bar{E} over 64 time indexes after the error accumulation process. In our simulation, \bar{E} s of the rSDFT, mSDFT, and gSDFT are 9.72×10^{-3} , 7.41×10^{-11} , and 7.17×10^{-11} , respectively. We can observe that the error of the gSDFT is even smaller than that of the mSDFT. Therefore, the proposed gSDFT consistently outperforms the other algorithms.

SUMMARY

In this article, we proposed a new guaranteed stable SDFT algorithm by excluding the imprecise twiddle factor from the feedback of the resonator. We first derived the general formula between the DFT bins with multihop distance. Then the proposed gSDFT algorithm was designed by exploiting the special relationship between the



[FIG2] The measured errors of the stable SDFT algorithms.

(continued on page 156)



IEEE TRANSACTIONS ON

SIGNAL AND INFORMATION PROCESSING OVER NETWORKS



Now accepting paper submissions

The new *IEEE Transactions on Signal and Information Processing over Networks* publishes high-quality papers that extend the classical notions of processing of signals defined over vector spaces (e.g. time and space) to processing of signals and information (data) defined over networks, potentially dynamically varying. In signal processing over networks, the topology of the network may define structural relationships in the data, or may constrain processing of the data. Topics of interest include, but are not limited to the following:

Adaptation, Detection, Estimation, and Learning

- Distributed detection and estimation
- Distributed adaptation over networks
- Distributed learning over networks
- Distributed target tracking
- Bayesian learning; Bayesian signal processing
- Sequential learning over networks
- Decision making over networks
- Distributed dictionary learning
- Distributed game theoretic strategies
- Distributed information processing
- Graphical and kernel methods
- Consensus over network systems
- Optimization over network systems

Communications, Networking, and Sensing

- Distributed monitoring and sensing
- Signal processing for distributed communications and networking
- Signal processing for cooperative networking
- Signal processing for network security
- Optimal network signal processing and resource allocation

Modeling and Analysis

- Performance and bounds of methods
- Robustness and vulnerability
- Network modeling and identification

Modeling and Analysis (cont.)

- Simulations of networked information processing systems
- Social learning
- Bio-inspired network signal processing
- Epidemics and diffusion in populations

Imaging and Media Applications

- Image and video processing over networks
- Media cloud computing and communication
- Multimedia streaming and transport
- Social media computing and networking
- Signal processing for cyber-physical systems
- Wireless/mobile multimedia

Data Analysis

- Processing, analysis, and visualization of big data
- Signal and information processing for crowd computing
- Signal and information processing for the Internet of Things
- Emergence of behavior

Emerging topics and applications

- Emerging topics
- Applications in life sciences, ecology, energy, social networks, economic networks, finance, social sciences, smart grids, wireless health, robotics, transportation, and other areas of science and engineering

Editor-in-Chief: Petar M. Djurić, Stony Brook University (USA)

To submit a paper, go to: <https://mc.manuscriptcentral.com/tsipn-ieee>



The High Efficiency Image File Format Standard

The High Efficiency Image File Format (HEIF) is a standard developed by the Moving Picture Experts Group (MPEG) for the storage of images and image sequences. The standard facilitates file encapsulation of data coded according to the High Efficiency Video Coding (HEVC) standard. The compression performance of HEVC is superior to any alternative image or image sequence coding format. HEIF includes a rich set of features building on top of the widely used ISO Base Media File Format (ISO/BMFF), making HEIF superior feature-wise compared to other image file formats. This article provides an overview of the performance, features, and design of HEIF.

BACKGROUND

Version 1 of the HEVC standard [1] was finalized in January 2013 and has been reported to achieve the same image quality with about half the bit rate compared to the Advanced Video Coding (AVC) standard. After the finalization of version 1 of the HEVC standard, it was realized that to take the good still picture compression performance of HEVC into use, e.g., in cameras, a format to associate photographic metadata, such as various tags of the exchangeable image file format (Exif), to the coded image is needed. Various use cases concerning sequences of images, such as photo bursts, exposure stacks, or animations, were also considered important. A requirements-gathering process was hence performed by the MPEG, resulting in accepted requirements documents for storage of HEVC still images and image sequences in August 2013.

Digital Object Identifier 10.1109/MSP.2015.2419292
Date of publication: 15 June 2015

Subsequently, a standardization project was established by MPEG, in coordination with the Joint Photographic Experts Group (JPEG). In summary, the goal of the project was to specify an image file format capable of storing one or more HEVC-coded still images (i.e., independent static images) or image sequences and associated photographic metadata into a file.

The fundamental design decisions on file format structures to store still images and image sequences were agreed upon in January 2014. Various features were then refined until a public review period was held from April to June 2015. This column is based on the draft standard available for the aforementioned public review (see the “Resources” section). The technical finalization of the standard took place in June 2015, and the standard will be published after a final approval ballot.

USE CASES AND REQUIREMENTS

HEVC-coded content in HEIF enables a wide range of use cases varying from traditional still picture capture, storage, and sharing to sophisticated multi-image use cases, such as sharing of image bursts or storing sets of images for the purpose of processing those by means of computational photography. As HEVC provides support for various chroma formats and sample fidelities up to lossless coding, the format can serve the whole spectrum of use cases from today’s consumer devices storing images typically at 8 bits per sample to high-end professional devices with sample fidelity and dynamic range requirements going all the way up to 16 bits per sample.

Computational photography forms a new category of use cases that can benefit from HEIF. Now a set of related images can be stored in a single file with

associated metadata indicating relationships between different pictures. Examples of such emerging use cases include refocusing the shot by selecting an image with a desired focus from a set of pictures captured with different focal lengths, high dynamic range photography by combining pictures with different exposures, and building omnidirectional or panoramic images from a set of pictures with connected scenery.

As opposed to legacy image codecs, HEVC also provides interesting opportunities for hybrid video and imaging use cases, such as simultaneous capture of video and still pictures, extracting still pictures from a video sequence, storing bursts of images efficiently in a single file, and creating animations that can be shared and edited easily.

In addition to the aforementioned use cases, the capability of fast “random access” to each picture in an image sequence was included in the requirements to allow for flexibility in rendering the sequence and allow editing individual images in the sequence without affecting coding of other output images. Moreover, the support for auxiliary pictures, such as alpha channels, was included in the requirements.

STILL-IMAGE CODING PERFORMANCE

HEVC includes sophisticated intra coding tools [2], making it an excellent codec for still image compression. To assess coding efficiency of the HEVC intra picture coding, we carried out a set of experiments using the common test conditions [3] of the Joint Collaborative Team on Video Coding (JCT-VC). Tests were performed with constant quality settings, aligning the objective quality—measured as peak signal-to-noise ratio

(PSNR)—for each measurement. The resulting picture quality typically varied in the range of 34–44 dB, illustrating a wide range of quality levels from typical Web usage to visually lossless. More information on the experimental setup is available in [4].

Table 1 illustrates the coding efficiency of HEVC intra coding with respect to well-known still picture codecs. The results indicate that JPEG would require, on average, 139% higher bit rates than HEVC (i.e., 2.39 times the file size) to achieve the same objective picture quality. For JPEG-XR and JPEG-2000, the average increase in bit rate is 66% and 44%, respectively. Table 1 also provides information about the objective quality differences between HEVC and legacy codecs measured using the Bjøntegaard delta PSNR metrics [5]. Subjective testing performed for different test sets appear to also verify the results when it comes to the perceived quality of material coded with different codecs. For instance, [6] reports that typically subjective quality of HEVC-intra coded pictures are comparable to that of JPEG-coded pictures using two or sometimes even four times the bit rate of HEVC.

CODING OF IMAGE SEQUENCES

HEVC image sequences can be coded either by applying intra picture coding or using predictive video-like encoding,

where dependencies between coded pictures are restricted to guarantee quick random access to individual pictures in the file. In the case of intra picture coding, each image in the sequence is simply coded as an independent still picture without reference to any of the other pictures in the sequence. While providing suboptimal coding efficiency compared to traditional video coding, HEVC intra pictures still provide substantial improvement over legacy image coding alternatives. The results in Table 1 also apply when comparing HEVC intra coded image sequences to motion JPEG and motion JPEG-2000. The benefits of intra only coding include minimal random access latency to each output picture and the possibility to edit individual pictures without affecting coding and reconstruction of the other pictures in the file.

Inter coding can provide significant coding efficiency improvement to image sequences when images are correlated. Such correlation is especially evident in content like image bursts or animation clips. However, applying unrestricted predictive video coding can lead to unacceptable decoding delay when extracting only a specific picture out of a coded file. In the worst-case scenario, the whole sequence needs to be completely decoded to be able to access the last picture in the stream. To avoid such undesired behavior, the HEVC

image sequences rely on a predefined set of intra coded reference pictures, while the use of inter coded pictures as reference pictures for predicting other inter coded pictures is disallowed.

To measure coding efficiency of low-latency HEVC encoding structures for different use cases, the following experiments were performed. First, the JCT-VC test set of Table 1 was used to mimic image bursts. Eight first frames of each sequence were coded using the fourth picture in each clip as a reference picture for inter coding. Second, additional use cases with exposure stack, focal stack, and cinemagraph content were simulated by coding a representable sequence in each category. For an exposure stack a well-known “Memorial” sequence with 16 different exposures was used. In the focal stack case a “Mersu” sequence with 13 different focus distances was selected and finally a “Car and Tractor” cinemagraph was used to represent an animated clip where the majority of the picture is frozen and only a certain area of the scene is undergoing motion. The results reported in Table 2 indicate that for natural content the restricted inter coding can typically provide two to three times better compression than intra picture coding. In special cases like animations where the majority of the scene is static, the compression efficiency can

[TABLE 1] THE HEVC INTRA CODING PERFORMANCE WITH RESPECT TO LEGACY FORMATS. BIT RATE INCREASE REQUIRED TO ACHIEVE THE OBJECTIVE QUALITY PROVIDED BY HEVC INTRA CODING IS REPORTED FOR EACH TEST CATEGORY. SIMILARLY, THE AVERAGE PSNR DIFFERENCES WITH RESPECT TO HEVC AT ALIGNED BIT RATES ARE REPORTED IN DECIBELS.

CLASS	RESOLUTION	CHARACTERISTICS	JPEG		JPEG-XR		JPEG-2000	
			DBR	DPSNR	DBR	DPSNR	DBR	DPSNR
CLASS A	2,560 × 1,600	CROPPED 4K × 2K SEQUENCES FOR ULTRA HDTV SERVICES	87%	-3.6	44%	-2.1	48%	-2.1
CLASS B	1,920 × 1,080	HIGH-RESOLUTION SEQUENCES FOR STREAMING AND BROADCAST SERVICES	124%	-3.2	62%	-1.9	15%	-0.2
CLASS C	832 × 480	MEDIUM-RESOLUTION SEQUENCES FOR INTERNET/MOBILE VIDEO SERVICES	122%	-5.5	53%	-2.6	50%	-2.5
CLASS D	416 × 240	LOW-RESOLUTION SEQUENCES FOR SERVICES TO RESOURCE-CONSTRAINED DEVICES	110%	-5.8	47%	-2.7	43%	-2.2
CLASS E	1,280 × 720	720p SEQUENCES FOR VIDEO CONFERENCING APPLICATIONS	170%	-5.2	73%	-2.9	23%	-1.0
CLASS F	1,024 × 768, 1,280 × 720	COMPUTER SCREEN CONTENT AND COMPUTER-GENERATED CONTENT	223%	-11.8	118%	-8.2	87%	-6.2
AVERAGE			139%	-5.8	66%	-3.4	44%	-2.4

standards **IN A NUTSHELL** continued
[TABLE 2] CODING EFFICIENCY IMPROVEMENTS PROVIDED BY LOW-LATENCY PREDICTIVE CODING OF THE HEVC IMAGE FILE FORMAT. BIT RATE IMPACT AND CODING GAIN ARE REPORTED WITH RESPECT TO HEVC INTRA CODING.

CONTENT	TYPE	FRAMES	BIT RATE CHANGE	CODING GAIN [dB]
CLASS A	IMAGE BURST	8	-46%	1.9
CLASS B	IMAGE BURST	8	-51%	2.0
CLASS C	IMAGE BURST	8	-60%	2.5
CLASS D	IMAGE BURST	8	-63%	2.7
CLASS E	IMAGE BURST	8	-79%	4.8
CLASS F	IMAGE BURST	8	-55%	2.2
MEMORIAL	EXPOSURE STACK	16	-29%	1.4
MERSU	FOCAL STACK	13	-25%	1.3
CAR AND TRACTOR	CINEMAGRAPH	48	-97%	33.3

significantly exceed those levels and be tens of times more efficient than intra coding. More information on the experiment is available in [4].

DERIVING FROM THE ISO BASE MEDIA FILE FORMAT

The ISOBMFF was originally designed to be a container format for timed media data, such as audio and video. It is used as the basis from which widely used container file formats, such as MP4, and media encapsulation formats, such as ISO/IEC 14496-15 for AVC and HEVC video streams, are derived.

Since one of the initial requirements of HEIF was to support multipurpose files, e.g., for the storage of still images associated with video, it was justified to take ISOBMFF as the basis of the development. To reuse existing implementations, the ISOBMFF structures and features are used as much as possible in the design of HEIF.

Files conforming to ISOBMFF consist of a sequence of data structures called *boxes*, each comprising a four-character type, the size of the box in terms of bytes,

and the payload of the box. Boxes may be nested, i.e., a box may contain other boxes. ISOBMFF specifies constraints on the allowed box order and hierarchy.

In ISOBMFF, a continuous or timed media or metadata stream is conceptually organized in a track, whereas static media or metadata is conceptually stored in items. Consequently, the following basic design was chosen for HEIF:

1) Still images are stored as items. All image items are independently coded and do not depend on any other item in their decoding. Any number of image items can be included in the same file.

2) Image sequences are stored as tracks. An image sequence track is used when there is coding dependency between images or when the playback of the images is timed. As opposed to video tracks, the timing in the image sequence track is advisory.

While the standardization work started from the requirements to store HEVC-coded images and image sequences, it became apparent that the specified structures could be equally used

for any coding format. The HEIF specification is written in a way that other image container file formats can be derived and the support of other coding formats in addition to HEVC can be added in derived formats, similarly to how ISOBMFF can be used as a basis for derived file formats. The nickname of the standard, HEIF, refers to the structural file format specification, whereas the nickname HEVC Image File Format, or HEIC, can be used for the specified encapsulation of HEVC-coded images in HEIF files.

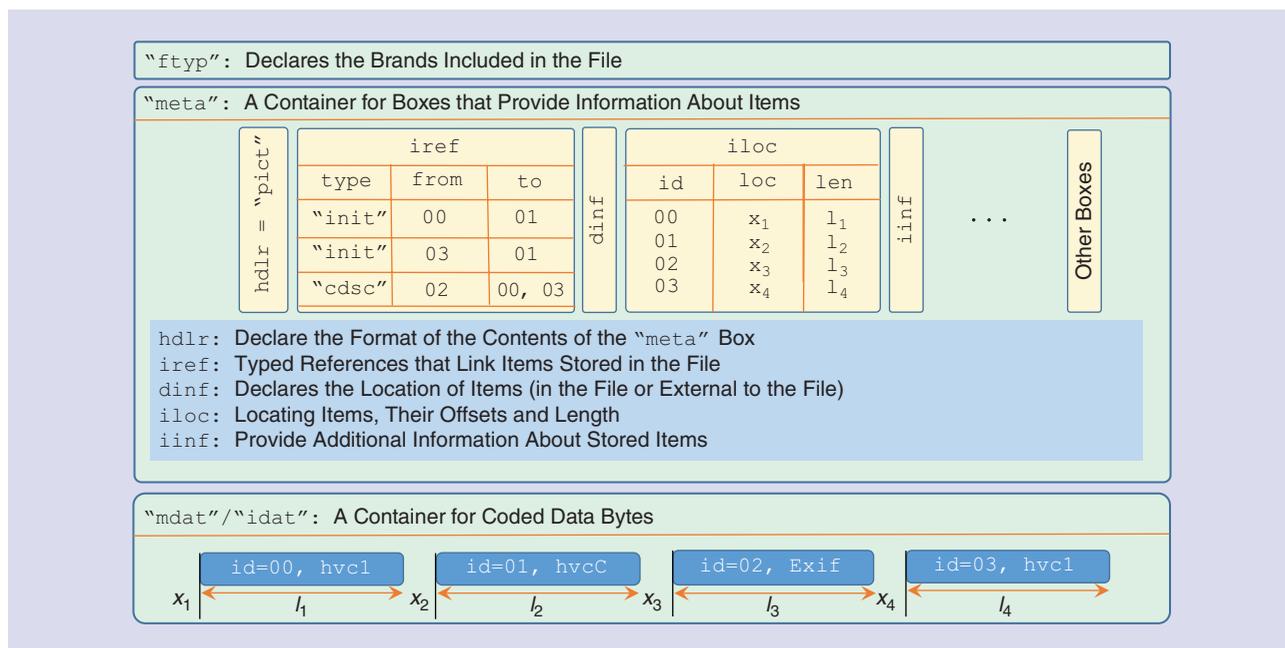
BRANDING AND SIGNALING

Files conforming to HEIF start with a File Type box (“`ftyp`”), which contains a list of brands to which the file conforms. Each brand is identified by its unique four-character code. The specification of a brand can include requirements and constraints for files of the brand and for file players supporting the brand. A brand included in the File Type box permits a player that supports the requirements of the brand to play the file.

The brands specified in the HEIF standard are presented in Table 3. The HEIF standard specifies the `mif1` and `msf1` structural brands, where requirements on file structures present in the file and to be supported by players, are given, but any image coding format can be used. Additionally, HEVC-specific brands are specified as listed in Table 3. As the File Type box is located at the start of the file, it provides easily accessible indications of the file contents to file players. It can be expected that the Main profile of HEVC will be most widely implemented out of all the HEVC profiles. Hence, it was decided to specify dedicated brand names, `heic` and

[TABLE 3] A MAPPING OF BRANDS SPECIFIED IN THE HEIF STANDARD TO THE CODING FORMAT, THE TYPE OF THE BRAND (A STILL IMAGE BRAND OR AN IMAGE SEQUENCE BRAND), THE RESPECTIVE MIME SUBTYPE, AND THE RECOMMENDED FILE NAME EXTENSION.

BRAND	CODING FORMAT	TYPE	MIME SUBTYPE	FILE EXTENSION
<code>mif1</code>	ANY	IMAGE	<code>heif</code>	<code>.heif</code>
<code>msf1</code>	ANY	SEQUENCE	<code>heif-sequence</code>	<code>.heif</code>
<code>heic</code>	HEVC (MAIN OR MAIN STILL PICTURE PROFILE)	IMAGE	<code>heic</code>	<code>.heic</code>
<code>heix</code>	HEVC (MAIN 10 OR FORMAT RANGE EXTENSIONS PROFILE)	IMAGE	<code>heic</code>	<code>.heic</code>
<code>hevc</code>	HEVC (MAIN OR MAIN STILL PICTURE PROFILE)	SEQUENCE	<code>heic-sequence</code>	<code>.heic</code>
<code>hevx</code>	HEVC (MAIN 10 OR FORMAT RANGE EXTENSIONS PROFILE)	SEQUENCE	<code>heic-sequence</code>	<code>.heic</code>



[FIG1] An example image file including two HEVC-coded images with Exif metadata.

hevc, for the Main profile compatible image files. This design allows players that support only the Main profile to determine whether the playback of the file is possible by inspecting the File Type box.

Internet media types, also known as *multipurpose Internet mail extensions (MIME)* types, are used by various applications to identify the type of a resource or a file. MIME types consist of a media type ("image" in the case of HEIF files), a subtype, and zero or more optional parameters. For multipurpose files, the selection of the subtype can be made on the basis of the primary use of the file. An optional `codecs` parameter can be present to indicate the used coding formats and the information on the tracks and items present in the file. MIME types can be used, e.g., with the picture element of Hypertext Markup Language, version 5 (HTML5) when offering several versions of the same original image using different image file formats and letting the browser to select which offered file suits its capabilities. This mechanism provides a forward compatible mechanism to take HEIF into use, while legacy browsers can continue to receive and decode, e.g., respective JPEG images.

STORAGE OF STILL IMAGES

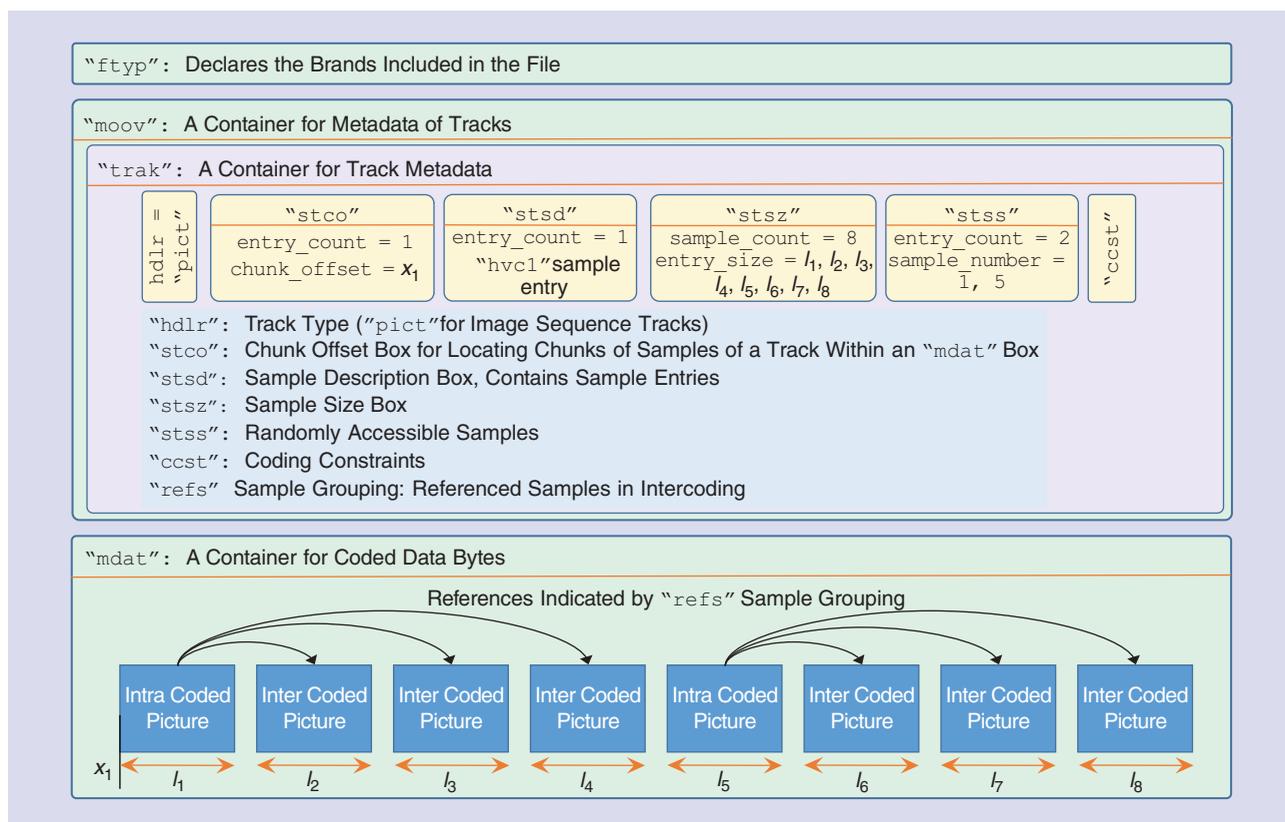
The illustration in Figure 1 shows an example of two HEVC-coded images along with their Exif metadata stored in a file conforming to the HEVC Image File Format. The file metadata for items is stored within a Meta box ("meta"). The handler type is set to "pict" indicating to a reader that this Meta box handles images. The coded images themselves are stored as items of "hvc1", indicating HEVC compression. The coded data for the images is contained either in a Media Data box ("mdat") or in an Item Data box ("idat"). The syntax of the "hvc1" items is identical to the sample syntax of HEVC video and image sequence tracks, which allows the data of intra coded images to be described both as an item and as a sample of a track. Basic image metadata, such as the width and the height, and all configuration information (e.g., parameter sets and information about the coding itself) required to initialize the decoder is stored as an item of type "hvcC" (for HEVC-coded images).

Relations between items are provided as entries in the Item Reference box ("iref"). A relation between the coded image item ("hvc1") and its initialization item ("hvcC") is included as an entry

type "init" in the Item Reference box. Exif metadata for the image can be optionally included in a file as an item of type "Exif" and linked to the image item using the "cdsc" reference type. Other metadata schemes, such as Extensible Metadata Platform (XMP), could alternatively or additionally be used in a similar fashion as Exif. Image items can (but need not) share the same initialization data, as the same item of type "hvcC" can be linked to more than one "hvc1" image items in the Item Reference box, as illustrated in Figure 1. The same applies to associating the same metadata item with multiple image items, which is also demonstrated in Figure 1.

Given a set of images in a file, it is sometimes essential to qualify certain relationships between images. HEIF facilitates indicating a cover image for a file, providing thumbnail images, and associating some or all of the images with an auxiliary image such as an alpha plane.

The design of HEIF also considers the use case of nondestructive image editing, in which the original coded image itself is kept in the file, and the decoded image is transformed in a signaled way to a derived image before presentation. A few basic image editing operations—rotation by 90° multiples, cropping, overlaying,

standards **IN A NUTSHELL** continued

[FIG2] An example image sequence file where all reference pictures are intra coded.

and tiling—are included in HEIF, while the design is flexible for adding other editing operations as extensions. Derived images are treated similarly to coded images in HEIF; e.g, a derived image can act as an input image in the construction of another derived image. Players supporting the HEVC image brands are required to handle 90° rotations and cropping, hence re-encoding of the images is not needed to guarantee that these simple operations are correctly carried out in all players.

STORAGE OF IMAGE SEQUENCES

Figure 2 provides a simplified example of a file containing an image sequence. To keep the figure simple, some mandatory boxes are omitted and the box nesting structure is not illustrated fully. The file metadata for the tracks is stored in the Movie box ("`moov`") and its child boxes. The file metadata includes for example information where in the file to locate the coded media data for the tracks. Each track has a type, identified by a four-character handler code. HEIF specifies a new

handler type "`pict`" for image sequences. Tracks comprise samples, such as audio or video frames, which are stored in the Media Data box ("`mdat`") or in an external file. Samples that occupy a contiguous byte range form a chunk. The location of chunks within the file is described with the Chunk Offset box ("`stco`"). In Figure 2, all samples of the track are in the same chunk. Each chunk is associated with one of the sample entries that are provided within the Sample Description box ("`stsd`"). A sample entry includes decoder initialization information, such as HEVC parameter sets, for the linked samples. In Figure 2, all samples use the same initialization data and hence only one sample entry is defined for the track. The Sample Size box ("`stsz`") contains the length of the samples, hence enabling to locate individual samples within each chunk.

As mentioned previously, random access was one of the key requirements for HEIF. ISO/BMFF specifies the Sync Sample box ("`stss`"), which lists the intra pictures and hence facilitates their random access. HEIF introduces two

features assisting in accessing inter coded pictures: the Coding Constraints box ("`ccst`") and the referenced samples ("`refs`") sample group. The Coding Constraints box is required to be present in image sequence tracks and indicates constraints that were used in the encoding of the image sequence. For example, in HEVC-specific image sequence brands, the Coding Constraints box is required to indicate that all reference pictures for inter prediction are intra coded pictures. The "`refs`" sample group is used to indicate the reference pictures for each inter coded picture. Consequently, decoding of any particular desired picture in an image sequence can be done by first inspecting from the Sync Sample box if the picture is intra coded, in which case the picture can be decoded right away. Otherwise, the "`refs`" sample group can be inspected to identify and decode the reference pictures first, before decoding the desired picture.

An image sequence track can be used for a set of images with inter picture prediction for improved compression performance, or when an image sequence is

[TABLE 4] A COMPARISON OF THE FEATURES OF SOME IMAGE FILE FORMATS.

	.HEIC	JPEG/EXIF	PNG	GIF (89A)	WEBP
FORMATS AND EXTENSIBILITY					
BASE CONTAINER FILE FORMAT	ISOBMFF	TIFF	–	–	RIFF
LOSSY COMPRESSION	YES (HEVC)	YES (JPEG)	NO	NO	YES (VP8)
LOSSLESS COMPRESSION	YES (HEVC)	YES (TIFF REV 6.0)	YES (PNG) ¹	YES (GIF) ¹	YES (VP8L)
EXTENSIBLE TO OTHER CODING FORMATS	YES	YES	NO	NO	NO
METADATA FORMAT (ON TOP OF INTERNAL)	EXIF, XMP, MPEG-7	EXIF	–	–	EXIF, XMP
EXTENSIBLE TO OTHER METADATA FORMATS	YES	NO	NO	NO	NO
OTHER MEDIA TYPES (AUDIO, TEXT, ETC.)	YES	ONLY AUDIO ²	NO	NO	NO
MULTIPICTURE FEATURES					
MULTIPLE IMAGES IN THE SAME FILE	YES	YES (MP EXT.)	NO	YES ³	YES ³
IMAGE SEQUENCES / ANIMATIONS	YES	NO	NO	YES	YES
INTER CODING	YES	NO	NO	NO	NO
DERIVED IMAGES					
MULTIPLE-OF-90° ROTATIONS	YES	YES	NO	NO	NO
CROPPING	YES	NO	NO	NO	NO
TILING/OVERLAYING	YES	NO	NO	NO	YES
EXTENSIBLE TO OTHER EDITING OPERATIONS	YES	NO	NO	NO	NO
AUXILIARY PICTURE INFORMATION					
TRANSPARENCY	YES	NO	YES	FULL TRANSPARENCY	YES
THUMBNAIL IMAGE	YES	YES	NO	NO	NO

¹ In GIF and indexed color PNG encoding, lossy color quantization is applied while the color-quantized image is losslessly compressed.

² PCM, μ -Law PCM, and ADPCM encapsulated in RIFF WAV.

³ Only for animations and tiling/overlying.

associated with timing information. In the former case, an untimed playlist, such as showing the images in a gallery view, should be chosen by the player, while in the latter case, the image sequence should be played as a timed slideshow. HEIF provides the means to indicate which one of these two is the preferred playback method.

ISOBMFF provides rich features for controlling the playback of tracks, such as cutting away parts of the track and changing the playback pace. For a tutorial on the timing and track-editing features of ISOBMFF, see [7]. In addition, HEIF introduces a functionality to loop tracks, which can be useful, e.g., in cinemagraphs and animations to achieve infinite playback.

HEIF enables the grouping of image items and tracks of a file into alternate groups. Players should select exactly one image item or track of an alternate group for playback. The items and tracks of an alternate group are listed in their preference order. This feature enables

the creation of multibranded files, where, e.g., a cinemagraph and a still image are indicated to be alternatives to each other, and both a still image brand and an image sequence brand are included in the File Type box. In another example, the same original image is represented by two alternative images in the file, one coded with the HEVC Main profile at $1,920 \times 1,080$ resolution and another coded with the HEVC Main 10 profile at $3,840 \times 2,160$ resolution, and both “heic” and “heix” brands are included in the File Type box.

Similar to image items, image sequence tracks may be accompanied by thumbnail and auxiliary image sequence tracks. Rotation by multiples of 90° and cropping can be indicated for image sequence tracks, and HEVC-specific image sequence brands are required to support them.

Metadata storage for image sequences reuses the existing ISOBMFF mechanisms, i.e., untimed metadata can be included in the Meta box of a track,

whereas timed metadata can use a metadata track linked to the image sequence track. For example, Exif metadata that applies to all samples of an image sequence track can be stored in the Meta box of a track. Metadata applying to certain but not all samples of a track is included as samples of a timed metadata track, and a metadata sample is associated with the sample of an image sequence track that has the equivalent timing information.

CONCLUSIONS

In this article, we reviewed the features and the design of HEIF that can be used for encapsulating HEVC-coded images and image sequences. We also gave some insights on the compression performance of HEVC for still image coding and for some image sequence use cases. A comparison of the features of HEIF with other available image file formats is presented in Table 4. It can be observed that HEIF is more extensible and comprehensive than the others. Particularly, the

standards **IN A NUTSHELL** continued

possibility to include other media types, the advanced multipicture features, and the support for nondestructive editing make HEIF more advanced than the legacy formats. The rich set of features makes HEIF suitable for a broad range of devices and applications, including, e.g., burst photography.

Video codecs are an integral part of many consumer devices. HEVC has made a successful market entry already, e.g., in high-end smartphones, digital still cameras, and television sets, and is expected to become widely supported. HEVC-coded images and image sequences encapsulated in HEIF files can complement and gradually replace older formats to achieve better compression and a wider range of features.

RESOURCES**IMAGE FILE FORMAT SPECIFICATIONS**

More information on the formats compared in Table 4 can be obtained through the following links.

- A draft HEIF standard (03/2015) is available at <http://mpeg.chiariglione.org/standards/mpeg-h/image-file-format/draft-text-isoiec-fdis-23008-12-carriage-still-image-and-image> and the final

HEIF standard (ISO/IEC 23008-12) is likely to appear among public ISO standards at <http://standards.iso.org/ittf/PubliclyAvailableStandards/>.

- The JPEG, Exif, PNG, and GIF specifications are available at <http://www.w3.org/Graphics/JPEG/itu-t81.pdf>, http://www.cipa.jp/std/documents/e/DC-008-2012_E.pdf, <http://www.w3.org/TR/PNG/>, and <http://www.w3.org/Graphics/GIF/spec-gif89a.txt>, respectively.

- The WebP container format is specified at https://developers.google.com/speed/webp/docs/riff_container, while the VP8L format is specified at https://developers.google.com/speed/webp/docs/webp_lossless_bitstream_specification.

- The MP4 registration authority, available at <http://www.mp4ra.org/>, contains the registrations of all four-character codes and other identifiers used in ISOBMFF and formats derived from it.

AUTHORS

Miska M. Hannuksela (miska.hannuksela@nokia.com) is a distinguished scientist with Nokia Technologies. He is an editor of the High Efficiency Image File Format standard.

Jani Lainema (jani.lainema@nokia.com) is a principal scientist with Nokia Technologies. He has contributed to the

designs of the High Efficiency Video Coding codec and the High Efficiency Image File Format standard.

Vinod K. Malamal Vadakital (vinod.malamalvadakital@nokia.com) is a senior researcher with Nokia Technologies. He is an editor of the High Efficiency Image File Format standard.

REFERENCES

- [1] J.-R. Ohm and G. J. Sullivan, "High efficiency video coding: The next frontier in video compression," *IEEE Signal Processing Mag.*, vol. 30, no. 1, pp. 152–158, Jan. 2013.
- [2] J. Lainema, F. Bossen, W.-J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792–1801, Dec. 2012.
- [3] F. Bossen, "Common HM test conditions and software reference configurations," *JCTVC-K1100*, Shanghai, China, Oct. 2012. [Online]. Available: http://phenix.int-evry.fr/jct/doc_end_user/documents/11_Shanghai/wg11/JCTVC-K1100-v1.zip
- [4] H. Roodaki-Lavasani and J. Lainema, "Efficient burst image compression using H.265/HEVC," in *Proc. SPIE Electronic Imaging*, San Francisco, Feb. 2014. doi:10.1117/12.2039419
- [5] G. Bjontegaard, "Calculation of average PSNR differences between RD-Curves," ITU-T SG16 Q.6 (VCEG) document VCEG-M33, Apr. 2001.
- [6] P. Hanhart, M. Rerabek, P. Korshunov, and T. Ebrahimi, "Subjective evaluation of HEVC intra coding for still image compression," *JCTVC-L0380*, Geneva, Switzerland, Jan. 14–23, 2013. [Online]. Available: http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L0380-v2.zip
- [7] M. M. Hannuksela, S. Döhla, and K. Murray, "The DVB file format," *IEEE Signal Processing Mag.*, vol. 29, no. 2, pp. 148–155, Mar. 2012.

SP **TIPS&TRICKS** (continued from page 148)

DFT outputs, which can be expressed without using the twiddle factor. The experimental results clearly demonstrated that the numerical error of the gSDFT is smaller than those of the existing stable SDFT algorithms. In addition, when the window size is relatively small, its total number of addition and multiplication operations is the lowest among the algorithms.

ACKNOWLEDGMENTS

I would like to thank anonymous reviewers for their valuable comments

which were very helpful in improving this article.

AUTHOR

Chun-Su Park (cspark@sejong.ac.kr) is an assistant professor in the Department of Digital Contents at Sejong University, South Korea.

REFERENCES

- [1] E. Jacobsen and R. Lyons, "An update to the sliding DFT," *IEEE Signal Processing Mag.*, vol. 21, no. 1, pp. 110–111, Jan. 2004.
- [2] S. Douglas and J. Soh, "A numerically-stable sliding-window estimator and its application to adaptive

filters," in *Proc. 31st Annu. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 1997, vol. 1, pp. 111–115.

- [3] E. Jacobsen and R. Lyons, "The sliding DFT," *IEEE Signal Processing Mag.*, vol. 20, no. 2, pp. 74–80, Mar. 2003.

- [4] K. Duda, "Accurate, guaranteed stable, sliding discrete Fourier transform," *IEEE Signal Processing Mag.*, vol. 27, no. 6, pp. 124–127, Nov. 2010.

- [5] C. S. Park and S. J. Ko, "The hopping discrete Fourier transform," *IEEE Signal Processing Mag.*, vol. 31, no. 2, pp. 135–139, Mar. 2014.

- [6] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1999, p. 635.





IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING

The new IEEE Transactions on Computational Imaging seeks original manuscripts for publication. This new journal will publish research results where computation plays an integral role in the image formation process. All areas of computational imaging are appropriate, ranging from the principles and theory of computational imaging, to modeling paradigms for computational imaging, to image formation methods, to the latest innovative computational imaging system designs. Topics of interest include, but are not limited to the following:

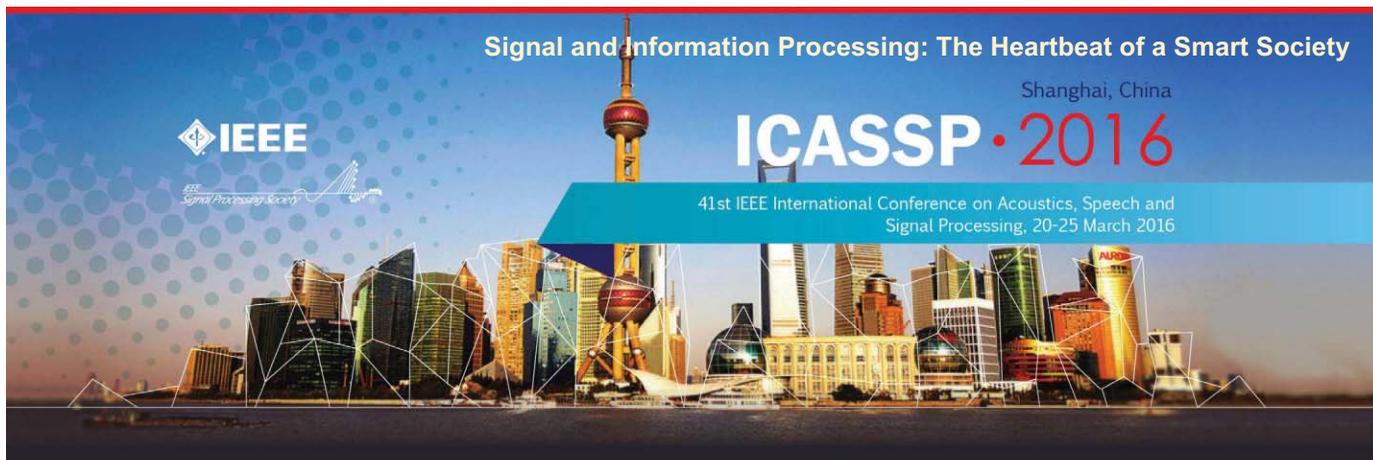
<p>Imaging Models and Representation</p> <ul style="list-style-type: none"> • Statistical-model based methods • System and image prior models • Noise models • Graphical and tree-based models • Perceptual models 	<p>Computational Photography</p> <ul style="list-style-type: none"> • Non-classical image capture, Generalized illumination • Time-of-flight imaging • High dynamic range imaging • Focal stacks 	<p>Tomographic Imaging</p> <ul style="list-style-type: none"> • X-ray CT • PET • SPECT
<p>Computational Sensing</p> <ul style="list-style-type: none"> • Coded source methods • Structured light • Coded aperture methods • Compressed sensing • Light-field sensing • Plenoptic imaging • Hardware and software systems 	<p>Computational Consumer Imaging</p> <ul style="list-style-type: none"> • Cell phone imaging • Camera-array systems • Depth cameras 	<p>Magnetic Resonance Imaging</p> <ul style="list-style-type: none"> • Diffusion tensor imaging • Fast acquisition
<p>Computational Image Creation</p> <ul style="list-style-type: none"> • Sparsity-based methods • Statistically-based inversion methods, Bayesian regularization • Super-resolution, multi-image fusion • Learning-based methods, Dictionary-based methods • Optimization-based methods; proximal iterative methods, ADMM 	<p>Computational Acoustic Imaging</p> <ul style="list-style-type: none"> • Multi-static ultrasound imaging • Photo-acoustic imaging • Acoustic tomography 	<p>Radar Imaging</p> <ul style="list-style-type: none"> • Synthetic aperture imaging • Inverse synthetic imaging • Terahertz imaging
	<p>Computational Microscopic Imaging</p> <ul style="list-style-type: none"> • Holographic microscopy • Quantitative phase imaging • Multi-illumination microscopy • Lensless microscopy 	<p>Geophysical Imaging</p> <ul style="list-style-type: none"> • Multi-spectral imaging • Ground penetrating radar • Seismic tomography
		<p>Multi-spectral Imaging</p> <ul style="list-style-type: none"> • Multi-spectral imaging • Hyper-spectral imaging • Spectroscopic imaging

Editor-in-Chief: W. Clem Karl, Boston University.

To submit a paper go to: <https://mc.manuscriptcentral.com/tci-ieee>



Digital Object Identifier 10.1109/MSP.2015.2434236



General Chairs

Zhi Ding, Univ. of California, Davis, USA
 Zhi-Quan Luo, Univ. of Minnesota, USA
 Wenjun Zhang, Shanghai Jiao Tong Univ., China

Technical Program Chairs

P. C. Ching, Chinese Univ. of Hong Kong, HK
 Dominic K.C. Ho, Univ. of Missouri, USA

Finance Chairs

Shuguang Cui, Texas A&M Univ., USA
 Rong Xie, Shanghai Jiao Tong Univ., China

Plenaries Chairs

Zhi-Pei Liang, UIUC, USA
 Björn Ottersten, Univ. of Luxembourg, Luxembourg

Special Sessions Chairs

Tim Davidson, McMaster Univ., Canada
 Jianguo Huang, Northwestern Polytech. Univ., China

Tutorials Chairs

Jian Li, Univ. of Florida, USA
 Jose Principe, Univ. of Florida, USA

Student Session Chair

Wei Zhang, Univ. of New South Wales, AU

Registration Chairs

Tongtong Li, Michigan State Univ., USA
 Xiaojun Yuan, ShanghaiTech Univ., China

Publicity Chairs

Xiaokang Yang, Shanghai Jiao Tong Univ., China
 Mounir Ghogho, Leeds Univ., UK
 Ignacio Santamaria, Univ. of Cantabria, Spain

Publication Chairs

Min Dong, Univ. of Ontario Inst. of Tech., Canada
 Thomas Fang Zheng, Tsinghua Univ., China

Industrial & Exhibit Chairs

Li Deng, Microsoft, USA
 Jinyu Li, Microsoft, USA
 Cathy Wicks, Texas Instruments, USA

Local Arrangement Chairs

Ning Liu, Shanghai Jiao Tong Univ., China
 Meixia Tao, Shanghai Jiao Tong Univ., China

Webmaster

Yi Xu, Shanghai Jiao Tong Univ., China

Workshop Chairs

Jianguo Huang, Northwestern Polytech. Univ., China
 Jiwu Huang, Sun Yat-sen Univ., China

ICASSP2016: Signal and information processing is the driving heartbeat in the development of technologies that enrich our lives and advance our society. The 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP) will be held in the Shanghai International Convention Center, Shanghai, China between March 20 and 25, 2016. The conference provides, both for researchers and developers, an engaging forum to exchange ideas and propel new developments in this field. The 2016 conference will showcase world-class presentations by internationally renowned speakers and will facilitate a fantastic opportunity to network with like-minded professionals from around the world. Topics include but are not limited to:

- Audio and acoustic signal processing
- Bio-imaging and biomedical signal processing
- Signal processing education
- Speech processing
- Industry technology tracks
- Information forensics and security
- Machine learning for signal processing
- Signal processing for Big Data
- Multimedia signal processing
- Sensor array & multichannel signal processing
- Design & implementation of signal processing systems
- Signal processing for communications & networking
- Image, video & multidimensional signal processing
- Signal processing theory & methods
- Spoken language processing
- Signal processing for the Internet of Things

Shanghai: Shanghai is the most populous city in China and one of the most populous cities in the world. A global city, Shanghai exerts influence over global commerce, finance, culture, art, fashion, research and entertainment. The city is located in the middle portion of the Chinese coast, and sits at the mouth of the Yangtze River. The city is a tourist destination renowned for its historical landmarks, such as the Bund and City God Temple, and its modern and ever-expanding Pudong skyline including the Oriental Pearl Tower. Today, Shanghai is the largest center of commerce and finance in mainland China, and has been described as the "showpiece" of the world's fastest-growing major economy.

Submission of Papers: Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possible references, and with one additional optional 5th page containing only references. A selection of best student papers will be made by the ICASSP 2016 committee upon recommendations from the Technical Committees.

Tutorial and Special Session Proposals: Tutorials will be held on March 20 and 21, 2016. Tutorial proposals must include title, outline, contact information, biography and selected publications for the presenter(s), and a description of the tutorial and the material to be distributed to participants. Special session proposals must include a topical title, rationale, session outline, contact information, and a list of invited speakers. Additional information can be found at the ICASSP 2016 website.

Signal Processing Letters: Authors of IEEE Signal Processing Letters (SPL) papers will be given the opportunity to present their work at ICASSP 2016, subject to space availability and approval by the ICASSP Technical Program Chairs. SPL papers published between January 1, 2015 and December 31, 2015 are eligible for presentation at ICASSP 2016.

Show and Tell: S&T offers a perfect stage to showcase innovative ideas in all technical areas of interest at ICASSP. S&T sessions contain demos that are highly interactive and visible. Please refer to the ICASSP 2016 website for additional information regarding demo submission.

Important Deadlines:

Special session & tutorial proposals	August 3, 2015
Notification of special session & tutorial acceptance	September 11, 2015
Submission of regular papers	September 25, 2015
Signal processing letters	December 16, 2015
Notification of paper acceptance	December 21, 2015
Revised paper upload	January 22, 2016
Author registration	January 22, 2016



Digital Object Identifier 10.1109/MSP.2015.2411564

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
IEEE Marketing Dept.	17	innovate.ieee.org	
IEEE MDL/Marketing	3	www.ieee.org/go/trymdl	
IEEE USA	9	www.ieeeusa.org/policy/govfel	+1 202 530 8347
Mathworks	CVR 4	www.mathworks.com/ltc	+1 508 647 7040
Mini-Circuits	CVR 2, 5, CVR 3	www.minicircuits.com	+1 718 934 4500

advertisers SALES OFFICES

James A. Vick

Sr. Director, Advertising

Phone: +1 212 419 7767;

Fax: +1 212 419 7589

jv.ieeemediamedia@ieee.org

Marion Delaney

Advertising Sales Director

Phone: +1 415 863 4717;

Fax: +1 415 863 4717

md.ieeemediamedia@ieee.org

Mindy Belfer

Advertising Sales Coordinator

Phone: +1 732 562 3937

Fax: +1 732 981 1855

m.belfer@ieee.org

Product Advertising

MIDATLANTIC

Lisa Rinaldo

Phone: +1 732 772 0160;

Fax: +1 732 772 0164

lr.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@ieee.org

VA, NC, SC, GA, FL, AL, MS, TN

**MIDWEST/SOUTH CENTRAL/
CENTRAL CANADA**

Darcy Giovino

Phone: +224 616 3034;

Fax: +1 847 729 4269

dg.ieeemediamedia@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.ieeemediamedia@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.ieeemediamedia@ieee.org

Europe, Africa, Middle East

Digital Object Identifier 10.1109/MSP.2015.2388979

[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

2015

[JUNE]

Third IEEE International Workshop on Compressed Sensing Theory and Its Applications to Radar, Sonar, and Remote Sensing (CoSeRa)
22–24 June, Pisa, Italy.
General Chairs: Fulvio Gini and Joachim Ender
URL: <http://www.cosera2015.iet.unipi.it/>

IEEE Signal Processing Society Summer School on Foundations and Advances in Stochastic Filtering (FASF)
22–26 June, Barcelona, Spain.
Organizers: Pau Closas and Joaquín Míguez
URL: <http://fasf2015.cttc.cat/>

16th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)
28 June–1 July, Stockholm, Sweden.
General Chairs: Joakim Jaldén and Björn Ottersten
URL: <http://www.spawc2015.org/>

IEEE International Conference on Multimedia and Expo (ICME)
29 June–3 July, Turin, Italy.
General Chairs: Enrico Magli, Stefano Tubaro, and Anthony Vetro
URL: <http://www.icme2015.ieee-icme.org/index.php>

[JULY]

Third IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)
12–15 July, Chengdu, China.
General Chairs: Yingbo Hua and Dezhong Yao
URL: <http://www.chinasip2015.org/>

[AUGUST]

IEEE Signal Processing and SP Education Workshop (SPW)
9–12 August, Salt Lake City, Utah, United States.
General Chair: Todd Moon
URL: <http://spw2015.coe.utah.edu/>

12th IEEE International Conference on Advanced Video- and Signal-Based Surveillance (AVSS)
25–28 August, Karlsruhe, Germany.
General Chairs: Jürgen Beyerer and Rainer Stiefelhagen
URL: <http://avss2015.org>

2015 23rd European Signal Processing Conference (EUSIPCO)
31 August–4 September, Nice, France.
General Chairs: Jean-Luc Dugelay and Dirk Slock
URL: <http://www.eusipco2015.org>

[SEPTEMBER]

IEEE Signal Processing Society Italy Chapter Summer School on Signal Processing (S3P)
7–11 September, Brescia, Italy.

Sensor Signal Processing for Defence (SSPD)
9–10 September, Edinburgh, United Kingdom.
General Chairs: Mike Davies, Jonathon Chambers, and Paul Thomas
URL: <http://www.sspdconference.org>

IEEE International Conference on Image Processing (ICIP)
28 September–1 October, Quebec City, Quebec, Canada.
General Chairs: Jean-Luc Dugelay and André Morin
URL: <http://www.icip2015.org/>

[OCTOBER]

IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB)
4–7 October, Montreal, Canada.
URL: <http://www.icuwb2015.org/index.html>

IEEE Workshop on Signal Processing Systems (SiPS)

14–16 October, Hangzhou, China.
General Chairs: Chaitali Chakrabarti and Nam Ling
URL: <http://www.sips2015.org/>

IEEE International Workshop on Multimedia Signal Processing (MMSP)

19–21 October, Xiamen, China.
General Chairs: Xiao-Ping Zhang, Oscar C. Au, and Jonathan Li
URL: <http://www.mmsp2015.org/>

[NOVEMBER]

Seventh IEEE International Workshop on Information Forensics and Security (WIFS)

16–19 November, Rome, Italy.
General Chairs: Patrizio Campisi and Nasir Memon
URL: <http://www.wifs2015.org/>

[DECEMBER]

IEEE 6th International Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP)
13–16 December, Cancun, Mexico.
URL: <http://inspire.rutgers.edu/camsap2015/>

IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)

13–17 December, Scottsdale, Arizona, United States.
URL: <http://www.asru2015.org/>

IEEE Global Conference on Signal and Information Processing (GlobalSIP)

14–16 December, Orlando, Florida, United States.
General Chairs: José M.F. Moura and Dapeng Oliver Wu
URL: <http://2015.ieeeglobalsip.org/>

2016

[MARCH]

41st IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
21–25 March, Shanghai, China.
General Chairs: Zhi Ding, Zhi-Quan Luo, and Wenjun Zhang
URL: <http://icassp2016.org>

Digital Object Identifier 10.1109/MSP.2015.2409018

Date of publication: 15 June 2015





MODULAR TEST SYSTEMS

Built Your Way and Delivered within 2 Weeks!

Signal Routing & Attenuation Control for Production Test, R&D and More!

Mini-Circuits' new ZTM-Series RF test systems dramatically accelerate custom solutions for a wide range of applications in test environments. Choose from our lineup of extra-long-life SPDT, SP4T, SP6T and transfer switches, and programmable attenuators with attenuation ranges of 0 to 30, 60, or 90, 110 or 120 dB. We'll build and ship a solution tailored to your exact requirements *within just 2 weeks!*

It's that simple! Give us a call and talk to our engineers about how Mini-Circuits' ZTM-Series custom rack mount test solutions can improve efficiency, increase throughput, and save cost in your business!

Features

- Rugged 19" Rack Mountable Chassis
- Customizable Front Panel Layout
- Light Weight
- USB and Ethernet Control
- User-friendly GUI and DLLs Included
- Qualified to 100 Million Switch Cycles
- Affordable Cost
- **Delivery within 2 Weeks!**

Choose from hundreds of possible configurations!



SPDT Switches
DC – 18 GHz



SP4T Switches
DC – 18 GHz



SP6T Switches
DC – 12 GHz



Transfer Switches
DC – 18 GHz



0 – 30, 60, 90, 110 or 120 dB
Programmable Attenuators
1 MHz – 6 GHz

Mini-Circuits®

www.minicircuits.com

P.O. Box 350166, Brooklyn, NY 11235-0003 (718) 934-4500 sales@minicircuits.com

527 rev. B



您会说 MATLAB 吗?

Over one million people around the world speak MATLAB. Engineers and scientists in every field from aerospace and semiconductors to biotech, financial services, and earth and ocean sciences use it to express their ideas. Do you speak MATLAB?



Solar Image taken by the X-Ray Telescope: supplied courtesy of Smithsonian Astrophysical Observatory.

See related article at mathworks.com/ltc

MATLAB[®]
The language of technical computing

Image Credit: SAO, NASA, JAXA/ISAS, NAOJ ©2011 The MathWorks, Inc.

IEEE SIGNAL PROCESSING SOCIETY

CONTENT GAZETTE

[ISSN 2167-5023]



JULY 2015



IEEE International Symposium on Biomedical Imaging

April 13-16, 2016, Prague, Czech Republic



CALL FOR PAPERS

The IEEE International Symposium on Biomedical Imaging (ISBI) is a scientific conference dedicated to mathematical, algorithmic, and computational aspects of biomedical imaging, across all scales of observation. It fosters knowledge transfer among different imaging communities and contributes to an integrative approach to biomedical imaging.

ISBI is a joint initiative from the IEEE Signal Processing Society (SPS) and the IEEE Engineering in Medicine and Biology Society (EMBS). The 2016 meeting will include tutorials, and a scientific program composed of plenary talks, invited special sessions, challenges, as well as oral and poster presentations of peer-reviewed papers.

High-quality papers are requested containing original contributions to the topics of interest including image formation and reconstruction, computational and statistical image processing and analysis, dynamic imaging, visualization, image quality assessment, and physical, biological, and statistical modeling. Accepted 4-page regular papers will be published in the symposium proceedings published by IEEE and included in IEEE Xplore.

To encourage attendance by a broader audience of imaging scientists and offer additional presentation opportunities, ISBI 2016 will continue to propose a second track featuring posters selected from 1-page abstract submissions without subsequent archival publication.

Venue:

ISBI 2016 will be held in the 4-star Clarion Congress Hotel Prague, one of the most modern congress hotels in Prague, Czech Republic. The hotel has space for up to 2500 delegates and a corresponding accommodation capacity, including a wellness and fitness center. It takes less than 15 minutes by Underground to reach the historical center of Prague, a UNESCO Heritage Site. Do not miss Prague's world-famous Old Town Square, Charles Bridge, and Prague Castle.



Important Dates:

4-page paper submission
August 3rd - October 26th, 2015

Author Notification for 4-page papers
December 23rd, 2015

**Final version of 4-page papers
& registration**
January 11th, 2016

Conference Chairs

Jan Kybic
Czech Technical University in Prague

Milan Sonka
The University of Iowa

Program Chairs

Karl Rohr
University of Heidelberg

Boudewijn Lelieveldt
Leiden University Medical Center, Netherlands

Organizing Committee

Joe Reinhardt, Bram van Ginneken, Franjo Pernus, Punam Saha, Mathews Jacob, Arrate Munoz-Barrutia, Zoltan Szabo, Radim Krupicka, Michal Kozubek, Ipek Oguz, Tomaz Vrtovec, Jiri Jan, Jiri Janacek, Lucie Kubinova, Pavel Tomancak, Eduardo Romero, Juan David Garcia, Jan Petr, Michal Sofka

Contact

Janice Sandler j.sandler@ieee.org

<http://biomedicalimaging.org/2016>

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



MAY 15, 2015

VOLUME 63

NUMBER 10

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

Estimation of Overspread Scattering Functions http://dx.doi.org/10.1109/TSP.2015.2403309	<i>G. Pfander and P. Zheltov</i>	2451
A New Model for Array Spatial Signature for Two-Layer Imaging With Applications to Nondestructive Testing Using Ultrasonic Arrays http://dx.doi.org/10.1109/TSP.2015.2403273	<i>N. Moallemi and S. Shahbazpanahi</i>	2464
Polynomial-Time Algorithms for the Exact MMOSPA Estimate of a Multi-Object Probability Density Represented by Particles http://dx.doi.org/10.1109/TSP.2015.2403292	<i>M. Baum, P. Willett, and U. D. Hanebeck</i>	2476
Joint Source Estimation and Localization http://dx.doi.org/10.1109/TSP.2015.2404311	<i>S. Sahnoun and P. Comon</i>	2485
Actor Merging for Dataflow Process Networks http://dx.doi.org/10.1109/TSP.2015.2411229	<i>J. Boutellier, J. Ersfolk, J. Lilius, M. Mattavelli, G. Roquier, and O. Silvén</i>	2496
Unified Codebook Design for Vector Channel Quantization in MIMO Broadcast Channels http://dx.doi.org/10.1109/TSP.2015.2414895	<i>J. H. Lee and W. Choi</i>	2509

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 \$602.50 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 of \$31.00 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 © 2015 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **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.



Performance Analysis of Antenna Selection in Two-Way Relay Networks http://dx.doi.org/10.1109/TSP.2015.2414904	
..... <i>K. Song, B. Ji, Y. Huang, M. Xiao, and L. Yang</i>	2520
Shallow Water Acoustic Channel Modeling Based on Analytical Second Order Statistics for Moving Transmitter/Receiver http://dx.doi.org/10.1109/TSP.2015.2411219	
..... <i>E. Baktash, M. J. Dehghani, M. R. F. Nasab, and M. Karimi</i>	2533
Low Complexity CFO Compensation in Uplink OFDMA Systems With Receiver Windowing http://dx.doi.org/10.1109/TSP.2015.2413379 ..	
..... <i>A. Farhang, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny</i>	2546
Channel State Tracking for Large-Scale Distributed MIMO Communication Systems http://dx.doi.org/10.1109/TSP.2015.2407316	
..... <i>D. R. Brown, R. Wang, and S. Dasgupta</i>	2559
Signal Recovery from Random Measurements via Extended Orthogonal Matching Pursuit http://dx.doi.org/10.1109/TSP.2015.2413384	
..... <i>S. K. Sahoo and A. Makur</i>	2572
Sparsity-Aware Sensor Collaboration for Linear Coherent Estimation http://dx.doi.org/10.1109/TSP.2015.2413381	
..... <i>S. Liu, S. Kar, M. Fardad, and P. K. Varshney</i>	2582
Estimation of Spatially Correlated Random Fields in Heterogeneous Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2015.2412917	
..... <i>I. Nevat, G. W. Peters, F. Septier, and T. Matsui</i>	2597
On Decentralized Estimation With Active Queries http://dx.doi.org/10.1109/TSP.2015.2411223	
..... <i>T. Tsiligkaridis, B. M. Sadler, and A. O. Hero</i>	2610
Asymptotically Efficient Estimation of a Nonlinear Model of the Heteroscedasticity and the Calibration of Measurement Systems http://dx.doi.org/10.1109/TSP.2015.2413375	
..... <i>I. Nikiforov</i>	2623
Robust Joint Source-Relay-Destination Design Under Per-Antenna Power Constraints http://dx.doi.org/10.1109/TSP.2015.2414892	
..... <i>H. Tang, W. Chen, and J. Li</i>	2639
Performance Analysis of Time-Reversal MUSIC http://dx.doi.org/10.1109/TSP.2015.2417507	
..... <i>D. Ciuonzo, G. Romano, and R. Solimene</i>	2650
Subspace Learning and Imputation for Streaming Big Data Matrices and Tensors http://dx.doi.org/10.1109/TSP.2015.2417491	
..... <i>M. Mardani, G. Mateos, and G. B. Giannakis</i>	2663
Constrained and Preconditioned Stochastic Gradient Method http://dx.doi.org/10.1109/TSP.2015.2412919	
..... <i>H. Jiang, G. Huang, P. A. Wilford, and L. Yu</i>	2678
Performance Tradeoffs for Networked Jump Observer-Based Fault Diagnosis http://dx.doi.org/10.1109/TSP.2015.2419188	
..... <i>D. Dolz, I. Peñarrocha, and R. Sanchis</i>	2692
Sensing and Recognition When Primary User Has Multiple Transmit Power Levels http://dx.doi.org/10.1109/TSP.2015.2415751	
..... <i>F. Gao, J. Li, T. Jiang, and W. Chen</i>	2704
Source Enumeration in Array Processing Using a Two-Step Test http://dx.doi.org/10.1109/TSP.2015.2414894	
..... <i>Z. Lu and A. M. Zoubir</i>	2718

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



JUNE 1, 2015

VOLUME 63

NUMBER 11

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

Diffusion LMS Over Multitask Networks http://dx.doi.org/10.1109/TSP.2015.2412918	<i>J. Chen, C. Richard, and A. H. Sayed</i>	2733
Multicell Coordinated Beamforming With Rate Outage Constraint—Part I: Complexity Analysis http://dx.doi.org/10.1109/TSP.2015.2414907	<i>W.-C. Li, T.-H. Chang, and C.-Y. Chi</i>	2749
Multicell Coordinated Beamforming With Rate Outage Constraint—Part II: Efficient Approximation Algorithms http://dx.doi.org/10.1109/TSP.2015.2414896	<i>W.-C. Li, T.-H. Chang, and C.-Y. Chi</i>	2763
Model Order Selection Based on Information Theoretic Criteria: Design of the Penalty http://dx.doi.org/10.1109/TSP.2015.2414900	<i>A. Mariani, A. Giorgetti, and M. Chiani</i>	2779
Detection of Dependent Heavy-Tailed Signals http://dx.doi.org/10.1109/TSP.2015.2408570	<i>A. Subramanian, A. Sundaresan, and P. K. Varshney</i>	2790
Generalized Consistent Error Estimator of Linear Discriminant Analysis http://dx.doi.org/10.1109/TSP.2015.2419190	<i>A. Zollanvari and E. R. Dougherty</i>	2804

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 \$602.50 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 of \$31.00 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 © 2015 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **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.



Maximum Entropy PDF Design Using Feature Density Constraints: Applications in Signal Processing http://dx.doi.org/10.1109/TSP.2015.2419189	<i>P. M. Baggenstoss</i>	2815
FFT Interpolation From Nonuniform Samples Lying in a Regular Grid http://dx.doi.org/10.1109/TSP.2015.2419178	<i>J. Selva</i>	2826
On the Efficiency of Far-Field Wireless Power Transfer http://dx.doi.org/10.1109/TSP.2015.2417497	<i>M. Xia and S. Aïssa</i>	2835
Distortion Minimization in Multi-Sensor Estimation Using Energy Harvesting and Energy Sharing http://dx.doi.org/10.1109/TSP.2015.2416682	<i>S. Knorn, S. Dey, A. Ahlén, and D. E. Quevedo</i>	2848
Distributed Power System State Estimation Using Factor Graphs http://dx.doi.org/10.1109/TSP.2015.2413297	<i>P. Chavali and A. Nehorai</i>	2864
Performance Bounds for Grouped Incoherent Measurements in Compressive Sensing http://dx.doi.org/10.1109/TSP.2015.2412912	<i>A. C. Polak, M. F. Duarte, and D. L. Goeckel</i>	2877
Stability and Performance Limits of Adaptive Primal-Dual Networks http://dx.doi.org/10.1109/TSP.2015.2415759	<i>Z. J. Towfic and A. H. Sayed</i>	2888
ABORT-Like Detection Strategies to Combat Possible Deceptive ECM Signals in a Network of Radars http://dx.doi.org/10.1109/TSP.2015.2415754	<i>A. Coluccia and G. Ricci</i>	2904
On the Characterization of ℓ_p -Compressible Ergodic Sequences http://dx.doi.org/10.1109/TSP.2015.2419183	<i>J. F. Silva and M. S. Derpich</i>	2915
Asymptotically Optimal Anomaly Detection via Sequential Testing http://dx.doi.org/10.1109/TSP.2015.2416674	<i>K. Cohen and Q. Zhao</i>	2929
Compressive Temporal Higher Order Cyclostationary Statistics http://dx.doi.org/10.1109/TSP.2015.2415760	<i>C. W. Lim and M. B. Wakin</i>	2942
Linear Convergence of Adaptively Iterative Thresholding Algorithms for Compressed Sensing http://dx.doi.org/10.1109/TSP.2015.2412915	<i>Y. Wang, J. Zeng, Z. Peng, X. Chang, and Z. Xu</i>	2957
Consensus on State and Time: Decentralized Regression With Asynchronous Sampling http://dx.doi.org/10.1109/TSP.2015.2416687	<i>H.-T. Wai and A. Scaglione</i>	2972
Distributed Sparse Total Least-Squares Over Networks http://dx.doi.org/10.1109/TSP.2015.2416671	<i>S. Huang and C. Li</i>	2986



IEEE TRANSACTIONS ON



SIGNAL AND INFORMATION PROCESSING OVER NETWORKS

Now accepting paper submissions

The new *IEEE Transactions on Signal and Information Processing over Networks* publishes high-quality papers that extend the classical notions of processing of signals defined over vector spaces (e.g. time and space) to processing of signals and information (data) defined over networks, potentially dynamically varying. In signal processing over networks, the topology of the network may define structural relationships in the data, or may constrain processing of the data. Topics of interest include, but are not limited to the following:

Adaptation, Detection, Estimation, and Learning

- Distributed detection and estimation
- Distributed adaptation over networks
- Distributed learning over networks
- Distributed target tracking
- Bayesian learning; Bayesian signal processing
- Sequential learning over networks
- Decision making over networks
- Distributed dictionary learning
- Distributed game theoretic strategies
- Distributed information processing
- Graphical and kernel methods
- Consensus over network systems
- Optimization over network systems

Communications, Networking, and Sensing

- Distributed monitoring and sensing
- Signal processing for distributed communications and networking
- Signal processing for cooperative networking
- Signal processing for network security
- Optimal network signal processing and resource allocation

Modeling and Analysis

- Performance and bounds of methods
- Robustness and vulnerability
- Network modeling and identification

Modeling and Analysis (cont.)

- Simulations of networked information processing systems
- Social learning
- Bio-inspired network signal processing
- Epidemics and diffusion in populations

Imaging and Media Applications

- Image and video processing over networks
- Media cloud computing and communication
- Multimedia streaming and transport
- Social media computing and networking
- Signal processing for cyber-physical systems
- Wireless/mobile multimedia

Data Analysis

- Processing, analysis, and visualization of big data
- Signal and information processing for crowd computing
- Signal and information processing for the Internet of Things
- Emergence of behavior

Emerging topics and applications

- Emerging topics
- Applications in life sciences, ecology, energy, social networks, economic networks, finance, social sciences, smart grids, wireless health, robotics, transportation, and other areas of science and engineering

Editor-in-Chief: Petar M. Djurić, Stony Brook University (USA)

To submit a paper, go to: <https://mc.manuscriptcentral.com/tsipn-ieee>



NEW PUBLICATION:**Transactions on Signal and Information Processing over Networks (T-SIPN)***

<http://www.signalprocessingsociety.org/publications/periodicals/tsipn/>

>>We are accepting paper submissions: please [submit a manuscript here](#)<<

There has been an explosion of research in network systems of various types, including physical, engineered, biological and social systems. Its aim is to find answers to fundamental questions about the systems and with them be able to understand, predict, and control them better. To that end, a core area of work is signal and information processing over networks.

Network systems represent a growing research field encompassing numerous disciplines in science and engineering. Their complexity is reflected in the diversity and the interconnectivity of their elements, which have the capacity to adapt and learn from experience. Applications of network systems are wide and include communications (wireless sensor networks, peer-to-peer networks, pervasive mobile networks, the Internet of Things), the electric power grid, biology, the Internet, the stock market, ecology, and in animal and human societies.

The Transactions on Signal and Information Processing over Networks (T-SIPN) publishes timely peer-reviewed technical articles on advances in the theory, methods, and algorithms for signal and information processing, inference, and learning in network systems. The following core topics define the scope of the Transaction:

Adaptation, Detection, Estimation, and Learning (ADEL)

- Distributed detection and estimation (ADEL-DDE)
- Distributed adaptation over networks (ADEL-DAN)
- Distributed learning over networks (ADEL-DLN)
- Distributed target tracking (ADEL-DTT)
- Bayesian learning; Bayesian signal processing (ADEL-BLSP)
- Sequential learning over networks (ADEL-SLN)
- Decision making over networks (ADEL-DMN)
- Distributed dictionary learning (ADEL-DDL)
- Distributed game theoretic strategies (ADEL-DGTS)
- Distributed information processing (ADEL-DIP)
- Graphical and kernel methods (ADEL-GKM)
- Consensus over network systems (ADEL-CNS)
- Optimization over network systems (ADEL-ONS)

Communications, Networking, and Sensing (CNS)

- Distributed monitoring and sensing (CNS-DMS)
- Signal processing for distributed communications and networking (CNS-SPDCN)
- Signal processing for cooperative networking (CNS-SPCN)
- Signal processing for network security (CNS-SPNS)
- Optimal network signal processing and resource allocation (CNS-NSPRA)

(continued on next page)

Modeling and Analysis (MA)

- Performance and bounds of methods (MA-PBM)
- Robustness and vulnerability (MA-RV)
- Network modeling and identification (MA-NMI)
- Simulations of networked information processing systems (MA-SNIPS)
- Social learning (MA-SL)
- Bio-inspired network signal processing (MA-BNSP)
- Epidemics and diffusion in populations (MA-EDP)

Imaging and Media Applications (IMA)

- Image and video processing over networks (IMA-IVPN)
- Media cloud computing and communication (IMA-MCCC)
- Multimedia streaming and transport (IMA-MST)
- Social media computing and networking (IMA-SMCN)
- Signal processing for cyber-physical systems (IMA-SPCPS)
- Wireless/mobile multimedia (IMA-WMM)

Data Analysis (DA)

- Processing, analysis, and visualization of big data (DA-BD)
- Signal and information processing for crowd computing (DA-CC)
- Signal and information processing for the Internet of Things (DA-IOT)
- Emergence of behavior (DA-EB)

Emerging topics and applications (ETA)

- Emerging topics (ETA-ET)
- Applications in life sciences, ecology, energy, social networks, economic networks, finance, social sciences etc. smart grids, wireless health, robotics, transportation, and other areas of science and engineering (ETA-APP)

>>We are accepting paper submissions: please [submit a manuscript here](#)<<

***T-SIPN is co-sponsored by the Signal Processing, Communications and Computer societies**

IEEE/ACM 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



JUNE 2015

VOLUME 23

NUMBER 6

ITASFA

(ISSN 2329-9290)

REGULAR PAPERS

Combining Relevance Language Modeling and Clarity Measure for Extractive Speech Summarization http://dx.doi.org/10.1109/TASLP.2015.2414820	<i>S.-H. Liu, K.-Y. Chen, B. Chen, H.-M. Wang, H.-C. Yen, and W.-L. Hsu</i>	957
Elimination of Impulsive Disturbances From Stereo Audio Recordings Using Vector Autoregressive Modeling and Variable-order Kalman Filtering http://dx.doi.org/10.1109/TASLP.2015.2414823	<i>M. Niedźwiecki, M. Ciołek, and K. Cisowski</i>	970
Learning Spectral Mapping for Speech Dereverberation and Denoising http://dx.doi.org/10.1109/TASLP.2015.2416653	<i>K. Han, Y. Wang, D. Wang, W. S. Woods, I. Merks, and T. Zhang</i>	982

IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 2329-9290) 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 \$339.00 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 of \$31.00 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 © 2015 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **Postmaster:** Send address changes to IEEE/ACM 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.



Identifying Cover Songs Using Information-Theoretic Measures of Similarity http://dx.doi.org/10.1109/TASLP.2015.2416655	993
..... <i>P. Foster, S. Dixon, and A. Klapuri</i>	
Coherent-to-Diffuse Power Ratio Estimation for Dereverberation http://dx.doi.org/10.1109/TASLP.2015.2418571	1006
..... <i>A. Schwarz and W. Kellermann</i>	
Incremental Syllable-Context Phonetic Vocoding http://dx.doi.org/10.1109/TASLP.2015.2418577	1019
..... <i>M. Cernak, P. N. Garner, A. Lazaridis, P. Motlicek, and X. Na</i>	
Audio-Based Video Genre Identification http://dx.doi.org/10.1109/TASLP.2014.2387411	1031
..... <i>M. Rouvier, S. Oger, G. Linarès, D. Matrouf, B. Merialdo, and Y. Li</i>	
Generative Modeling of Voice Fundamental Frequency Contours http://dx.doi.org/10.1109/TASLP.2015.2418576	1042
..... <i>H. Kameoka, K. Yoshizato, T. Ishihara, K. Kadowaki, Y. Ohishi, and K. Kashino</i>	
Multiview Soundfield Imaging in the Projective Ray Space http://dx.doi.org/10.1109/TASLP.2015.2419076	1054
..... <i>D. Marković, F. Antonacci, A. Sarti, and S. Tubaro</i>	
Novel Sampling Scheme on the Sphere for Head-Related Transfer Function Measurements http://dx.doi.org/10.1109/TASLP.2015.2419971 ..	1068
..... <i>A. P. Bates, Z. Khalid, and R. A. Kennedy</i>	
Encoding Multiple Audio Objects Using Intra-Object Sparsity http://dx.doi.org/10.1109/TASLP.2015.2419980	1082
..... <i>M. Jia, Z. Yang, C. Bao, X. Zheng, and C. Ritz</i>	

EDICS—Editor’s Information and Classification Scheme http://dx.doi.org/10.1109/TASLP.2015.2436471	1096
Information for Authors http://dx.doi.org/10.1109/TASLP.2015.2436472	1098



The Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop



10th-13th July 2016, Rio de Janeiro, Brazil



Call for Papers

General Chairs

Rodrigo C. de Lamare,
PUC-Rio, Brazil and University of York, United Kingdom

Martin Haardt,
TU Ilmenau, Germany

Technical Chairs

Aleksandar Dogandzic,
Iowa State University, USA

Vítor Nascimento,
University of São Paulo, Brazil

Special Sessions Chair

Cédric Richard,
University of Nice, France

Publicity Chair

Maria Sabrina Greco,
University of Pisa, Italy

Important Dates

Special Session Proposals
29th January, 2016

Submission of Papers
26th February, 2016

Notification of Acceptance
29th April, 2016

Final Manuscript Submission
16th May, 2016

Advance Registration
16th May, 2016

Technical Program

The SAM Workshop is an important IEEE Signal Processing Society event dedicated to sensor array and multichannel signal processing. The organizing committee invites the international community to contribute with state-of-the-art developments in the field. SAM 2016 will feature plenary talks by leading researchers in the field as well as poster and oral sessions with presentations by the participants.

Welcome to Rio de Janeiro! – The workshop will be held at the Pontifical Catholic University of Rio de Janeiro, located in Gávea, in a superb area surrounded by beaches, mountains and the Tijuca National Forest, the world's largest urban forest. Rio de Janeiro is a world renowned city for its culture, beautiful landscapes, numerous tourist attractions and international cuisine. The workshop will take place during the first half of July about a month before the 2016 Summer Olympic Games when Rio will offer plenty of cultural activities and festivities, which will make SAM 2016 a memorable 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
- Distributed and adaptive signal processing
- Intelligent systems and knowledge-based signal processing
- Microphone and loudspeaker array applications
- MIMO radar
- Multi-antenna systems: multiuser MIMO, massive MIMO and space-time coding
- Multi-channel imaging and hyperspectral processing
- 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
- Source Localization, Classification and Tracking
- Synthetic aperture techniques
- Space-time adaptive processing
- Statistical modelling for sensor arrays
- Waveform diverse sensors and systems

Submission of papers – Full-length four-page papers will be accepted only electronically.

Special session proposals – They should be submitted by e-mail to the Technical Program Chairs and the Special Sessions Chair and include a topical title, rationale, session outline, contact information, and list of invited speakers.

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



JUNE 2015

VOLUME 24

NUMBER 6

IIPRE4

(ISSN 1057-7149)

PAPERS

Scalable Mobile Image Retrieval by Exploring Contextual Saliency http://dx.doi.org/10.1109/TIP.2015.2411433	1709
..... <i>X. Yang, X. Qian, and Y. Xue</i>	
Undersampled Face Recognition via Robust Auxiliary Dictionary Learning http://dx.doi.org/10.1109/TIP.2015.2409738	1722
..... <i>C.-P. Wei and Y.-C. F. Wang</i>	
Multiscale Logarithm Difference Edgemaps for Face Recognition Against Varying Lighting Conditions http://dx.doi.org/10.1109/TIP.2015.2409988	1735
..... <i>Z.-R. Lai, D.-Q. Dai, C.-X. Ren, and K.-K. Huang</i>	
Randomized Spatial Context for Object Search http://dx.doi.org/10.1109/TIP.2015.2405337	1748
..... <i>Y. Jiang, J. Meng, J. Yuan, and J. Luo</i>	
Structured Sparse Priors for Image Classification http://dx.doi.org/10.1109/TIP.2015.2409572	1763
..... <i>U. Srinivas, Y. Suo, M. Dao, V. Monga, and T. D. Tran</i>	
Adaptive Regularization With the Structure Tensor http://dx.doi.org/10.1109/TIP.2015.2409562	1777
..... <i>V. Estellers, S. Soatto, and X. Bresson</i>	
Generation of Spatial Orders and Space-Filling Curves http://dx.doi.org/10.1109/TIP.2015.2409571	1791
..... <i>G. Schrack and L. Stocco</i>	
Head Pose Estimation From a 2D Face Image Using 3D Face Morphing With Depth Parameters http://dx.doi.org/10.1109/TIP.2015.2405483	1801
..... <i>S. G. Kong and R. O. Mbouna</i>	



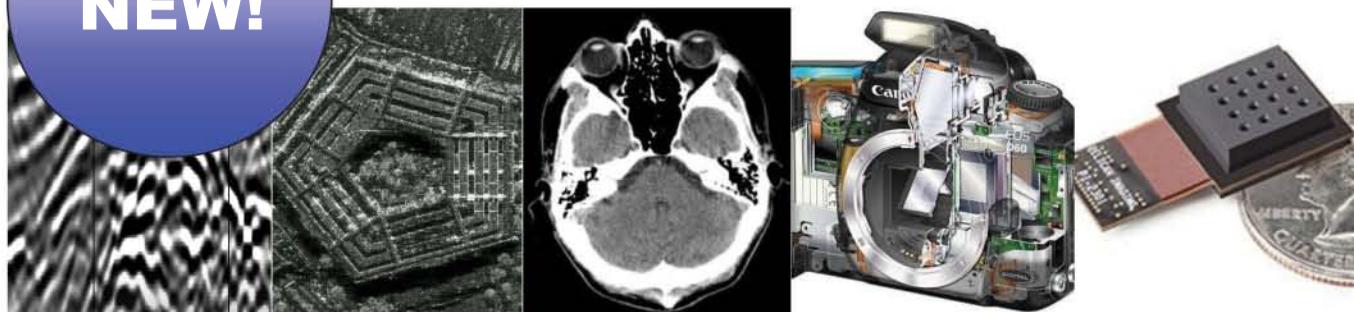
Exemplar-Based Inpainting: Technical Review and New Heuristics for Better Geometric Reconstructions http://dx.doi.org/10.1109/TIP.2015.2411437	<i>P. Buysseens, M. Daisy, D. Tschumperlé, and O. Lézoray</i>	1809
A Low-Rank Approximation-Based Transductive Support Tensor Machine for Semisupervised Classification http://dx.doi.org/10.1109/TIP.2015.2403235	<i>X. Liu, T. Guo, L. He, and X. Yang</i>	1825
Hashing on Nonlinear Manifolds http://dx.doi.org/10.1109/TIP.2015.2405340	<i>F. Shen, C. Shen, Q. Shi, A. van den Hengel, Z. Tang, and H. T. Shen</i>	1839
Anisotropic Spectral-Spatial Total Variation Model for Multispectral Remote Sensing Image Destriping http://dx.doi.org/10.1109/TIP.2015.2404782	<i>Y. Chang, L. Yan, H. Fang, and C. Luo</i>	1852
Event Oriented Dictionary Learning for Complex Event Detection http://dx.doi.org/10.1109/TIP.2015.2413294	<i>Y. Yan, Y. Yang, D. Meng, G. Liu, W. Tong, A. G. Hauptmann, and N. Sebe</i>	1867
Utilizing Image Scales Towards Totally Training Free Blind Image Quality Assessment http://dx.doi.org/10.1109/TIP.2015.2411436	<i>A. Saha and Q. M. J. Wu</i>	1879
A Chi-Squared-Transformed Subspace of LBP Histogram for Visual Recognition http://dx.doi.org/10.1109/TIP.2015.2409554	<i>J. Ren, X. Jiang, and J. Yuan</i>	1893
Learning Person-Person Interaction in Collective Activity Recognition http://dx.doi.org/10.1109/TIP.2015.2409564	<i>X. Chang, W.-S. Zheng, and J. Zhang</i>	1905
Recursive Approximation of the Bilateral Filter http://dx.doi.org/10.1109/TIP.2015.2403238	<i>Q. Yang</i>	1919
Combining Facial Dynamics With Appearance for Age Estimation http://dx.doi.org/10.1109/TIP.2015.2412377	<i>H. Dibeklioglu, F. Alnajar, A. A. Salah, and T. Gevers</i>	1928
Estimation of Illuminants From Projections on the Planckian Locus http://dx.doi.org/10.1109/TIP.2015.2405414	<i>B. Mazin, J. Delon, and Y. Gousseau</i>	1944
Double Nuclear Norm-Based Matrix Decomposition for Occluded Image Recovery and Background Modeling http://dx.doi.org/10.1109/TIP.2015.2400213	<i>F. Zhang, J. Yang, Y. Tai, and J. Tang</i>	1956
Image Denoising by Exploring External and Internal Correlations http://dx.doi.org/10.1109/TIP.2015.2412373	<i>H. Yue, X. Sun, J. Yang, and F. Wu</i>	1967
Depth Reconstruction From Sparse Samples: Representation, Algorithm, and Sampling http://dx.doi.org/10.1109/TIP.2015.2409551	<i>L.-K. Liu, S. H. Chan, and T. Q. Nguyen</i>	1983
EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIP.2015.2425519		1997
Information for Authors http://dx.doi.org/10.1109/TIP.2015.2425518		1998

CALLS FOR PAPERS

Call for Papers IEEE J-STSP Special Issue on Structured Matrices in Signal and Data Processing http://dx.doi.org/10.1109/TIP.2015.2417692		2000
---	--	------



IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING



The new IEEE Transactions on Computational Imaging seeks original manuscripts for publication. This new journal will publish research results where computation plays an integral role in the image formation process. All areas of computational imaging are appropriate, ranging from the principles and theory of computational imaging, to modeling paradigms for computational imaging, to image formation methods, to the latest innovative computational imaging system designs. Topics of interest include, but are not limited to the following:

<p>Imaging Models and Representation</p> <ul style="list-style-type: none"> • Statistical-model based methods • System and image prior models • Noise models • Graphical and tree-based models • Perceptual models 	<p>Computational Photography</p> <ul style="list-style-type: none"> • Non-classical image capture, Generalized illumination • Time-of-flight imaging • High dynamic range imaging • Focal stacks 	<p>Tomographic Imaging</p> <ul style="list-style-type: none"> • X-ray CT • PET • SPECT
<p>Computational Sensing</p> <ul style="list-style-type: none"> • Coded source methods • Structured light • Coded aperture methods • Compressed sensing • Light-field sensing • Plenoptic imaging • Hardware and software systems 	<p>Computational Consumer Imaging</p> <ul style="list-style-type: none"> • Cell phone imaging • Camera-array systems • Depth cameras 	<p>Magnetic Resonance Imaging</p> <ul style="list-style-type: none"> • Diffusion tensor imaging • Fast acquisition
<p>Computational Image Creation</p> <ul style="list-style-type: none"> • Sparsity-based methods • Statistically-based inversion methods, Bayesian regularization • Super-resolution, multi-image fusion • Learning-based methods, Dictionary-based methods • Optimization-based methods; proximal iterative methods, ADMM 	<p>Computational Acoustic Imaging</p> <ul style="list-style-type: none"> • Multi-static ultrasound imaging • Photo-acoustic imaging • Acoustic tomography 	<p>Radar Imaging</p> <ul style="list-style-type: none"> • Synthetic aperture imaging • Inverse synthetic imaging • Terahertz imaging
	<p>Computational Microscopic Imaging</p> <ul style="list-style-type: none"> • Holographic microscopy • Quantitative phase imaging • Multi-illumination microscopy • Lensless microscopy 	<p>Geophysical Imaging</p> <ul style="list-style-type: none"> • Multi-spectral imaging • Ground penetrating radar • Seismic tomography
		<p>Multi-spectral Imaging</p> <ul style="list-style-type: none"> • Multi-spectral imaging • Hyper-spectral imaging • Spectroscopic imaging

Editor-in-Chief: W. Clem Karl, Boston University.

To submit a paper go to: <https://mc.manuscriptcentral.com/tci-ieee>



**General Chair**

Lina Karam
Arizona State University

General Co-Chair

Aggelos Katsaggelos
Northwestern University

Technical Program Chairs

Fernando Pereira
Instituto Superior Técnico
Gaurav Sharma
University of Rochester

Innovation Program Chairs

Haohong Wang
TCL Research America

Jeff Bier
BDTI & Embedded Vision Alliance

Finance Chair

Sohail Dianat
Rochester Institute of Technology

Plenary Chairs

Michael Marcellin
University of Arizona
Sethuraman Panchanathan
Arizona State University

Special Sessions Chairs

Dinei Florencio
Microsoft Research
Chaker Larabi
Poitiers University
Zhou Wang
University of Waterloo

Tutorials Chairs

Ghassan AlRegib
Georgia Tech
Rony Ferzli
Intel

Publicity Chair

Michel Sarkis
Qualcomm Technologies Inc.

Awards Chairs

Vivek Goyal
Boston University
Ivana Tosic
Ricoh Innovations

Exhibits Chair

David Frakes
Arizona State University &
Google

Publication Chairs

Patrick Le Callet
Nantes University
Baixin Li
Arizona State University

Local Arrangement Chairs

Jorge Caviedes
Intel
Pavan Turaga
Arizona State University

Registration Chair

Ricardo De Queiroz
Universidade de Brasilia

Conference Management

Conference Management Services

The 23rd IEEE International Conference on Image Processing (ICIP) will be held in the Phoenix Convention Centre, Phoenix, Arizona, USA, on September 25 - 28, 2016. ICIP is the world's largest and most comprehensive technical conference focused on image and video processing and computer vision. In addition to the Technical Program, ICIP 2016 will feature an **Innovation Program** focused on **innovative vision technologies and fostering innovation, entrepreneurship, and networking**. The conference will feature world-class speakers, tutorials, exhibits, and a vision technology showcase.

Topics in the ICIP 2016 Technical Program include but are not limited to the following:

<i>Filtering, Transforms, Multi-Resolution Processing</i>	<i>Biological and Perceptual-based Processing</i>
<i>Restoration, Enhancement, Super-Resolution</i>	<i>Visual Quality Assessment</i>
<i>Computer Vision Algorithms and Technologies</i>	<i>Scanning, Display, and Printing</i>
<i>Compression, Transmission, Storage, Retrieval</i>	<i>Document and Synthetic Visual Processing</i>
<i>Computational Imaging</i>	<i>Applications to various fields (e.g., biomedical,</i>
<i>Color and Multispectral Processing</i>	<i>Advanced Driving Assist Systems, assistive</i>
<i>Multi-View and Stereoscopic Processing</i>	<i>living, security, learning,</i>
<i>Multi-Temporal and Spatio-Temporal Processing</i>	<i>health and environmental monitoring,</i>
<i>Video Processing and Analytics</i>	<i>manufacturing, consumer electronics)</i>
<i>Authentication and Biometrics</i>	

The ICIP 2016 innovation program will feature a **vision technology showcase** of state-of-the-art vision technologies, innovation challenges, talks by innovation leaders and entrepreneurs, tutorials, and networking.

Paper Submission: Prospective authors are invited to submit full-length papers at the conference website, with up to four pages for technical content including figures and references, and with one additional optional 5th page for references only. Submission instructions, templates for the required paper format, and information on "no show" policy are available at www.icip2016.com.

Tutorials and Special Sessions Proposals: Tutorials will be held on September 25, 2016. Tutorial proposals should be submitted to tutorials@icip2016.com and must include title, outline, contact information, biography and selected publications for the presenter(s), and a description of the tutorial and material to be distributed to participants. Special Sessions proposals should be submitted to specialsessions@icip2016.com and must include a topical title, rationale, session outline, contact information, and a list of invited papers. For detailed submission guidelines, please refer the ICIP 2016 website at www.icip2016.com.

Important Deadlines:

Special Session and Tutorial Proposals: November 16, 2015
 Notification of Special Session and Tutorial Acceptance: December 18, 2015
 Paper Submissions: January 25, 2016
 Notification of Paper Acceptance: April 30, 2016
 Visual Technology Innovator Award Nomination: March 30, 2016
 Revised Paper Upload Deadline: May 30, 2016
 Authors' Registration Deadline: May 30, 2016



<http://www.facebook.com/icip2016>



<https://twitter.com/icip2016/>



<https://www.linkedin.com/groups/ICIP-2016-6940658>



IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

MAY 2015

VOLUME 10

NUMBER 5

ITIFA6

(ISSN 1556-6013)

PAPERS

- Randomized Component and Its Application to (t,m,n)-Group Oriented Secret Sharing <http://dx.doi.org/10.1109/TIFS.2014.2384393> 889
 *M. Fuyou, X. Yan, W. Xingfu, and M. Badawy*
- Hill-Climbing Attacks on Multibiometrics Recognition Systems <http://dx.doi.org/10.1109/TIFS.2014.2384735> 900
 *E. Maiorana, G. E. Hine, and P. Campisi*
- On Criminal Identification in Color Skin Images Using Skin Marks (RPPVSM) and Fusion With Inferred Vein Patterns
<http://dx.doi.org/10.1109/TIFS.2014.2387575> *A. Nurhudatiana and A. W.-K. Kong* 916
- An Energy-Ratio-Based Approach for Detecting Pilot Spoofing Attack in Multiple-Antenna Systems
<http://dx.doi.org/10.1109/TIFS.2015.2392564> *Q. Xiong, Y.-C. Liang, K. H. Li, and Y. Gong* 932
- On the Error Region for Channel Estimation-Based Physical Layer Authentication Over Rayleigh Fading
<http://dx.doi.org/10.1109/TIFS.2015.2392565> *A. Ferrante, N. Laurenti, C. Masiero, M. Pavon, and S. Tomasin* 941
- An Image Recapture Detection Algorithm Based on Learning Dictionaries of Edge Profiles <http://dx.doi.org/10.1109/TIFS.2015.2392566> .. 953
 *T. Thongkamwitoon, H. Muammar, and P.-L. Dragotti*
- Real-World and Rapid Face Recognition Toward Pose and Expression Variations via Feature Library Matrix
<http://dx.doi.org/10.1109/TIFS.2015.2393553> *A. Moeini and H. Moeini* 969
- Security Considerations in Minutiae-Based Fuzzy Vaults <http://dx.doi.org/10.1109/TIFS.2015.2392559> 985
 *B. Tams, P. Mihăilescu, and A. Munk*
- Blurred Image Splicing Localization by Exposing Blur Type Inconsistency <http://dx.doi.org/10.1109/TIFS.2015.2394231> 999
 *K. Bahrami, A. C. Kot, L. Li, and H. Li*
- Joint Substation-Transmission Line Vulnerability Assessment Against the Smart Grid <http://dx.doi.org/10.1109/TIFS.2015.2394240> 1010
 *Y. Zhu, J. Yan, Y. Tang, Y. Sun, and H. He*
- Using Visual Rhythms for Detecting Video-Based Facial Spoof Attacks <http://dx.doi.org/10.1109/TIFS.2015.2395139> 1025
 *A. Pinto, W. R. Schwartz, H. Pedrini, and A. de Rezende Rocha*
- Visualizing and Modeling the Scanning Behavior of the Conficker Botnet in the Presence of User and Network Activity
<http://dx.doi.org/10.1109/TIFS.2015.2396478> *R. Weaver* 1039
- A Hybrid Scheme of Public-Key Encryption and Somewhat Homomorphic Encryption <http://dx.doi.org/10.1109/TIFS.2015.2398359> 1052
 *J. H. Cheon and J. Kim*
- One-Time Spectrum Coexistence in Dynamic Spectrum Access When the Secondary User May Be Malicious
<http://dx.doi.org/10.1109/TIFS.2015.2398360> *A. Garnaev and W. Trappe* 1064
- Median Filtered Image Quality Enhancement and Anti-Forensics via Variational Deconvolution
<http://dx.doi.org/10.1109/TIFS.2015.2398362> *W. Fan, K. Wang, F. Cayre, and Z. Xiong* 1076
- Privacy-Preserving Detection of Sensitive Data Exposure <http://dx.doi.org/10.1109/TIFS.2015.2398363> *X. Shu, D. Yao, and E. Bertino* 1092

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

JUNE 2015

VOLUME 10

NUMBER 6

ITIFA6

(ISSN 1556-6013)

PAPERS

Short Dynamic Group Signature Scheme Supporting Controllable Linkability http://dx.doi.org/10.1109/TIFS.2015.2390497	1109
..... J. Y. Hwang, L. Chen, H. S. Cho, and D. Nyang	
ENF-Based Region-of-Recording Identification for Media Signals http://dx.doi.org/10.1109/TIFS.2015.2398367	1125
..... A. Hajj-Ahmad, R. Garg, and M. Wu	
A PUF-FSM Binding Scheme for FPGA IP Protection and Pay-Per-Device Licensing http://dx.doi.org/10.1109/TIFS.2015.2400413	1137
..... J. Zhang, Y. Lin, Y. Lyu, and G. Qu	
On the Secrecy Rate Region of Multiple-Access Wiretap Channel With Noncausal Side Information http://dx.doi.org/10.1109/TIFS.2015.2400416	1151
..... A. Sonee and G. A. Hodtani	
Enabling Cloud Storage Auditing With Key-Exposure Resistance http://dx.doi.org/10.1109/TIFS.2015.2400425	1167
..... J. Yu, K. Ren, C. Wang, and V. Varadharajan	
Authorized and Rogue Device Discrimination Using Dimensionally Reduced RF-DNA Fingerprints http://dx.doi.org/10.1109/TIFS.2015.2400426	1180
..... D. R. Reising, M. A. Temple, and J. A. Jackson	
A New Biocryptosystem-Oriented Security Analysis Framework and Implementation of Multibiometric Cryptosystems Based on Decision Level Fusion http://dx.doi.org/10.1109/TIFS.2015.2402593	1193
..... C. Li, J. Hu, J. Pieprzyk, and W. Susilo	
Strong, Neutral, or Weak: Exploring the Impostor Score Distribution http://dx.doi.org/10.1109/TIFS.2015.2403136	1207
..... A. Sgroi, P. J. Flynn, K. Bowyer, and P. J. Phillips	
Tally-Based Simple Decoders for Traitor Tracing and Group Testing http://dx.doi.org/10.1109/TIFS.2015.2403575	1221
..... B. Škorić	
The CEO Problem With Secrecy Constraints http://dx.doi.org/10.1109/TIFS.2015.2404134	1234
..... F. Naghibi, S. Salimi, and M. Skoglund	
On the Optimality of Keyless Authentication in a Noisy Model http://dx.doi.org/10.1109/TIFS.2015.2405891	1250
..... S. Jiang	
Weak Secrecy in the Multiway Untrusted Relay Channel With Compute-and-Forward http://dx.doi.org/10.1109/TIFS.2015.2405903	1262
..... J. Richter, C. Scheunert, S. Engelmann, and E. A. Jorswieck	
White-Box Traceable Ciphertext-Policy Attribute-Based Encryption Supporting Flexible Attributes http://dx.doi.org/10.1109/TIFS.2015.2405905	1274
..... J. Ning, X. Dong, Z. Cao, L. Wei, and X. Lin	
Ranking Radically Influential Web Forum Users http://dx.doi.org/10.1109/TIFS.2015.2407313	1289
..... T. Anwar and M. Abulaish	
Secure and Privacy Preserving Protocol for Cloud-Based Vehicular DTNs http://dx.doi.org/10.1109/TIFS.2015.2407326	1299
..... J. Zhou, X. Dong, Z. Cao, and A. V. Vasilakos	

COMMENTS AND CORRECTIONS

Comments on "DAC-MACS: Effective Data Access Control for Multiauthority Cloud Storage Systems"/Security Analysis of Attribute Revocation in Multiauthority Data Access Control for Cloud Storage Systems http://dx.doi.org/10.1109/TIFS.2015.2407327 ..	1315
..... J. Hong, K. Xue, and W. Li	
EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIFS.2015.2431485	1318
Information for Authors http://dx.doi.org/10.1109/TIFS.2015.2431484	1319

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/>

MAY 2015

VOLUME 17

NUMBER 5

ITMUF8

(ISSN 1520-9210)

PAPERS

Audio/Video Analysis and Synthesis

Spatio-Temporally Consistent Color and Structure Optimization for Multiview Video Color Correction
<http://dx.doi.org/10.1109/TMM.2015.2412879> S.-P. Lu, B. Ceulemans, A. Munteanu, and P. Schelkens 577

Distributed and Parallel Processing

Characterization of SURF and BRISK Interest Point Distribution for Distributed Feature Extraction in Visual Sensor Networks
<http://dx.doi.org/10.1109/TMM.2015.2406574> G. Dán, M. A. Khan, and V. Fodor 591

Multimodal Human-Human and Human-Computer Dialog

TCD-TIMIT: An Audio-Visual Corpus of Continuous Speech
<http://dx.doi.org/10.1109/TMM.2015.2407694> N. Harte and E. Gillen 603

Content Description and Annotation

PixNet: A Localized Feature Representation for Classification and Visual Search
<http://dx.doi.org/10.1109/TMM.2015.2410734> N. Pourian and B. S. Manjunath 616

Efficient Mining of Optimal AND/OR Patterns for Visual Recognition
<http://dx.doi.org/10.1109/TMM.2015.2414720> C. Weng and J. Yuan 626

Multimedia Search and Retrieval

Fine-Grained Image Search
<http://dx.doi.org/10.1109/TMM.2015.2408566> L. Xie, J. Wang, B. Zhang, and Q. Tian 636

Fast Image Retrieval: Query Pruning and Early Termination
<http://dx.doi.org/10.1109/TMM.2015.2408563> L. Zheng, S. Wang, Z. Liu, and Q. Tian 648

Structured Visual Feature Learning for Classification via Supervised Probabilistic Tensor Factorization
<http://dx.doi.org/10.1109/TMM.2015.2410135> X. Tan, F. Wu, X. Li, S. Tang, W. Lu, and Y. Zhuang 660

Global-Scale Location Prediction for Social Images Using Geo-Visual Ranking
<http://dx.doi.org/10.1109/TMM.2015.2413351> X. Li, M. Larson, and A. Hanjalic 674



<i>Realtime Communication and Video Conferencing</i>		
Rate and Power Allocation for Joint Coding and Transmission in Wireless Video Chat Applications http://dx.doi.org/10.1109/TMM.2015.2413354	<i>S.-P. Chuah, Y.-P. Tan, and Z. Chen</i>	687
<i>Web and Internet</i>		
Semantic-Improved Color Imaging Applications: It Is All About Context http://dx.doi.org/10.1109/TMM.2015.2410175	<i>A. Lindner and S. Süsstrunk</i>	700
<i>Media Transmission and Communication Standardization</i>		
A Novel Traffic Rate Measurement Algorithm for Quality of Experience-Aware Video Admission Control http://dx.doi.org/10.1109/TMM.2015.2416637	<i>Q. M. Qadir, A. A. Kist, and Z. Zhang</i>	711
<i>Media Cloud Computing and Communication</i>		
Bucket-Filling: An Asymptotically Optimal Video-on-Demand Network With Source Coding http://dx.doi.org/10.1109/TMM.2015.2416636	<i>Z. Chang and S.-H. G. Chan</i>	723
<i>Wireless/Mobile Multimedia</i>		
Joint Time-Domain Resource Partitioning, Rate Allocation, and Video Quality Adaptation in Heterogeneous Cellular Networks http://dx.doi.org/10.1109/TMM.2015.2408254	<i>A. Argyriou, D. Kosmanos, and L. Tassiulas</i>	736
<i>Multimodal Human Behavior</i>		
Automatic Recognition of Emergent Social Roles in Small Group Interactions http://dx.doi.org/10.1109/TMM.2015.2408437	<i>A. Sapru and H. Bourlard</i>	746
<hr/>		
Information for Authors http://dx.doi.org/10.1109/TMM.2015.2422256		761
<hr/>		
CALLS FOR PAPERS		
IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING http://dx.doi.org/10.1109/TMM.2015.2422257		763
IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS http://dx.doi.org/10.1109/TMM.2015.2422258		764
<hr/>		

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING


www.ieee.org/sp/index.html

APRIL 2015

VOLUME 9

NUMBER 3

IJSTGY

(ISSN 1932-4553)

ISSUE ON INTERACTIVE MEDIA PROCESSING FOR IMMERSIVE COMMUNICATION

EDITORIAL

Introduction to the Issue on Interactive Media Processing for Immersive Communication <http://dx.doi.org/10.1109/IJSTSP.2015.2406092>
 *G. Cheung, D. Florencio, P. Callet, C.-W. Lin, and E. Magli* 381

PAPERS

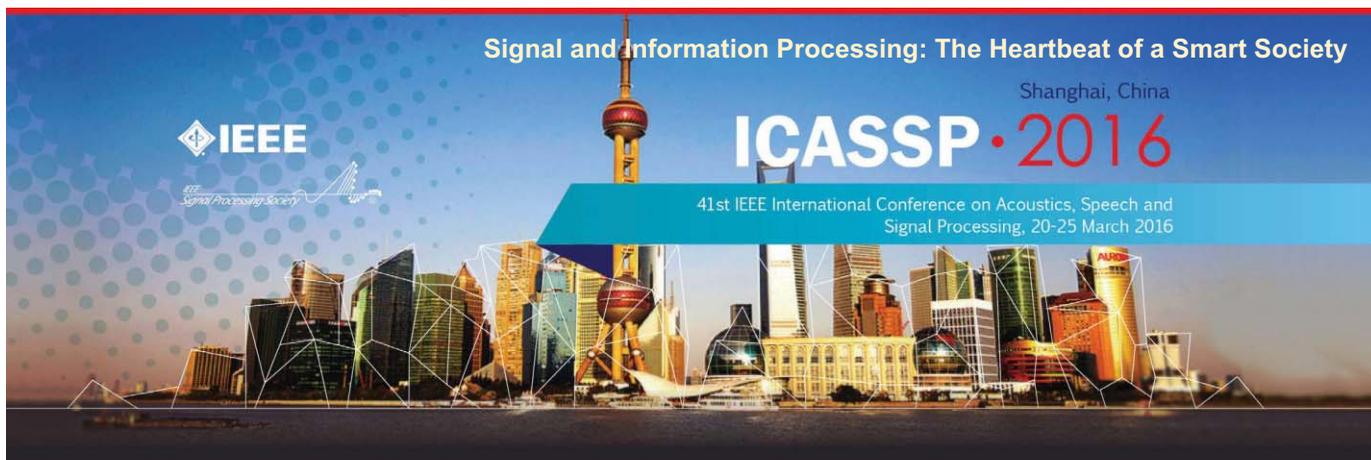
Single Shot Dual-Frequency Structured Light Based Depth Sensing <http://dx.doi.org/10.1109/IJSTSP.2015.2403794>
 *F. Li, S. Gao, G. Shi, Q. Li, L. Yang, R. Li, and X. Xie* 384

Accurate Dynamic 3D Sensing With Fourier-Assisted Phase Shifting <http://dx.doi.org/10.1109/IJSTSP.2014.2378217>
 *P. Cong, Z. Xiong, Y. Zhang, S. Zhao, and F. Wu* 396

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (ISSN 1932-4553) is published eight times a year in February, March, April, June, August, September, October, and December 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 \$371.00 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 of \$31.00 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 © 2015 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: PitneyBowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ONM5W3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



Fast, Automated, Scalable Generation of Textured 3D Models of Indoor Environments http://dx.doi.org/10.1109/JSTSP.2014.2381153	409
..... <i>E. Turner, P. Cheng, and A. Zakhor</i>	
Large-Scale 3D Point Cloud Compression Using Adaptive Radial Distance Prediction in Hybrid Coordinate Domains http://dx.doi.org/10.1109/JSTSP.2014.2370752	422
..... <i>J.-K. Ahn, K.-Y. Lee, J.-Y. Sim, and C.-S. Kim</i>	
Probabilistic Multiview Depth Image Enhancement Using Variational Inference http://dx.doi.org/10.1109/JSTSP.2014.2373331	435
..... <i>P. K. Rana, J. Taghia, Z. Ma, and M. Flierl</i>	
Real-Time and Temporal-Coherent Foreground Extraction With Commodity RGBD Camera http://dx.doi.org/10.1109/JSTSP.2014.2382476 ..	449
..... <i>M. Zhao, C.-W. Fu, J. Cai, and T.-J. Cham</i>	
Perceptual and Bitrate-Scalable Coding of Haptic Surface Texture Signals http://dx.doi.org/10.1109/JSTSP.2014.2374574	462
..... <i>R. Chaudhari, C. Schuwerk, M. Danaei, and E. Steinbach</i>	
View-Popularity-Driven Joint Source and Channel Coding of View and Rate Scalable Multi-View Video http://dx.doi.org/10.1109/JSTSP.2015.2402633	474
..... <i>J. Chakareski, V. Velisavljevic, and V. Stankovic</i>	
Optimizing Multiview Video Plus Depth Prediction Structures for Interactive Multiview Video Streaming http://dx.doi.org/10.1109/JSTSP.2015.2407320	487
..... <i>A. De Abreu, P. Frossard, and F. Pereira</i>	
Streaming Codes With Partial Recovery Over Channels With Burst and Isolated Erasures http://dx.doi.org/10.1109/JSTSP.2014.2388191	501
..... <i>A. Badr, A. Khisti, W. Tan, and J. Apostolopoulos</i>	
Cloud Mobile 3D Display Gaming User Experience Modeling and Optimization by Asymmetric Graphics Rendering http://dx.doi.org/10.1109/JSTSP.2015.2396475	517
..... <i>Y. Lu, Y. Liu, and S. D. De</i>	
3D Perception Based Quality Pooling: Stereopsis, Binocular Rivalry, and Binocular Suppression http://dx.doi.org/10.1109/JSTSP.2015.2393296	533
..... <i>K. Lee and S. Lee</i>	
ITEM: Immersive Telepresence for Entertainment and Meetings—A Practical Approach http://dx.doi.org/10.1109/JSTSP.2014.2375819	546
..... <i>V. A. Nguyen, J. Lu, S. Zhao, D. T. Vu, H. Yang, D. L. Jones, and M. N. Do</i>	
<i>withyou</i> —An Experimental End-to-End Telepresence System Using Video-Based Reconstruction http://dx.doi.org/10.1109/JSTSP.2015.2402635	562
..... <i>D. J. Roberts, A. J. Fairchild, S. P. Campion, J. O'Hare, C. M. Moore, R. Aspin, T. Duckworth, P. Gasparello, and F. Tecchia</i>	



General Chairs

Zhi Ding, Univ. of California, Davis, USA
 Zhi-Quan Luo, Univ. of Minnesota, USA
 Wenjun Zhang, Shanghai Jiao Tong Univ., China

Technical Program Chairs

P. C. Ching, Chinese Univ. of Hong Kong, Hong Kong
 Dominic K.C. Ho, Univ. of Missouri, USA

Finance Chairs

Shuguang Cui, Texas A&M Univ., USA
 Rong Xie, Shanghai Jiao Tong Univ., China

Plenaries Chairs

Zhi-Pei Liang, UIUC, USA
 Björn Ottersten, Univ. of Luxembourg, Luxembourg

Special Sessions Chairs

Tim Davidson, McMaster Univ., Canada
 Jianguo Huang, Northwestern Polytech. Univ., China

Tutorials Chairs

Jian Li, Univ. of Florida, USA
 Jose Principe, Univ. of Florida, USA

Student Session Chair

Wei Zhang, Univ. of New South Wales, Australia

Registration Chairs

Tongtong Li, Michigan State Univ., USA
 Xiaojun Yuan, ShanghaiTech Univ., China

Publicity Chairs

Xiaokang Yang, Shanghai Jiao Tong Univ., China
 Mounir Ghogho, Leeds Univ., UK
 Ignacio Santamaria, Univ. of Cantabria, Spain

Publication Chairs

Min Dong, Univ. of Ontario Inst. of Tech., Canada
 Thomas Fang Zheng, Tsinghua Univ., China

Industrial & Exhibit Chairs

Li Deng, Microsoft, USA
 Jinyu Li, Microsoft, USA
 Cathy Wicks, Texas Instruments, USA

Local Arrangement Chairs

Ning Liu, Shanghai Jiao Tong Univ., China
 Meixia Tao, Shanghai Jiao Tong Univ., China

Webmaster

Yi Xu, Shanghai Jiao Tong Univ., China

Workshop Chairs

Jianguo Huang, Northwestern Polytech. Univ., China
 Jiwu Huang, Sun Yat-sen Univ., China

ICASSP2016: Signal and information processing is the driving heartbeat in the development of technologies that enrich our lives and advance our society. The 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP) will be held in the Shanghai International Convention Center, Shanghai, China between March 20 and 25, 2016. The conference provides, both for researchers and developers, an engaging forum to exchange ideas and propel new developments in this field. The 2016 conference will showcase world-class presentations by internationally renowned speakers and will facilitate a fantastic opportunity to network with like-minded professionals from around the world. Topics include but are not limited to:

- Audio and acoustic signal processing
- Bio-imaging and biomedical signal processing
- Signal processing education
- Speech processing
- Industry technology tracks
- Information forensics and security
- Machine learning for signal processing
- Signal processing for Big Data
- Multimedia signal processing
- Sensor array & multichannel signal processing
- Design & implementation of signal processing systems
- Signal processing for communications & networking
- Image, video & multidimensional signal processing
- Signal processing theory & methods
- Spoken language processing
- Signal processing for the Internet of Things

Shanghai: Shanghai is the most populous city in China and one of the most populous cities in the world. A global city, Shanghai exerts influence over global commerce, finance, culture, art, fashion, research and entertainment. The city is located in the middle portion of the Chinese coast, and sits at the mouth of the Yangtze River. The city is a tourist destination renowned for its historical landmarks, such as the Bund and City God Temple, and its modern and ever-expanding Pudong skyline including the Oriental Pearl Tower. Today, Shanghai is the largest center of commerce and finance in mainland China, and has been described as the "showpiece" of the world's fastest-growing major economy.

Submission of Papers: Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possible references, and with one additional optional 5th page containing only references. A selection of best student papers will be made by the ICASSP 2016 committee upon recommendations from the Technical Committees.

Tutorial and Special Session Proposals: Tutorials will be held on March 20 and 21, 2016. Tutorial proposals must include title, outline, contact information, biography and selected publications for the presenter(s), and a description of the tutorial and the material to be distributed to participants. Special session proposals must include a topical title, rationale, session outline, contact information, and a list of invited speakers. Additional information can be found at the ICASSP 2016 website.

Signal Processing Letters: Authors of IEEE Signal Processing Letters (SPL) papers will be given the opportunity to present their work at ICASSP 2016, subject to space availability and approval by the ICASSP Technical Program Chairs. SPL papers published between January 1, 2015 and December 31, 2015 are eligible for presentation at ICASSP 2016.

Show and Tell: S&T offers a perfect stage to showcase innovative ideas in all technical areas of interest at ICASSP. S&T sessions contain demos that are highly interactive and visible. Please refer to the ICASSP 2016 website for additional information regarding demo submission.

Important Deadlines:

Special session & tutorial proposals	August 3, 2015
Notification of special session & tutorial acceptance	September 11, 2015
Submission of regular papers	September 25, 2015
Signal processing letters	December 16, 2015
Notification of paper acceptance	December 21, 2015
Revised paper upload	January 22, 2016
Author registration	January 22, 2016



IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

JULY 2015

VOLUME 22

NUMBER 7

ISPLEM

(ISSN 1070-9908)

LETTERS

Higher-Order-Statistics-Based Fractal Dimension for Noisy Bowel Sound Detection http://dx.doi.org/10.1109/LSP.2014.2369856	789
..... <i>M.-J. Sheu, P.-Y. Lin, J.-Y. Chen, C.-C. Lee, and B.-S. Lin</i>	
Visual Tracking via Structure Constrained Grouping http://dx.doi.org/10.1109/LSP.2014.2369476	794
..... <i>L. Wang, H. Lu, and D. Wang</i>	
Adaptive Array Thinning for Enhanced DOA Estimation http://dx.doi.org/10.1109/LSP.2014.2370632	799
..... <i>X. Wang, E. Aboutanios, and M. G. Amin</i>	
Feasible Point Pursuit and Successive Approximation of Non-Convex QCQPs http://dx.doi.org/10.1109/LSP.2014.2370033	804
..... <i>O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, and N. D. Sidiropoulos</i>	
Gradient Radial Basis Function Based Varying-Coefficient Autoregressive Model for Nonlinear and Nonstationary Time Series http://dx.doi.org/10.1109/LSP.2014.2369415	809
..... <i>M. Gan, C. L. P. Chen, H.-X. Li, and L. Chen</i>	
Performance of Transmit Beamforming Codebooks with Separate Amplitude and Phase Quantization http://dx.doi.org/10.1109/LSP.2014.2370762	813
..... <i>A. Dowhuszko and J. Hämmäläinen</i>	
On the Behavior of EMD and MEMD in Presence of Symmetric α -Stable Noise http://dx.doi.org/10.1109/LSP.2014.2371132	818
..... <i>A. Komaty, A. Boudraa, J. P. Nolan, and D. Dare</i>	
Joint Power Splitting and Antenna Selection in Energy Harvesting Relay Channels http://dx.doi.org/10.1109/LSP.2014.2369748	823
..... <i>Z. Zhou, M. Peng, Z. Zhao, and Y. Li</i>	
Colocated MIMO Radar Transmit Beamspace Design for Randomly Present Target Detection http://dx.doi.org/10.1109/LSP.2014.2371241 ..	828
..... <i>H. Xu, J. Wang, J. Yuan, and X. Shan</i>	
Non-Local Means Image Denoising With a Soft Threshold http://dx.doi.org/10.1109/LSP.2014.2371332	833
..... <i>L. Lu, W. Jin, and X. Wang</i>	
No-Reference Quality Assessment of Contrast-Distorted Images Based on Natural Scene Statistics http://dx.doi.org/10.1109/LSP.2014.2372333	838
..... <i>Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai</i>	
Two Efficient Algorithms for Approximately Orthogonal Nonnegative Matrix Factorization http://dx.doi.org/10.1109/LSP.2014.2371895	843
..... <i>B. Li, G. Zhou, and A. Cichocki</i>	

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 \$309.00 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 © 2015 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage 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.

Bi-DCT: DCT-based Local Binary Descriptor for Dense Stereo Matching http://dx.doi.org/10.1109/LSP.2014.2369211	847
..... <i>S. Kim, K. Paeng, J.-W. Seo, and S. D. Kim</i>	
Long Term Relevance Feedback: A Probabilistic Axis Re-Weighting Update Scheme http://dx.doi.org/10.1109/LSP.2014.2372073	852
..... <i>A. Lakshmi, M. Nema, and S. Rakshit</i>	
Robust One-Bit Bayesian Compressed Sensing with Sign-Flip Errors http://dx.doi.org/10.1109/LSP.2014.2373380	857
..... <i>F. Li, J. Fang, H. Li, and L. Huang</i>	
Fast Decomposition of Large Nonnegative Tensors http://dx.doi.org/10.1109/LSP.2014.2374838	862
..... <i>J. Cohen, R. C. Farias, and P. Comon</i>	
On the Performance of Secure Full-Duplex Relaying under Composite Fading Channels http://dx.doi.org/10.1109/LSP.2014.2374994	867
..... <i>H. Alves, G. Brante, R. D. Souza, D. B. da Costa, and M. Latva-aho</i>	
Local Topographic Shape Patterns for Texture Description http://dx.doi.org/10.1109/LSP.2014.2374608	871
..... <i>C. He, T. Zhuo, X. Su, F. Tu, and D. Chen</i>	
A Novel Biometric Approach for Human Identification and Verification Using Eye Blinking Signal http://dx.doi.org/10.1109/LSP.2014.2374338	876
..... <i>M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas</i>	
Perfect Gaussian Integer Sequences of Odd Period $2^{2^m} - 1$ http://dx.doi.org/10.1109/LSP.2014.2375313	881
..... <i>C.-D. Lee, Y.-P. Huang, Y. Chang, and H.-H. Chang</i>	
A Multi-Modal Topic Model for Image Annotation Using Text Analysis http://dx.doi.org/10.1109/LSP.2014.2375341	886
..... <i>J. Tian, Y. Huang, Z. Guo, X. Qi, Z. Chen, and T. Huang</i>	
Minor Surfaces are Boundaries of Mode-Based Clusters http://dx.doi.org/10.1109/LSP.2014.2376192	891
..... <i>E. Ataer-Cansizoglu, M. Akcakaya, and D. Erdogmus</i>	
An Efficient Frame-Content Based Intra Frame Rate Control for High Efficiency Video Coding http://dx.doi.org/10.1109/LSP.2014.2377032 ..	896
..... <i>M. Wang, K. N. Ngan, and H. Li</i>	
A General Solution to Optimal Fixed-Gain (α - β - γ etc.) Filters http://dx.doi.org/10.1109/LSP.2014.2376876	901
..... <i>D. F. Crouse</i>	
The Effects of Multiple Carrier Frequency Offsets on the Performance of Virtual MISO FSK Systems http://dx.doi.org/10.1109/LSP.2014.2375172	905
..... <i>M. Hussain and S. A. Hassan</i>	
Kernelized Relaxed Margin Components Analysis for Person Re-identification http://dx.doi.org/10.1109/LSP.2014.2377204	910
..... <i>H. Liu, M. Qi, and J. Jiang</i>	
On Computing Jeffrey's Divergence Between Time-Varying Autoregressive Models http://dx.doi.org/10.1109/LSP.2014.2377473	915
..... <i>C. Magnan, A. Giremus, and E. Grivel</i>	
Binary Tomography Reconstructions With Stochastic Level-Set Methods http://dx.doi.org/10.1109/LSP.2014.2375511	920
..... <i>L. Wang, B. Sixou, and F. Peyrin</i>	
Non-Coherent Direction of Arrival Estimation from Magnitude-Only Measurements http://dx.doi.org/10.1109/LSP.2014.2377556	925
..... <i>H. Kim, A. M. Haimovich, and Y. C. Eldar</i>	
Slow Adaptive Power Control and Outage Avoidance in Composite Fading Wireless Channels http://dx.doi.org/10.1109/LSP.2014.2378773 ..	930
..... <i>Y. Seyedi, E. Bolouki, and J.-F. Frigon</i>	
A Probabilistic Interpretation of the Exponential Mixture http://dx.doi.org/10.1109/LSP.2014.2379014	935
..... <i>S. Kay</i>	
Maximizing Network Capacity with Optimal Source Selection: A Network Science Perspective http://dx.doi.org/10.1109/LSP.2014.2378264 ..	938
..... <i>C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu</i>	
Efficient Sub-Regional Multiple-Source Detection Based on Subspace Matrix Filtering http://dx.doi.org/10.1109/LSP.2014.2379619	943
..... <i>Z. Fan, G. Liang, Y. Wang, Y. Wang, and Q. Li</i>	
On the Ozarow-Leung Scheme for the Gaussian Broadcast Channel with Feedback http://dx.doi.org/10.1109/LSP.2014.2375893	948
..... <i>Y. Murin, Y. Kaspi, and R. Dabora</i>	
Kernel Least Mean Square with Single Feedback http://dx.doi.org/10.1109/LSP.2014.2377726	953
..... <i>J. Zhao, X. Liao, S. Wang, and C. K. Tse</i>	
Discrete Receive Beamforming http://dx.doi.org/10.1109/LSP.2014.2379333	958
..... <i>J. Israel, A. Fischer, J. Martinovic, E. A. Jorswieck, and M. Mesyagutov</i>	
Scene Text Identification by Leveraging Mid-level Patches and Context Information http://dx.doi.org/10.1109/LSP.2014.2379625	963
..... <i>R. Wang, N. Sang, and C. Gao</i>	
Estimating Long Tail Models for Risk Trends http://dx.doi.org/10.1109/LSP.2014.2378787	968
..... <i>S. Shenoy and D. Gorinevsky</i>	
Fast Digital Filtering of Spectrometric Data for Pile-up Correction http://dx.doi.org/10.1109/LSP.2014.2377352	973
..... <i>T. Trigano, E. Barat, T. Dautremer, and T. Montagu</i>	
A Time Regularization Technique for Discrete Spectral Envelopes Through Frequency Derivative http://dx.doi.org/10.1109/LSP.2014.2380412	978
..... <i>G. Degottex</i>	
Performance Analysis of Time of Arrival Estimation on OFDM Signals http://dx.doi.org/10.1109/LSP.2014.2378994	983
..... <i>M. Driusso, M. Comisso, F. Babich, and C. Marshall</i>	
Empirical Non-Parametric Estimation of the Fisher Information http://dx.doi.org/10.1109/LSP.2014.2378514	988
..... <i>V. Berisha and A. O. Hero</i>	
Soft-Output MIMO Detectors with Channel Estimation Error http://dx.doi.org/10.1109/LSP.2014.2374425	993
..... <i>L. M.A. Jalloul, S. P. Alex, and M. M. Mansour</i>	

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

AUGUST 2015

VOLUME 22

NUMBER 8

ISPLEM

(ISSN 1070-9908)

LETTERS

Can we Automatically Transform Speech Recorded on Common Consumer Devices in Real-World Environments into Professional Production Quality Speech?—A Dataset, Insights, and Challenges http://dx.doi.org/10.1109/LSP.2014.2379648 <i>G. J. Mysore</i>	1006
Nonlinear Regression with Logistic Product Basis Networks http://dx.doi.org/10.1109/LSP.2014.2380791	
..... <i>M. Sajjadi, M. Seyedhosseini, and T. Tasdizen</i>	1011
Determining Autocorrelation Matrix Size and Sampling Frequency for MUSIC Algorithm http://dx.doi.org/10.1109/LSP.2014.2366638	
..... <i>A. Naha, A. K. Samanta, A. Routray, and A. K. Deb</i>	1016
Polyphase Decomposition of Digital Fractional-Delay Filters http://dx.doi.org/10.1109/LSP.2014.2381603	
FSITM: A Feature Similarity Index For Tone-Mapped Images http://dx.doi.org/10.1109/LSP.2014.2381458	
..... <i>H. Ziaei Nafchi, A. Shahkolaei, R. Farrahi Moghaddam, and M. Cheriet</i>	1026
Stochastic Cramér-Rao Bound Analysis for DOA Estimation in Spherical Harmonics Domain http://dx.doi.org/10.1109/LSP.2014.2381361	
..... <i>L. Kumar and R. M. Hegde</i>	1030
Sequential Changepoint Approach for Online Community Detection http://dx.doi.org/10.1109/LSP.2014.2381553	
..... <i>D. Marangoni-Simonsen and Y. Xie</i>	1035
Perfect Gaussian Integer Sequences of Arbitrary Length http://dx.doi.org/10.1109/LSP.2014.2381642	
A Statistical Method for Improved 3D Surface Detection http://dx.doi.org/10.1109/LSP.2014.2382172	
Real-Time Impulse Noise Suppression from Images Using an Efficient Weighted-Average Filtering http://dx.doi.org/10.1109/LSP.2014.2381649	
A Fine Resolution Frequency Estimator Based on Double Sub-segment Phase Difference http://dx.doi.org/10.1109/LSP.2014.2385086	
..... <i>H. Hosseini, F. Hessar, and F. Marvasti</i>	1050
..... <i>X. Huang and X.-G. Xia</i>	1055
Training Design and Channel Estimation in Uplink Cloud Radio Access Networks http://dx.doi.org/10.1109/LSP.2014.2380776	
..... <i>X. Xie, M. Peng, W. Wang, and H. V. Poor</i>	1060
Expected Power Bound for Two-Dimensional Digital Filters in the Fornasini-Marchesini Local State-Space Model http://dx.doi.org/10.1109/LSP.2014.2382764	
..... <i>C. K. Ahn and H. Kar</i>	1065
Locality-Constrained Sparse Auto-Encoder for Image Classification http://dx.doi.org/10.1109/LSP.2014.2384196	
..... <i>W. Luo, J. Yang, W. Xu, and T. Fu</i>	1070
Compressed Sensing Performance of Random Bernoulli Matrices with High Compression Ratio http://dx.doi.org/10.1109/LSP.2014.2385813 ..	
..... <i>W. Lu, W. Li, K. Kpalma, and J. Ronsin</i>	1074

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 \$309.00 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 © 2015 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage 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.

Energy Efficient Two-Way Non-Regenerative Relaying for Relays with Multiple Antennas http://dx.doi.org/10.1109/LSP.2014.2386310	<i>J. Zhang and M. Haardt</i>	1079
Efficient Soft-Input Soft-Output MIMO Chase Detectors for Arbitrary Number of Streams http://dx.doi.org/10.1109/LSP.2014.2385776	<i>A. Goma and L. M.-A. Jalloul</i>	1084
Automatic Feeding Control for Dense Aquaculture Fish Tanks http://dx.doi.org/10.1109/LSP.2014.2385794	<i>Y. Atoum, S. Srivastava, and X. Liu</i>	1089
Comparison of Adaptive and Model-Free Methods for Dynamic Measurement http://dx.doi.org/10.1109/LSP.2014.2388369	<i>I. Markovskiy</i>	1094
A Volumetric SRP with Refinement Step for Sound Source Localization http://dx.doi.org/10.1109/LSP.2014.2385864	<i>M. V. S. Lima, W. A. Martins, L. O. Nunes, L. W. P. Biscainho, T. N. Ferreira, M. V. M. Costa, and B. Lee</i>	1098
Person Re-Identification by Robust Canonical Correlation Analysis http://dx.doi.org/10.1109/LSP.2015.2390222	<i>L. An, S. Yang, and B. Bhanu</i>	1103
Intrinsic Integer-Periodic Functions for Discrete Periodicity Detection http://dx.doi.org/10.1109/LSP.2014.2387430	<i>S.-C. Pei and K.-S. Lu</i>	1108
Infinite Impulse Response Graph Filters in Wireless Sensor Networks http://dx.doi.org/10.1109/LSP.2014.2387204	<i>X. Shi, H. Feng, M. Zhai, T. Yang, and B. Hu</i>	1113
A 1.5-D Multi-Channel EEG Compression Algorithm Based on NLSPIHT http://dx.doi.org/10.1109/LSP.2015.2389856	<i>G. Xu, J. Han, Y. Zou, and X. Zeng</i>	1118
Enhancing Image Denoising by Controlling Noise Incursion in Learned Dictionaries http://dx.doi.org/10.1109/LSP.2015.2388712	<i>S. K. Sahoo and A. Makur</i>	1123
Primary-Ambient Extraction Using Ambient Phase Estimation with a Sparsity Constraint http://dx.doi.org/10.1109/LSP.2014.2387021	<i>J. He, W.-S. Gan, and E.-L. Tan</i>	1127
Second-Order Statistics Analysis to Cope With Contrast Enhancement Counter-Forensics http://dx.doi.org/10.1109/LSP.2015.2389241	<i>A. De Rosa, M. Fontani, M. Massai, A. Piva, and M. Barni</i>	1132
A Discrete Tchebichef Transform Approximation for Image and Video Coding http://dx.doi.org/10.1109/LSP.2015.2389899	<i>P. A. M. Oliveira, R. J. Cintra, F. M. Bayer, S. Kulasekera, and A. Madanayake</i>	1137
Divergence-Free Wavelet Frames http://dx.doi.org/10.1109/LSP.2015.2388794	<i>E. Bostan, M. Unser, and J. P. Ward</i>	1142
Secure Relay and Jammer Selection for Physical Layer Security http://dx.doi.org/10.1109/LSP.2014.2387860	<i>H. Hui, A. L. Swindlehurst, G. Li, and J. Liang</i>	1147
Parameter Estimation in Spherical Symmetry Groups http://dx.doi.org/10.1109/LSP.2014.2387206	<i>Y.-H. Chen, D. Wei, G. Newstadt, M. DeGraef, J. Simmons, and A. Hero</i>	1152
Reduced-Rank STAP for Slow-Moving Target Detection by Antenna-Pulse Selection http://dx.doi.org/10.1109/LSP.2015.2390148	<i>X. Wang, E. Aboutanios, and M. G. Amin</i>	1156
Word Segmentation Method for Handwritten Documents based on Structured Learning http://dx.doi.org/10.1109/LSP.2015.2389852	<i>J. Ryu, H. I. Koo, and N. I. Cho</i>	1161
A Robust Algorithm for Joint Sparse Recovery in Presence of Impulsive Noise http://dx.doi.org/10.1109/LSP.2014.2387435	<i>J. Shang, Z. Wang, and Q. Huang</i>	1166
On Transmission of a Continuous Signal via a Noiseless Binary Channel http://dx.doi.org/10.1109/LSP.2014.2387427	<i>N. Dokuchaev</i>	1171
Normality and Correlation Coefficient in Estimation of Insulators' Spectral Signature http://dx.doi.org/10.1109/LSP.2015.2390638	<i>G. Fontgalland and H. J. G. Pedro</i>	1175
Separation of Dependent Autoregressive Sources Using Joint Matrix Diagonalization http://dx.doi.org/10.1109/LSP.2014.2380312	<i>A. Boudjellal, A. Mesloub, K. Abed-Meraim, and A. Belouchrani</i>	1180
MR Image Reconstruction with Convolutional Characteristic Constraint (CoCCo) http://dx.doi.org/10.1109/LSP.2014.2376699	<i>X. Peng and D. Liang</i>	1184
Kernel-Based Face Hallucination via Dual Regularization Priors http://dx.doi.org/10.1109/LSP.2015.2390972	<i>J. Shi and C. Qi</i>	1189
An Expectation Propagation Perspective on Approximate Message Passing http://dx.doi.org/10.1109/LSP.2015.2391287	<i>X. Meng, S. Wu, L. Kuang, and J. Lu</i>	1194
An Efficient Jacobi-Like Deflationary ICA Algorithm: Application to EEG Denoising http://dx.doi.org/10.1109/LSP.2014.2385868	<i>S. H. Sardouie, L. Albera, M. B. Shamsollahi, and I. Merlet</i>	1198

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

SEPTEMBER 2015

VOLUME 22

NUMBER 9

ISPLEM

(ISSN 1070-9908)

LETTERS

Minimum Covariance Bounds for the Fusion under Unknown Correlations http://dx.doi.org/10.1109/LSP.2015.2390417	1210
..... <i>M. Reinhardt, B. Noack, P. O. Arambel, and U. D. Hanebeck</i>	
Local Diagonal Extrema Pattern: A New and Efficient Feature Descriptor for CT Image Retrieval http://dx.doi.org/10.1109/LSP.2015.2392623	1215
..... <i>S. R. Dubey, S. K. Singh, and R. K. Singh</i>	
Sparse Signal Recovery from a Mixture of Linear and Magnitude-Only Measurements http://dx.doi.org/10.1109/LSP.2015.2393295	1220
..... <i>M. Akçakaya and V. Tarokh</i>	
Gabor Filter Based on Stochastic Computation http://dx.doi.org/10.1109/LSP.2015.2392123	1224
..... <i>N. Onizawa, D. Katagiri, K. Matsumiya, W. J. Gross, and T. Hanyu</i>	
Fast Non-Negative Orthogonal Matching Pursuit http://dx.doi.org/10.1109/LSP.2015.2393637	1229
..... <i>M. Yaghoobi, D. Wu, and M. E. Davies</i>	
On the Reconstruction of Wavelet-Sparse Signals From Partial Fourier Information http://dx.doi.org/10.1109/LSP.2015.2393953	1234
..... <i>Y. Zhang and P. L. Dragotti</i>	
Preconditioning for Underdetermined Linear Systems with Sparse Solutions http://dx.doi.org/10.1109/LSP.2015.2392000	1239
..... <i>E. Tsiligianni, L. P. Kondi, and A. K. Katsaggelos</i>	
A Novel Generalization of Modified LMS Algorithm to Fractional Order http://dx.doi.org/10.1109/LSP.2015.2394301	1244
..... <i>Y. Tan, Z. He, and B. Tian</i>	
A Systematic Framework for Composite Hypothesis Testing of Independent Bernoulli Trials http://dx.doi.org/10.1109/LSP.2015.2395811	1249
..... <i>D. Ciunzo, A. De Maio, and P. S. Rossi</i>	
Robust Content Fingerprinting Algorithm Based on Sparse Coding http://dx.doi.org/10.1109/LSP.2015.2395726	1254
..... <i>Y. Li</i>	
On the Quality and Timeliness of Fusion in a Random Access Sensor Network http://dx.doi.org/10.1109/LSP.2015.2396573	1259
..... <i>A. Soganli, O. Ercetin, and M. Cetin</i>	
A Robust Scaling Approach for Implementation of HsMMs http://dx.doi.org/10.1109/LSP.2015.2397278	1264
..... <i>B.-C. Li and S.-Z. Yu</i>	
Reversible Color Spaces without Increased Bit Depth and Their Adaptive Selection http://dx.doi.org/10.1109/LSP.2015.2397034	1269
..... <i>T. Strutz and A. Leipnitz</i>	
Capture-Aware Estimation for Large-Scale RFID Tags Identification http://dx.doi.org/10.1109/LSP.2015.2396911	1274
..... <i>Y. Wang, H. Wu, and Y. Zeng</i>	
Quaternion Wiener Deconvolution for Noise Robust Color Image Registration http://dx.doi.org/10.1109/LSP.2015.2398033	1278
..... <i>M. Pedone, E. Bayro-Corrochano, J. Flusser, and J. Heikkilä</i>	
Compressive Sensing Time Reversal MIMO Radar: Joint Direction and Doppler Frequency Estimation http://dx.doi.org/10.1109/LSP.2015.2396650	1283
..... <i>M. H. S. Sajjadih and A. Asif</i>	
On Secrecy Capacity of Helper-Assisted Wiretap Channel with an Out-of-Band Link http://dx.doi.org/10.1109/LSP.2015.2398674	1288
..... <i>L. Li, Z. Chen, and J. Fang</i>	
Histogram-Based Locality-Preserving Contrast Enhancement http://dx.doi.org/10.1109/LSP.2015.2399612	1293
..... <i>J. Shin and R.-H. Park</i>	
Robust Parallel Analog Function Computation via Wireless Multiple-Access MIMO Channels http://dx.doi.org/10.1109/LSP.2015.2399657	1297
..... <i>J. Huang, Q. Zhang, Q. Li, and J. Qin</i>	
Tag Refinement for User-Contributed Images via Graph Learning and Nonnegative Tensor Factorization http://dx.doi.org/10.1109/LSP.2015.2399915	1302
..... <i>Z. Qian, P. Zhong, and R. Wang</i>	

On the Secrecy of Interference-Limited Networks under Composite Fading Channels http://dx.doi.org/10.1109/LSP.2015.2398514	1306
..... <i>H. Alves, C. H. M. de Lima, P. H. J. Nardelli, R. D. Souza, and M. Latva-aho</i>	
Distributed Compressed Estimation Based on Compressive Sensing http://dx.doi.org/10.1109/LSP.2015.2400372	1311
..... <i>S. Xu, R. C. de Lamare, and H. V. Poor</i>	
Simultaneous Blind Gamma Estimation http://dx.doi.org/10.1109/LSP.2015.2396299	1316
..... <i>J. Vazquez-Corral and M. Bernalmío</i>	
ICI-Free Equalization in OFDM Systems with Blanking Preprocessing at the Receiver for Impulsive Noise Mitigation http://dx.doi.org/10.1109/LSP.2015.2398366	1321
..... <i>D. Darsena, G. Gelli, F. Melito, and F. Verde</i>	
Restricted Isometry Property of Subspace Projection Matrix Under Random Compression http://dx.doi.org/10.1109/LSP.2015.2402206	1326
..... <i>X. Shen and Y. Gu</i>	
Visual Tracking via Locally Structured Gaussian Process Regression http://dx.doi.org/10.1109/LSP.2015.2402313	1331
..... <i>Y. Sui and L. Zhang</i>	
Spatial Stimuli Gradient Sketch Model http://dx.doi.org/10.1109/LSP.2015.2404827	1336
..... <i>J. J. Mathew and A. P. James</i>	
Multichannel Random Discrete Fractional Fourier Transform http://dx.doi.org/10.1109/LSP.2015.2402395	1340
..... <i>X. Kang, F. Zhang, and R. Tao</i>	
Structured Saliency Fusion Based on Dempster–Shafer Theory http://dx.doi.org/10.1109/LSP.2015.2399621	1345
..... <i>X. Wei, Z. Tao, C. Zhang, and X. Cao</i>	
Visual Tracking via Constrained Incremental Non-negative Matrix Factorization http://dx.doi.org/10.1109/LSP.2015.2404856	1350
..... <i>H. Zhang, S. Hu, X. Zhang, and L. Luo</i>	
Performance Analysis of Network Coded Cooperation with Channel Coding and Adaptive DF-Based Relaying in Rayleigh Fading Channels http://dx.doi.org/10.1109/LSP.2015.2405083	1354
..... <i>T. X. Vu, P. Duhamel, and M. D. Renzo</i>	
Adaptive Radar Detection and Range Estimation with Oversampled Data for Partially Homogeneous Environment http://dx.doi.org/10.1109/LSP.2015.2404923	1359
..... <i>C. Hao, D. Orlando, G. Foglia, X. Ma, and C. Hou</i>	
Artifact-Free Wavelet Denoising: Non-convex Sparse Regularization, Convex Optimization http://dx.doi.org/10.1109/LSP.2015.2406314	1364
..... <i>Y. Ding and I. W. Selesnick</i>	
Low Cost Pre-Coder Design for MIMO AF Two-Way Relay Channel http://dx.doi.org/10.1109/LSP.2014.2375276	1369
..... <i>Y. Zhang, L. Ping, and Z. Zhang</i>	
Component Extraction for Non-Stationary Multi-Component Signal Using Parameterized De-chirping and Band-Pass Filter http://dx.doi.org/10.1109/LSP.2014.2377038	1373
..... <i>Y. Yang, X. Dong, Z. Peng, W. Zhang, and G. Meng</i>	
Sparse-Distinctive Saliency Detection http://dx.doi.org/10.1109/LSP.2014.2382755	1378
..... <i>Y. Luo, P. Wang, W. Zhu, and H. Qiao</i>	
Adaptive Integral Operators for Signal Separation http://dx.doi.org/10.1109/LSP.2014.2352340	1383
..... <i>X. Hu, S. Peng, and W.-L. Hwang</i>	
Eavesdropping-Based Gossip Algorithms for Distributed Consensus in Wireless Sensor Networks http://dx.doi.org/10.1109/LSP.2015.2398191	1388
..... <i>S. Wu, B. Liu, X. Bai, and Y. Hou</i>	
Pileup Correction Algorithm using an Iterated Sparse Reconstruction Method http://dx.doi.org/10.1109/LSP.2015.2406911	1392
..... <i>T. Trigano, I. Gildin, and Y. Sepulcre</i>	
Salient Object Detection with Higher Order Potentials and Learning Affinity http://dx.doi.org/10.1109/LSP.2014.2377216	1396
..... <i>L. Zhang and X. Yuan</i>	
Semi-Local Structure Patterns for Robust Face Detection http://dx.doi.org/10.1109/LSP.2014.2382762	1400
..... <i>K. Jeong, J. Choi, and G.-J. Jang</i>	
Near-Duplicate Image Retrieval Based on Contextual Descriptor http://dx.doi.org/10.1109/LSP.2014.2377795	1404
..... <i>J. Yao, B. Yang, and Q. Zhu</i>	
Optimal Bartlett Detector Based SPRT for Spectrum Sensing in Multi-antenna Cognitive Radio Systems http://dx.doi.org/10.1109/LSP.2015.2407374	1409
..... <i>S. Dwivedi, A. Kota, and A. K. Jagannatham</i>	
Recovery of Discontinuous Signals Using Group Sparse Higher Degree Total Variation http://dx.doi.org/10.1109/LSP.2015.2407321	1414
..... <i>G. Ongie and M. Jacob</i>	
Structure-Induced Complex Kalman Filter for Decentralized Sequential Bayesian Estimation http://dx.doi.org/10.1109/LSP.2015.2407196	1419
..... <i>A. Mohammadi and K. N. Plataniotis</i>	
Weight-Space Viterbi Decoding Based Spectral Subtraction for Reverberant Speech Recognition http://dx.doi.org/10.1109/LSP.2015.2408371	1424
..... <i>S. M. Ban and H. S. Kim</i>	
Local Smoothness Enforced Cost Volume Regularization for Fast Stereo Correspondence http://dx.doi.org/10.1109/LSP.2015.2409203	1429
..... <i>Q. Yang</i>	
Quantifying the Transmit Diversity Order of Euclidean Distance Based Antenna Selection in Spatial Modulation http://dx.doi.org/10.1109/LSP.2015.2408574	1434
..... <i>R. Rajashekar, K. V. S. Hari, and L. Hanzo</i>	
Remarks on the Spatial Smoothing Step in Coarray MUSIC http://dx.doi.org/10.1109/LSP.2015.2409153	1438
..... <i>C.-L. Liu and P. P. Vaidyanathan</i>	
A Fast Low Complexity Method for Optimal Zero-Forcing Beamformer MU-MIMO System http://dx.doi.org/10.1109/LSP.2015.2407903	1443
..... <i>H. H. Dam, A. Cantoni, and B. Li</i>	
A Step Size Control Method for Deficient Length FBLMS Algorithm http://dx.doi.org/10.1109/LSP.2014.2382113	1448
..... <i>M. Wu and J. Yang</i>	
Visual Tracking via Temporally Smooth Sparse Coding http://dx.doi.org/10.1109/LSP.2014.2365363	1452
..... <i>T. Liu, G. Wang, L. Wang, and K. L. Chan</i>	
Datum-Adaptive Local Metric Learning for Person Re-identification http://dx.doi.org/10.1109/LSP.2014.2387847	1457
..... <i>K. Liu, Z. Zhao, and A. Cai</i>	
A Generic Construction of Z-Periodic Complementary Sequence Sets with Flexible Flock Size and Zero Correlation Zone Length http://dx.doi.org/10.1109/LSP.2014.2369512	1462
..... <i>P. Ke and Z. Zhou</i>	
Coherent Integration for Maneuvering Target Detection Based on Radon-Lv's Distribution http://dx.doi.org/10.1109/LSP.2015.2390777	1467
..... <i>X. Li, G. Cui, W. Yi, and L. Kong</i>	
Subcategory-Aware Object Detection http://dx.doi.org/10.1109/LSP.2014.2299571	1472
..... <i>X. Yu, J. Yang, Z. Lin, J. Wang, T. Wang, and T. Huang</i>	
Learning to Detect Anomalies in Surveillance Video http://dx.doi.org/10.1109/LSP.2015.2410031	1477
..... <i>T. Xiao, C. Zhang, and H. Zha</i>	
Aggregative Modeling of Nonlinear Systems http://dx.doi.org/10.1109/LSP.2015.2405613	1482
..... <i>P. Wachel and P. Śliwiński</i>	
Multimedia Retrieval via Deep Learning to Rank http://dx.doi.org/10.1109/LSP.2015.2410134	1487
..... <i>X. Zhao, X. Li, and Z. Zhang</i>	
Robust Transceiver Optimization for Power-Splitting Based Downlink MISO SWIPT Systems http://dx.doi.org/10.1109/LSP.2015.2410833	1492
..... <i>F. Wang, T. Peng, Y. Huang, and X. Wang</i>	
Cramér–Rao Bound for Sparse Signals Fitting the Low-Rank Model with Small Number of Parameters http://dx.doi.org/10.1109/LSP.2015.2409896	1497
..... <i>M. Shaghghi and S. A. Vorobyov</i>	

IEEE SignalProcessing

MAGAZINE

[VOLUME 32 NUMBER 4 JULY 2015]

ART INVESTIGATION

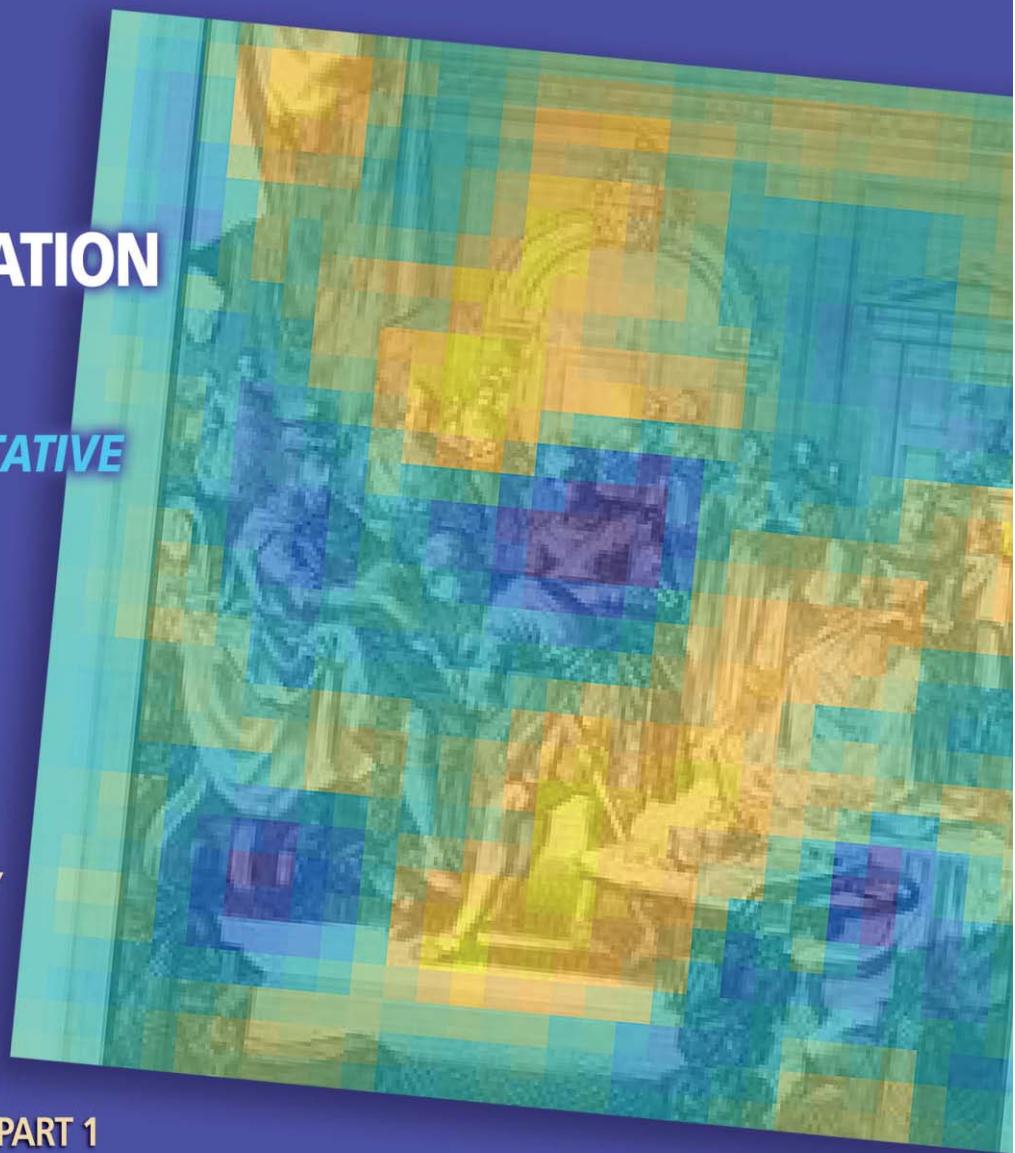
BRIDGING
ART HISTORY
AND QUANTITATIVE
SCIENCES

IMAGE ANALYSIS:
IN PLANT
PHENOTYPING

LECTURE NOTES
ON GAME THEORY
APPLICATIONS

AUTOMATION
IN ANESTHESIA

THE IEEE SIGNAL
PROCESSING CUP: PART 1



IEEE
Signal Processing Society

IEEE

CONTENTS

[VOLUME 32 NUMBER 4]

SPECIAL SECTION—SIGNAL PROCESSING FOR ART INVESTIGATION

14 FROM THE GUEST EDITORS

Patrice Abry, Andrew G. Klein,
William A. Sethares, and
C. Richard Johnson, Jr.

18 MULTISCALE ANISOTROPIC TEXTURE ANALYSIS AND CLASSIFICATION OF PHOTOGRAPHIC PRINTS

Patrice Abry, Stéphane G. Roux,
Herwig Wendt, Paul Messier,
Andrew G. Klein, Nicolas Tremblay,
Pierre Borgnat, Stéphane Jaffard,
Béatrice Vedel, Jim Coddington,
and Lee Ann Daffner

28 HUNTING FOR PAPER MOLDMATES AMONG REMBRANDT'S PRINTS

C. Richard Johnson, Jr., William A.
Sethares, Margaret Holben Ellis,
and Saira Haqqi

38 AUTOMATIC THREAD-LEVEL CANVAS ANALYSIS

Laurens van der Maaten
and Robert Erdmann

46 TOWARD DISCOVERY OF THE ARTIST'S STYLE

Nanne van Noord, Ella Hendriks,
and Eric Postma

55 QUANTITATIVE CANVAS WEAVE ANALYSIS USING 2-D SYNCHROQUEEZED TRANSFORMS

Haizhao Yang, Jianfeng Lu,
William P. Brown, Ingrid Daubechies,
and Lexing Ying

64 ANCIENT COIN CLASSIFICATION USING REVERSE MOTIF RECOGNITION

Hafeez Anwar, Sebastian Zambanini,
Martin Kampel, and Klaus Vondrovec

75 MULTIMEDIA ANALYSIS AND ACCESS OF ANCIENT MAYA EPIGRAPHY

Rui Hu, Gulcan Can, Carlos Pallán
Gayol, Guido Krempel, Jakub Spotak,
Gabrielle Vail, Stephane
Marchand-Maillet, Jean-Marc Odobez,
and Daniel Gatica-Perez

85 COMPUTERIZED FACE RECOGNITION IN RENAISSANCE PORTRAIT ART

Ramya Srinivasan, Conrad Rudolph,
and Amit Roy-Chowdhury

95 CHALLENGES IN CONTENT-BASED IMAGE INDEXING OF CULTURAL HERITAGE COLLECTIONS

David Picard, Philippe-Henri Gosselin,
and Marie-Claude Gaspard

103 PDE-BASED GRAPH SIGNAL PROCESSING FOR 3-D COLOR POINT CLOUDS

François Lozes, Abderrahim Elmoataz,
and Olivier Lézoray

112 DIGITAL IMAGE PROCESSING OF THE GHENT ALTARPIECE

Aleksandra Pižurica, Ljiljana Platiša,
Tijana Ružić, Bruno Cornelis,
Ann Dooms, Maximiliaan Martens,
Hélène Dubois, Bart Devolder,
Marc De Mey, and Ingrid Daubechies

COLUMNS

4 FROM THE EDITOR

Art, Engineering,
and Community
Min Wu

6 PRESIDENT'S MESSAGE

SigPort: A Paper Repository for
Signal Processing
Alex Acero

8 SPECIAL REPORTS

Human–Machine Interfaces:
Methods of Control
John Edwards

123 SP EDUCATION

Undergraduate Students
Compete in the IEEE Signal
Processing Cup: Part 1
Kin-Man Lam, Carlos Óscar S.
Sorzano, Zhilin Zhang,
and Patrizio Campisi

126 APPLICATIONS CORNER

Image Analysis: The New
Bottleneck in Plant Phenotyping
Massimo Minervini, Hanno Scharr,
and Sotirios A. Tsaftaris

132 LECTURE NOTES

Understanding Game Theory via
Wireless Power Control
Giacomo Bacci, Luca Sanguinetti,
and Marco Luise

138 LIFE SCIENCES

Signal Processing and Automation
in Anesthesia
Guy A. Dumont

145 SP TIPS & TRICKS

Fast, Accurate, and Guaranteed
Stable Sliding Discrete
Fourier Transform
Chun-Su Park

150 STANDARDS IN A NUTSHELL

The High Efficiency Image File
Format Standard
Miska M. Hannuksela, Jani Lainema,
and Vinod K. Malamal Vadakital

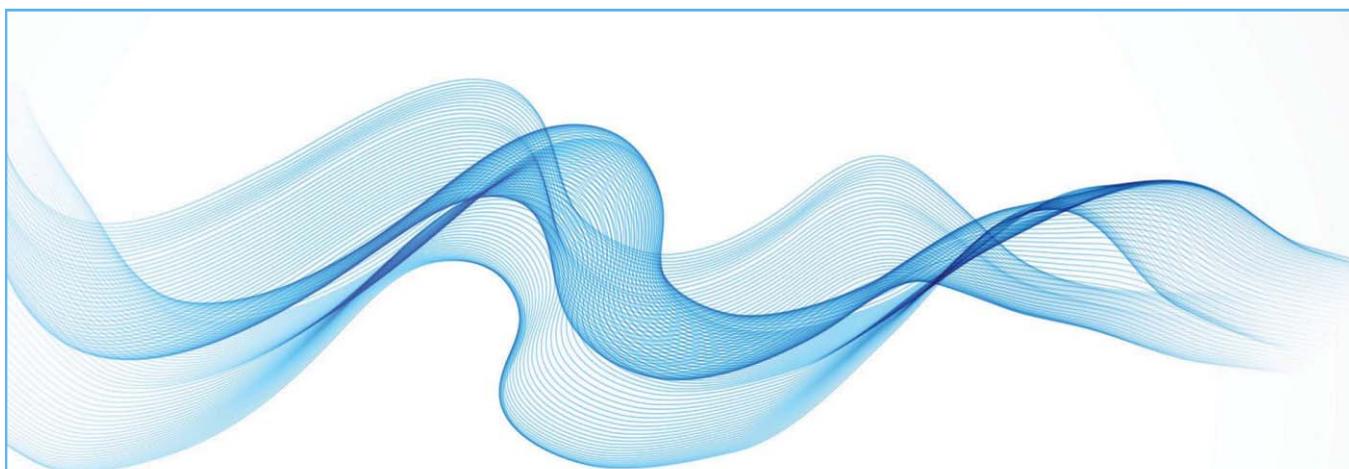
DEPARTMENTS

12 SOCIETY NEWS

New Society Officers for 2016 and
Election of Directors-at-Large and
Members-at-Large

160 DATES AHEAD

Digital Object Identifier 10.1109/MSP.2015.2419312



FREE SPS STUDENT MEMBERSHIP FOR 2015

You're in the beginning stages of your career. Membership in the IEEE Signal Processing Society can help you lay the groundwork for many years of success. You can have it all in 2015 - and for free! Membership includes:

- **Discounts** on conference registration fees;
- Eligibility to apply for **travel grants** to attend SPS flagship conferences including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) and IEEE International Conference on Image Processing (ICIP);
- **Networking** and **job opportunities** at the ICASSP Student Career Luncheon;
- Eligibility to enter our **student competition**, the Signal Processing Cup, for a US\$5,000 grand prize;
- **Involvement opportunities** through SPS's local Chapters - more than 130 worldwide;
- **Free** electronic and digital subscriptions to *IEEE Signal Processing Magazine*, *Inside Signal Processing eNewsletter*, and the *IEEE Signal Processing Society Content Gazette*;
- Access to cutting-edge **educational resources**, including SigView, SPS's online video tutorial portal.

See everything Signal Processing Society membership can do for you:
<http://signalprocessingsociety.org>

Already an IEEE member? Join SPS for free now!
 (You must have already renewed your IEEE membership for 2015 to use this offer)

- Visit <http://iee.org/join>
- On the left side, click "Societies and Special Interest Groups"
- Click "IEEE Signal Processing Society," then "Join the IEEE Signal Processing Society"
- When you reach the catalog page, click "Add Item(s)" and sign in with your IEEE account *Note: Free offer applies only to basic membership. For US\$8.00, enhance your membership for more great benefits!*
- Once logged in, click "Proceed to Checkout"
- When you reach the shopping cart, enter the promotion code **SP15STUAD** and click "Apply"
- Complete check out and congratulations! Welcome to SPS!

Not yet an IEEE Student Member?

Get a free SPS membership with the purchase of an IEEE Student membership!

- Visit <http://iee.org/join>
- Click "Join as a student" on the bottom right to create your new IEEE Student member account
- After your IEEE account is created, complete the membership application and proceed to "Do you want to add any memberships and subscriptions?"
- Select "Signal Processing Society membership" and click "add selected item"
- Click "Proceed to Checkout"
- When you reach the shopping cart, enter the promotion code **SP15STUNW** and click "Apply"
- Complete check out and congratulations! Welcome to SPS!

Note: Must be an active IEEE Student or Graduate Student member. This offer does not apply to SPS Students or Graduate Students renewing for 2015.



1st Call for Papers

Co-sponsored by



The 15th IEEE International Symposium on Signal Processing and Information Technology

December 7-10, 2015, Abu Dhabi, UAE

Supported by

**General Chair**

Adel Elmaghraby
University of Louisville, USA
Begoña García Zapirain
DeustoTech, Spain

Technical Prog. Chair

Esam Abdel-Raheem
University of Windsor, Canada

Technical Prog. Co-Chairs

Mohammed Ghazal
Abu Dhabi University, UAE
Walaa Sheta
City for Scientific Research, Egypt

Registration & Finance Chair

Reda Ammar
University of Connecticut, USA

Publication Co-Chairs

Hassan Hajjdiab
Abu Dhabi University, UAE
Zakaria Maamar
Zayed University, UAE

Tutorials Co-Chairs

Murad Elhadef
Abu Dhabi University, UAE
Sartaj Sahni
University of Florida, USA

Plenary & Special Sessions Co-Chairs

Braham Barkat
The Petroleum Institute, UAE
Christos Douligeris
University of Piraeus, Greece

Publicity Co-Chairs

Manar Abu Taleb
University of Sharjah, UAE
Adel Khelifi
Al Hosn University, UAE
Marco Re
Univ. of Rome "Tor Vergata", Italy

Industrial Liaison Co-Chairs

Mohammed Abu Khousa
The Petroleum Institute, UAE
Motasir Qasymeh
Abu Dhabi University, UAE

Local Arrangements Co-Chairs

Riad Kanan
Ashraf Khalil
Abu Dhabi University, UAE

Web Manager

Mostafa G. Mostafa
Mckendree University, USA

The **IEEE ISSPIT 2015** is the fifteenth in a series of international symposia that aims to cover most of the aspects in the fields of signal processing and information technology. Sessions will include tutorials in addition to presentations on new research results. Papers describing original work are invited in any of the areas listed below. Accepted papers will be published in the *Proceedings of IEEE ISSPIT 2015* and will be available via **IEEE Xplore**. Acceptance will be based on quality, relevance, and originality. Contest for Best Paper Award will be held and award will be given.

Papers are invited in the following **topics**:

- Signal Processing Theory and Methods
- Signal Processing for Communications and Networking
- Design & Implementation of Signal Processing Systems
- Image, Video & Multidimensional Signal Processing
- Multimedia Signal Processing
- Biological Image and signal processing
- Audio and Acoustic signal Processing
- Health Informatics and e-Health
- Sensor Arrays
- Radar Signal Processing
- Internet Software Architectures
- Multimedia and Image Based Systems
- Mobile Computing and Applications
- E-Commerce
- Bioinformatics and Bioengineering
- Information Processing
- Geographical Information Systems
- Object Based Software Engineering
- Speech Processing
- Computer Networks
- Neural Network

Prospective authors are invited to submit full-length, 6-page (max) papers in two-column formats including diagrams and references. Authors can submit their papers as PDF files through the online submission system found on the ISSPIT website: www.isspit.org. The title page should include author(s) name(s), affiliation, mailing address, telephone, fax, and e-mail address. The author should indicate one or two of the above categories that best describe the topic of the paper.

Important Dates

Proposals for Tutorials & Special Sessions	Sept. 4 th , 2015
Regular paper submission	Sept. 4 th , 2015
Notification of acceptance	Oct. 9 th , 2015
Final version paper with registration	Oct. 23 rd , 2015

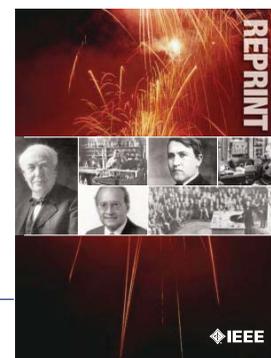
website: <http://www.isspit.org>



IEEE ORDER FORM FOR REPRINTS

Purchasing IEEE Papers in Print is easy, cost-effective and quick.

Complete this form, send via our secure fax (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
Piscataway, NJ 08855-1331

Email the Reprint Department at
reprints@ieee.org for questions regarding
this form

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.
- \$21.50 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: _____

2012 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	\$129	\$213	\$245	\$248	\$288	\$340	\$371	\$408	\$440	\$477	\$510	\$543
100	\$245	\$425	\$479	\$495	\$573	\$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.

2015 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
2. Engineering management
3. Project engineering management
4. Research and development management
5. Design engineering management—analogue
6. Design engineering management—digital
7. Research and development engineering
8. Design/development engineering—analogue
9. Design/development engineering—digital
10. Hardware engineering
11. Software design/development
12. Computer science
13. Science/physics/mathematics
14. Engineering (not elsewhere specified)
15. Marketing/sales/purchasing
16. Consulting
17. Education/teaching
18. Retired
19. Other _____

C. Principal responsibility

1. Engineering and scientific management
2. Management other than engineering
3. Engineering design
4. Engineering
5. Software: science/mngmnt/engineering
6. Education/teaching
7. Consulting
8. Retired
9. Other _____

D. Title

1. Chairman of the Board/President/CEO
2. Owner/Partner
3. General Manager
4. VP Operations
5. VP Engineering/Dir. Engineering
6. Chief Engineer/Chief Scientist
7. Engineering Management
8. Scientific Management
9. Member of Technical Staff
10. Design Engineering Manager
11. Design Engineer
12. Hardware Engineer
13. Software Engineer
14. Computer Scientist
15. Dean/Professor/Instructor
16. Consultant
17. Retired
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/Effective January 2015)

For Transactions and Journals:

Authors are encouraged to submit manuscripts of Regular papers (papers which provide a complete disclosure of a technical premise), or Comment Correspondences (brief items that provide comment on a paper previously published in these TRANSACTIONS).

Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere.

Every manuscript must:

- i. provide a clear statement of the problem and what the contribution of the work is to the relevant research community;
- ii. state why this contribution is significant (what impact it will have);
- iii. provide citation of the published literature most closely related to the manuscript; and
- iv. state what is distinctive and new about the current manuscript relative to these previously published works.

By submission of your manuscript to these TRANSACTIONS, all listed authors have agreed to the authorship list and all the contents and confirm that the work is original and that figures, tables and other reported results accurately reflect the experimental work. In addition, the authors all acknowledge that they 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. The corresponding author is responsible for obtaining consent from all co-authors and, if needed, from sponsors before submission.

In order to be considered for review, a paper must be within the scope of the journal and represent a novel contribution. A paper is a candidate for an Immediate Rejection if it is of limited novelty, e.g. a straightforward combination of theories and algorithms that are well established and are repeated on a known scenario. Experimental contributions will be rejected without review if there is insufficient experimental data. These TRANSACTIONS are published in English. Papers that have a large number of typographical and/or grammatical errors will also be rejected without review.

In addition to presenting a novel contribution, acceptable manuscripts must describe and cite related work in the field to put the contribution in context. Do not give theoretical derivations or algorithm descriptions that are easily found in the literature; merely cite the reference.

New and revised manuscripts should be prepared following the "Manuscript Submission" guidelines below, and submitted to the online manuscript system, ScholarOne Manuscripts. 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 Manuscript system.

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-column, single-spaced format using a font size of 10 points or larger, having a margin of at least 1 inch on all sides. Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOne-Manuscripts site. Since many reviewers prefer a larger font, you are strongly encouraged to also submit a single-column, double-spaced version (11 point font or larger), which is easy to create with the templates provided **IEEE Author Digital Toolbox** (http://www.ieee.org/publications_standards/publications/authors/authors_journals.html). Page length restrictions will be determined by the double-column

version. 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 .PDF 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. *EDICS (Not applicable to Journal of Selected Topics in Signal Processing).* All submissions must be classified by the author with an EDICS (Editors' Information Classification Scheme) selected from the list of EDICS published online at the at the publication's EDICS webpage (*please see the list below). 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.
4. *Additional Documents for Review.* Please upload pdf versions of all items in the reference list that are not publicly available, such as unpublished (submitted) papers. Graphical abstracts and supplemental materials intended to appear with the final paper (see below) must also be uploaded for review at the time of the initial submission for consideration in the review process. 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. *Supplemental Materials.* IEEE Xplore can publish multimedia files (audio, images, video), datasets, and software (e.g. 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 IEEE Author Digital Toolbox under "Multimedia." To make your work reproducible by others, these TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper.
6. *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.
7. *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.

Comment Correspondence. Comment Correspondences provide brief comments on material previously published in these TRANSACTIONS. These items may not exceed 2 pages in double-column, single spaced format, using 9 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 "Manuscript Submission" above for instructions).

Authors may also submit manuscripts of overview articles, but note that these include an additional white paper approval process <http://www.signalprocessingsociety.org/publications/overview-articles/>. [This does not apply to the Journal of Selected Topics in Signal Processing. Please contact the Editor-in-Chief.]

Digital Object Identifier

Manuscript Length. For the initial submission of a regular paper, the manuscript may not exceed 13 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. Supplemental materials and graphical abstracts are not included in the page count. For regular papers, the revised manuscript may not exceed 16 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. For Overview Papers, the maximum length is double that for regular submissions at each stage (please reference <http://www.signalprocessingsociety.org/publications/overview-articles/> for more information).

Note that any paper in excess of 10 pages will be subject to mandatory overlength page charges. 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. Note: Papers submitted to the TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.

Exceptions to manuscript length requirements may, under extraordinary circumstances, be granted by the Editor-in-Chief. However, such exception does not obviate your requirement to pay any and all overlength or additional charges that attach to the manuscript.

Resubmission of Previously Rejected Manuscripts. Authors of manuscripts rejected from any journal are allowed to resubmit their manuscripts only once. At the time of submission, you will be asked whether your manuscript is a new submission or a resubmission of an earlier rejected manuscript. If it is a resubmission of a manuscript previously rejected by any journal, you are expected to submit supporting documents identifying the previous submission and detailing how your new version addresses all of the reviewers' comments. Papers that do not disclose connection to a previously rejected paper or that do not provide documentation as to changes made may be immediately rejected.

Author Misconduct. Author misconduct includes plagiarism, self-plagiarism, and research misconduct, including falsification or misrepresentation of results. All forms of misconduct are unacceptable and may result in sanctions and/or other corrective actions. 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 author's previously published work that also involves other authors. Self-plagiarism involves the verbatim copying or reuse of an authors own prior work without appropriate citation, including 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. For more information on the definitions, investigation process, and corrective actions related to author misconduct, see the Signal Processing Society Policies and Procedures Manual, Section 6.1. <http://www.signalprocessingsociety.org/about-sps/governance/policy-procedure/part-2>. Author misconduct may also be actionable by the IEEE under the rules of Member Conduct.

Extensions of the Author's Prior 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 their related prior work; the papers cannot be identical; and the journal publication must include substantively novel aspects such as new experimental results and analysis or added theoretical work. The journal publication should clearly specify how the journal paper offers novel contributions when citing the prior work. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper, and the prior work must be cited as the primary source.

Submission Format. Authors are required to prepare manuscripts employing the on-line style files developed by IEEE, which include guidelines for abbreviations, mathematics, and graphics. 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 the **IEEE Author Digital Toolbox** under "Template for all TRANSACTIONS." (LaTeX and MS Word). Please note the following requirements about the abstract:

- The abstract must be a concise yet comprehensive reflection of what is in your article.
- The abstract must be self-contained, without abbreviations, footnotes, displayed equations, or references.

- The abstract must be between 150-250 words.
- The abstract should include a few keywords or phrases, as this will help readers to find it. Avoid over-repetition of such phrases as this can result in a page being rejected by search engines.

In addition to written abstracts, papers may include a graphical abstract; see http://www.ieee.org/publications_standards/publications/authors/authors_journals.html for options and format requirements.

IEEE supports the publication of author names in the native language alongside the English versions of the names in the author list of an article. For more information, see "Author names in native languages" (http://www.ieee.org/publications_standards/publications/authors/auth_names_native_lang.pdf) on the IEEE Author Digital Toolbox page.

Open Access. The publication is a hybrid journal, allowing either Traditional manuscript submission or Open Access (author-pays OA) manuscript submission. Upon submission, if you choose to have your manuscript be an Open Access article, you commit to pay the discounted \$1,750 OA fee if your manuscript is accepted for publication in order to enable unrestricted public access. Any other application charges (such as overlength page charge and/or charge for the use of color in the print format) will be billed separately once the manuscript formatting is complete but prior to the publication. If you would like your manuscript to be a Traditional submission, your article will be available to qualified subscribers and purchasers via IEEE Xplore. No OA payment is required for Traditional submission.

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 (two 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. (**NOTE: Papers accepted to IEEE TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.) 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. No mandatory overlength page charges will be applied to overview articles in the Society's journals.

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 charges for print are the responsibility of the author. Details of the associated charges can be found on the IEEE Publications page.

Payment of fees on color reproduction is not negotiable or voluntary, and the author's agreement to publish the manuscript in these TRANSACTIONS is considered acceptance of this requirement.

*EDICS Webpages:

IEEE TRANSACTIONS ON SIGNAL PROCESSING:

<http://www.signalprocessingsociety.org/publications/periodicals/tsp/TSP-EDICS/>

IEEE TRANSACTIONS ON IMAGE PROCESSING:

<http://www.signalprocessingsociety.org/publications/periodicals/image-processing/tip-edics/>

IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE / ACM:

<http://www.signalprocessingsociety.org/publications/periodicals/taslp/taslp-edics/>

IEEE TRANSACTIONS ON INFORMATION, FORENSICS AND SECURITY:

<http://www.signalprocessingsociety.org/publications/periodicals/forensics/forensics-edics/>

IEEE TRANSACTIONS ON MULTIMEDIA:

<http://www.signalprocessingsociety.org/tmm/tmm-edics/>

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING:

<http://www.signalprocessingsociety.org/publications/periodicals/tci/tci-edics/>

IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS:

<http://www.signalprocessingsociety.org/publications/periodicals/tsipn/tsipn-edics/>

