

IEEE Signal Processing

Volume 33 | Number 6 | November 2016

MAGAZINE

SMART VEHICLE TECHNOLOGIES

Signal Processing on the Move

The Race to Improve Radar Imagery

Matching Theory for Wireless Communications

Enriching Undergraduate Programs with SP Research

Indoor Wi-Fi GPS with Centimeter Accuracy

IEEE Signal Processing Society



Fuel your imagination.

The IEEE Member Digital Library gives you the latest technology research—so you can connect ideas, hypothesize new theories, and invent better solutions.

Get full-text access to the IEEE *Xplore*® digital library—at an exclusive price—with the only member subscription that includes any IEEE journal article or conference paper.

Choose from two great options designed to meet the needs of every IEEE member:

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.

Contents

Volume 33 | Number 6 | November 2016

SPECIAL SECTION

SIGNAL PROCESSING FOR SMART VEHICLE TECHNOLOGIES

12 FROM THE GUEST EDITORS

John H.L. Hansen,
Kazuya Takeda,
Sanjeev M. Naik,
Mohan M. Trivedi,
Gerhard U. Schmidt, and
Yingying (Jennifer) Chen

14 DRIVER-BEHAVIOR MODELING USING ON-ROAD DRIVING DATA

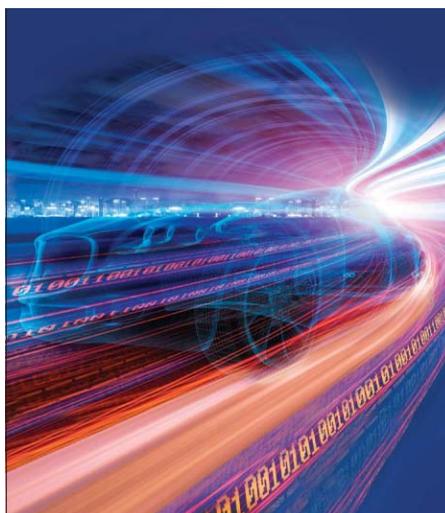
Chiyomi Miyajima
and Kazuya Takeda

22 DRIVER STATUS MONITORING SYSTEMS FOR SMART VEHICLES USING PHYSIOLOGICAL SENSORS

Youjun Choi, Sang Ik Han,
Seung-Hyun Kong,
and Hyunwoo Ko



PG. 123



ON THE COVER

Signal processing is playing an increasingly substantial role in advancing today's vehicles into "smart" ones. The special section in this issue provides a venue for summarizing, educating, and sharing the state of the art in signal processing applied to automotive systems.

COVER IMAGE: @ISTOCKPHOTO.COM/NADLA

35 SMART DRIVER MONITORING: WHEN SIGNAL PROCESSING MEETS HUMAN FACTORS

Amirhossein S. Aghaei,
Birsen Donmez,
Cheng Chen Liu, Dengbo He,
George Liu, Konstantinos N.
Plataniotis, Hwei-Yen Winnie Chen,
and Zohreh Sojoudi

49 CONVERSATIONAL IN-VEHICLE DIALOG SYSTEMS

Fuliang Weng,
Pongtep Angkititrakul,
Elizabeth E. Shriberg,
Larry Heck, Stanley Peters,
and John H.L. Hansen

61 RECENT ADVANCES IN ACTIVE NOISE CONTROL INSIDE AUTOMOBILE CABINS

Prasanga N. Samarasinghe,
Wen Zhang, and
Thushara D. Abhayapala

74 COORDINATION OF COOPERATIVE AUTONOMOUS VEHICLES

Robert Hult, Gabriel R. Campos,
Erik Steinmetz, Lars Hammarstrand,
Paolo Falcone, and Henk Wymeersch

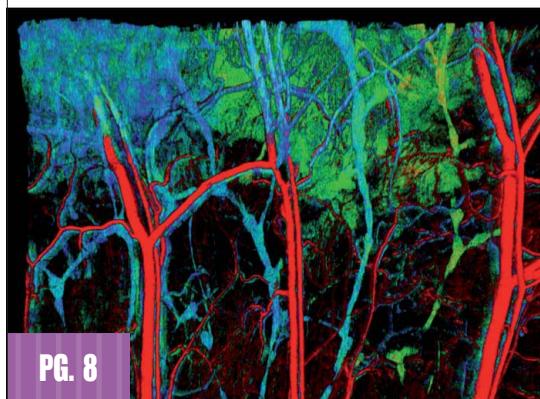
FEATURES

85 THE RACE TO IMPROVE RADAR IMAGERY

Lifan Zhao, Lu Wang,
Lei Yang, Abdelhak M. Zoubir,
and Guoan Bi

103 MATCHING THEORY

Siavash Bayat, Yonghui Li,
Lingyang Song, and Zhu Han



PG. 8

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 © 2016 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.2016.2610483

COLUMNS

- 6 Reader's Choice**
Top Downloads in IEEE *Xplore*
- 8 Special Reports**
Signal Processing at the Heart of a Health-Care Renaissance
John Edwards
- 123 SP Education**
Enriching the Undergraduate Program with Research Projects
Ed Richter and Arye Nehorai
- 128 Applications Corner**
Indoor Global Positioning System with Centimeter Accuracy Using Wi-Fi
Chen Chen, Yi Han, Yan Chen, and K.J. Ray Liu
- 135 Lecture Notes**
Efficient Adjoint Computation for Wavelet and Convolution Operators
James Folberth and Stephen Becker
- 141 Tips & Tricks**
Reducing Quantization Errors for Inner-Product Operations in Embedded Digital Signal Processing Systems
Zhuo Wang, Jintao Zhang, and Naveen Verma



IMAGE LICENSED BY INGRAM PUBLISHING

The 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP) will be held 7–9 December 2016 in the greater Washington, D.C., area.

DEPARTMENTS

- 3 From the Editor**
Publishing Articles in *IEEE Signal Processing Magazine*
Min Wu
- 4 President's Message**
Meeting the Needs of Our Members
Rabab Ward
- 148 Dates Ahead**



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

IEEE Signal Processing Magazine

EDITOR-IN-CHIEF

Min Wu—University of Maryland, College Park
U.S.A.

AREA EDITORS

Feature Articles

Shuguang Robert Cui—Texas A&M University,
U.S.A.

Special Issues

Douglas O'Shaughnessy—INRS, Canada

Columns and Forum

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, U.S.A.

EDITORIAL BOARD

Mrityunjoy Chakraborty—Indian Institute of
Technology, Kharagpur, India

George Chrisikos—Qualcomm, Inc.,
U.S.A.

Patrick Flandrin—ENS Lyon, France

Mounir Ghogho—University of Leeds,
U. K.

Lina Karam—Arizona State University, U.S.A.

Hamid Krim—North Carolina State University,
U.S.A.

Sven Lončarić—University of Zagreb, Croatia

Brian Lovell—University of Queensland, Australia

Jian Lu—Qihoo 360, China

Henrique (Rico) Malvar—Microsoft Research,
U.S.A.

Yi Ma—ShanghaiTech University, China

Stephen McLaughlin—Heriot-Watt University,
Scotland

Athina Petropulu—Rutgers University,
U.S.A.

Peter Ramadge—Princeton University,
U.S.A.

Shigeki Sagayama—Meiji University, Japan

Erchin Serpedin—Texas A&M University,
U.S.A.

Shihab Shamma—University of Maryland,
U.S.A.

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

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

Pramod K. Varshney—Syracuse University,
U.S.A.

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

Gregory Wornell—Massachusetts Institute
of Technology, U.S.A.

Dapeng Wu—University of Florida, U.S.A.

ASSOCIATE EDITORS—COLUMNS AND FORUM

Ivan Bajic—Simon Fraser University, Canada

Rodrigo Capobianco Guido—

São Paulo State University, Brazil

Ching-Te Chiu—National Tsing Hua University,
Taiwan

Michael Gormish—Ricoh Innovations, Inc.

Xiaodong He—Microsoft Research

Danilo Mandic—Imperial College, U.K.

Aleksandra Mojsilovic—

IBM T.J. Watson Research Center

Fatih Porikli—MERL

Shantanu Rane—PARC, U.S.A.

Saeid Sanei—University of Surrey, U.K.

Roberto Togneri—The University of
Western Australia

Alessandro Vinciarelli—IDIAP-EPFL

Azadeh Vosoughi—University of Central Florida

Stefan Winkler—UIUC/ADSC, Singapore

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, U.S.A.

Pierluigi Failla—Compass Inc, New York,
U.S.A.

Marco Guerriero—General Electric Research,
U.S.A.

Yang Li—Harbin Institute of Technology, China

Yuhong Liu—Penn State University at Altoona,
U.S.A.

Andreas Merentitis—University of Athens, Greece

Michael Muma—TU Darmstadt, Germany

Xiaorong Zhang—San Francisco State University,
U.S.A.

ASSOCIATE EDITOR—SOCIAL MEDIA/OUTREACH

Guijin Wang—Tsinghua University, China

IEEE SIGNAL PROCESSING SOCIETY

Rabab Ward—President

Ali Sayed—President-Elect

Carlo S. Regazzoni—Vice President, Conferences

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

Thrasivoulos (Thrasos) N. Pappas—
Vice President, Publications

Walter Kellerman—Vice President,
Technical Directions

IEEE SIGNAL PROCESSING SOCIETY STAFF

Denise Hurley—Senior Manager of Conferences
and Publications

Rebecca Wollman—Publications Administrator

IEEE PERIODICALS MAGAZINES DEPARTMENT

Jessica Barragué, *Managing Editor*

Geraldine Krolin-Taylor, *Senior Managing Editor*

Mark David, *Senior Manager
Advertising and Business Development*

Felicia Spagnoli, *Advertising Production Manager*

Janet Dudar, *Senior Art Director*

Gail A. Schnitzer, Mark Morrissey,
Associate Art Directors

Theresa L. Smith, *Production Coordinator*

Dawn M. Melley, *Editorial Director*

Peter M. Tuohy, *Production Director*

Fran Zappulla, *Staff Director,
Publishing Operations*

Digital Object Identifier 10.1109/MSP.2016.2610485

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.

Min Wu | Editor-in-Chief | minwu@umd.edu

Publishing Articles in *IEEE Signal Processing Magazine*

The 2016 update of *Journal Citation Report (JCR)* was recently released by Thomson Reuters. Based on 2015 citation data for literature within the sciences and social sciences, *IEEE Signal Processing Magazine (SPM)* has continued to score high in this annual citation study. Its journal impact factor in this 2016 release reached 6.67, the highest ever in *SPM*'s history! It is in the top 1% out of more than 250 publications in the electrical and electronics category examined by the *JCR* and continues to be the top-ranked magazine in the IEEE.

SPM's sustained success could not be possible without the tireless effort from our area editors and associate editors on both the current team and the teams before. They work with prospective authors to solicit articles on timely topics of broad interests, oversee peer reviews, and polish the accepted articles until the final stage of production. The Publications Board of the IEEE Signal Processing Society (SPS) initiated an annual recognition to honor the contributions of outstanding editors and editorial board members. For this inaugural year, Dr. Fulvio Gini, area editor for special issues in 2012–2014, and Dr. Gwenael Doerr, area editor for columns and forum in 2015–2016, were recognized for their outstanding contributions. I extend my heartfelt congratulations and appreciation to them!

As another essential part of our team, senior editorial board members offer vital advice. I would like to take this opportunity

to thank six retiring board members for their valuable contributions. We appreciate the prompt and thoughtful feedback by Dr. Isabel Trancoso and Dr. Hing Cheung So on nearly every proposal for special issues and feature articles sent to the editorial board; Dr. Pramod Varshney, Dr. Hamid Krim, and Dr. Patrick Flandrin offered constructive input on the magazine's policy and potential topics; Dr. Z. Jane Wang took the lead in organizing a wonderful pilot of a feature articles cluster—the May 2016 issue on brain analytics, allowing the magazine to develop a synergistic option between feature articles and a special issue. Thanks also go to two retiring area editors, Dr. Wade Trappe and Dr. Gwenael Doërr, for their contributions to special issues and columns, respectively, and a warm welcome to Dr. Douglas O'Shaughnessy as our new area editor for special issues.

By far, the most important part of any successful publication is the authors, who, in our case, also include the guest editors proposing and organizing special issues and article clusters or series. They are the heroes behind the magazine's sustained high impact.

When welcoming authors from around the world, we note that many initial submissions often have to be declined for being out of scope, because our magazine is very different from other publications that focus on reporting new research results. Instead, *SPM* publishes tutorial-style surveys, overviews, and column articles of interest to the broad signal processing professionals and related communities. We do not consider papers proposing brand

new algorithms that have not been sufficiently examined by the technical community, and these types of articles should be sent to journals. Articles also need to be written in an accessible style for a audience broader than experts only, using visually appealing illustrations and keeping mathematical equations and symbols to a minimum.

We strongly suggest prospective authors read *SPM*'s online "Information for Authors" [1] and refer to recently published articles in the magazine for examples. Before working on a full-length feature article or organizing a special issue/section on a central theme of sufficient technical interests and maturity, authors should first prepare a short proposal as outlined in the author guide. Complementing feature articles and special issue articles in each issue are shorter articles in columns and forums on a wide range of topics in signal processing, and they are divided in categories with different objectives and style. Authors wishing to submit an article or a proposal are encouraged to contact me or the respective area editor. We look forward to working with you and hope that you consider *SPM* as your top choice in which to publish when you have source material in the magazine's scope and style.

Reference

[1] IEEE Signal Processing Magazine: Information for Authors. [Online]. Available: <http://signalprocessing-society.org/publications-resources/ieee-signal-processing-magazine/information-authors-spm>

SP

PRESIDENT'S MESSAGE

Rabab Ward | SPS President | rababw@ece.ubc.ca

Meeting the Needs of Our Members

The purpose and face of professional associations and their memberships are ever-evolving, posing a set of challenges that are equal parts daunting and exciting. The IEEE Signal Processing Society (SPS), in particular, faces unique circumstances as our organization changes rapidly because of the technologies we support. Keeping up can be a challenge, but our dedicated volunteers and staff work aggressively to ensure that the SPS stays in the forefront in offering world-class benefits as nimble as signal processing itself.

We have more than 17,000 members in our vibrant community, spanning more than 120 countries and countless fields, technical interests, professions, and career stages. Each individual's needs are different. How do we attract and engage our growing population of student and young professional members, while still serving and honoring our mid-to-late-career and retired members? How do we manage all of that and still act as a signal processing resource to the general public? And how do we support members—and nonmembers—across all of these areas without compromising value among one, or all, of these groups? Offering valuable, flexible benefits is key to engaging members early and also by offering products, services, and benefits that will carry them through every stage of their careers.

*Digital Object Identifier 10.1109/MSP.2016.2610182
Date of publication: 4 November 2016*

Our population of student and graduate student members is growing rapidly—student members alone are up more than 69.0% since 2014—and figuring out how to serve this demographic has been a learning experience. It is critical to serve this audience and prove the value of SPS membership so as to ignite and retain interest for longtime member loyalty. Research shows that, in professional associations, these groups (categorized mostly as millennials) prioritize networking and job opportunities as member benefits as they begin their careers. That is why the SPS Student Career Luncheon, launched in 2013, is such an important event for our Society. Originally launched at ICASSP, the event connects our younger members with representatives in industry to explore job opportunities in signal processing. Since its inception, it has been attended by hundreds of students and is now expanded to all major SPS conferences, e.g., ICASSP, ICIP, and GlobalSIP. As this demographic continues to rise, we hope to further broaden our efforts to include virtual components that will facilitate networking and job opportunities to a larger audience who may not have access or resources to attend conferences in-person.

Once these younger members begin to network and are hired for these sought-after jobs, we must continue to foster their professional growth and make sure the Society remains relevant to their successes. Studies suggest members in the middle of their careers are

focused on establishing themselves professionally, prioritizing professional development, technical and industry information, and continuing education. To that end, in June of this year, we launched the highly anticipated SPS Resource Center. Previously known as SigView, the SPS Resource Center (<http://rc.signalprocessingsociety.org>) is our online hub of tutorials, conference recordings, educational content, and more, all offered at varying price points for IEEE Members and non members and SPS members. It is our goal that, as the resource center expands, it will become an all-purpose signal processing education powerhouse with webinars and comprehensive continuing education offerings to benefit our members. Paired with our conferences, workshops, Seasonal Schools program, and newly established networking events geared toward young professionals, we're looking ahead to provide a multifaceted professional development experience for members across varying demographics.

As members' careers begin to wind down and eventually edge into retirement, their needs and wants from SPS membership have evolved from what they were during their first years with the Society. Mid-to-late- and postcareer members want to remain active and involved in SPS and enhance their knowledge about signal processing but aren't necessarily seeking the same benefits tailored for newcomers. They turn to SPS for more networking, valuable

industry information and research, as well as recognition through our awards. Their resumes are already built, and they already know the value of SPS membership and its impact on their careers. As a result, many feel compelled to give back to the Society in volunteer positions, acting as ambassadors who serve for the benefit of future generations of members. This loyal demographic is the one that must be built and nurtured from the early stages, setting examples as SPS success stories.

Despite our successes, we recognize areas for growth and improvement for our diverse membership within all of these stages. How do we make our Society—and, thus, our field—more inclusive for women, and how do we promote and support their participation? What are industry members seeking out of SPS membership? How can we embrace and engage the general public in meaning-

ful discourse about the impact of signal processing? How do we not only recognize diversity but implement services and practices to encourage it? We have taken small steps—we've expanded the Women in Signal Processing Luncheon to take place at all major SPS conferences and have been proactive in an external outreach to increase public visibility and awareness of signal processing fields. With our collective knowledge, impact, and reach, we really have the power to better serve all of these groups and change the world together.

There is no “one-size-fits-all” approach to membership in any professional association, and SPS volunteer leadership and staff work daily to develop and implement programs and tools to benefit and serve our diverse member base from the onset of their careers to the end. Generally, we have relied on more traditional

avenues, such as high-ranking publications and conferences around the world, to serve our members. As a result, we've established a global community, with individuals acting as resources for one another and to the Society. Our role is evolving, though, and we look forward to building the SPS member experience to be an all-inclusive resource for meaningful lifelong success for members, for the Society, and for humanity.

We encourage you to let us know how we can better serve you. If you have ideas, suggestions, or would like to help us identify solutions to best provide for all members, please contact me at rababw@ece.ubc.ca.






Fourth IEEE Global Conference on Signal and Information Processing
December 7–9, 2016, Greater Washington D.C., USA




<http://2016.ieeeglobalsip.org/>

Join Us at GlobalSIP 2016 – Featuring 14 Symposia, 24 Keynotes & Panels

- General symposium
- Cognitive communications and Radar
- Compressed sensing and deep learning
- Big data analytics in medical imaging
- Signal and info. processing over networks
- Signal processing for understanding crowd dynamics
- Distributed information processing, optimization, and resource management
- Signal and information processing for smart grid infrastructure
- Transceivers and signal processing for 5G wireless and mm-wave systems
- Non-commutative theory and applications
- Information theoretic approaches to security and privacy
- Sparse signal processing for communications
- Signal processing of big data
- Emerging signal processing applications





Digital Object Identifier 10.1109/MSP.2016.2621061

Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels

Spencer, Q.H.; Swindlehurst, A.L.; Haardt, M.

This paper presents two closed-form constrained solutions to the capacity problem and the power control problem. The first solution, block diagonalization, is a generalization of channel inversion when there are multiple antennas at each receiver. The second, successive optimization, is a method for solving the power minimization problem one user at a time.

February 2004

Bayesian Compressive Sensing

J. Shihao; Xue Y.; Carin, L.

In this paper, a Bayesian formalism is employed for estimating the underlying signal f based on compressive-sensing measurements g . The proposed framework has the following properties: 1) in addition to estimating the underlying signal f , “error bars” are also estimated, these giving a measure of confidence in the inverted signal; 2) using knowledge of the error bars is used to determine when a sufficient number of compressive-sensing measurements have been performed; 3) the compressive sensing measurements are optimized adaptively and, hence, not determined randomly; and 4) the framework accounts for additive noise in the compressive-sensing measurements and provides an estimate of the noise variance.

June 2008

Matching Pursuits with Time-Frequency Dictionaries

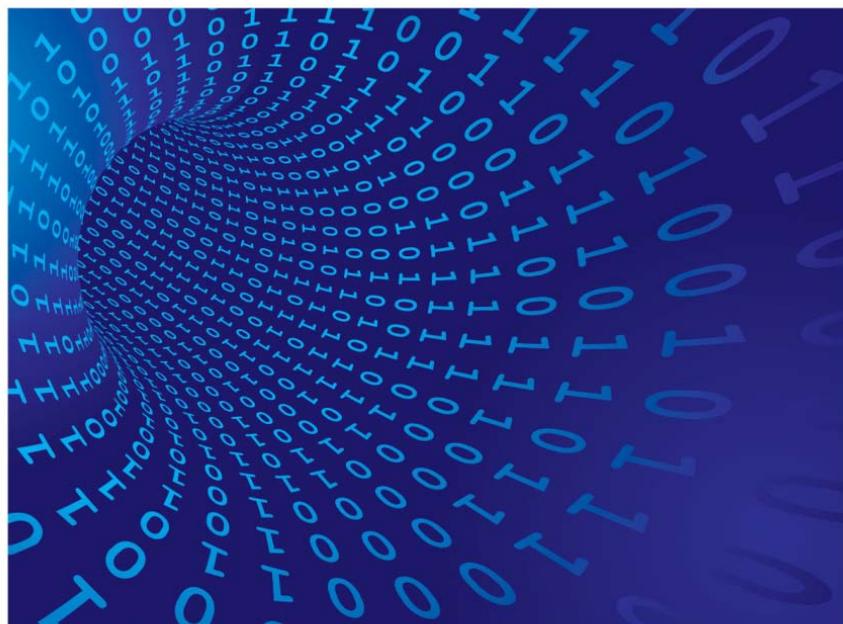
S.G. Mallat; Zhang, Z.

The authors introduce an algorithm, called *matching pursuit*, that decomposes any signal into a linear expansion of waveforms that are selected from a redundant dictionary of functions. This is a truly classic paper cited by thousands over the years that followed.

December 1993

A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers

Morgan, D.R.; Ma, Z.; J. Kim; M.G. Zierdt; J. Pastalan



©GRAPHIC STOCK

In this paper, the authors compare a predistortion model to classical methods of avoiding “spectral regrowth.” A new generalized memory polynomial that achieves the best performance to date is demonstrated with experimental results obtained from a testbed using an actual 30-W, 2-GHz power amplifier.

October 2006

On the Efficiency of Far-Field Wireless Power Transfer

Xia, M.; Aissa, S.

This paper investigates the power transfer efficiency of the wireless power transfer segment in future communication systems in support of simultaneous power and data transfer, by means of analytically computing the time-average output direct current (dc) power at user equipment (UE). If opportunistic scheduling is performed among N symmetric/asymmetric UE, the power scaling laws are attained by using extreme value theory and reveal that the gain in power transfer efficiency is $\ln N$ if UEs are symmetric whereas the gain is N if UEs are asymmetric, compared with that of conventional round-robin scheduling.

March 2015

Structured Compressed Sensing: From Theory to Applications

Duarte, M.F.; Eldar, Y.C.

In the overview, the theme of this paper is exploiting signal and measurement structure in compressive sensing (CS). The primary focus is bridging theory and practice; that is, to pinpoint the potential of structured CS strategies to emerge from the math to the hardware. The summary highlights new directions as well as relations to more traditional CS, with the hope of serving both as a review to practitioners wanting to join this emerging field, and as a reference for researchers that attempts to put some of the existing ideas in perspective of practical applications.

July 2011

Improving Wireless Physical Layer Security via Cooperating Relays

Dong, L.; Han, Z.; Petropulu, A.P.; Poor, H.V.

This paper addresses secure communications of one source-destination pair with the help of multiple cooperating relays in the presence of one or more eavesdroppers. Three cooperative schemes are considered: decode-and-forward (DF), amplify-and-forward (AF), and cooperative jamming (CJ).

December 2009

SP

Signal Processing at the Heart of a Health-Care Renaissance

As people live longer, new technologies promise better and less costly diagnostic services

Global life expectancy is rising thanks, in no small measure, to advancements in medical diagnostic and treatment technologies. According to the World Health Organization (WHO), 71.4 years was the average life expectancy at birth of the global population in 2015. There is reason to believe that this number will climb even higher in the years ahead as living standards increase, access to medical services improves, and innovative new health-care technologies become available.

On the opposite side of the coin, longer lifespans pose a serious health-care challenge. As more people survive into old age, the demand for high-quality medical services will almost certainly increase. Yet with health-care costs already climbing rapidly worldwide, governments and other health-care providers will soon face a sobering choice: reduce and/or ration medical services or turn to a new generation of medical technologies that can sustain current service levels at an affordable cost.

In some places, the future is already here. According to 2014 Japanese government estimates, 33.0% of the country's population is above age 60, 25.9% are aged 65 or above and 12.5% are aged 75 or above. Many other nations, particularly in Western Europe, are well on the way to posting similar numbers.

Digital Object Identifier 10.1109/MSP.2016.2598884
Date of publication: 4 November 2016

Three-dimensional skin cell imaging

As the world looks for ways of making health care both better and more affordable, signal processing is playing an important role in the creation of a new generation of promising diagnostic and treatment technologies. In areas such as medical imaging, medical device connectivity, and health monitors, signal processing is playing an essential role in technology development and operation.

At Stanford University, researchers are using signal processing in a technology that creates real-time three-dimensional (3-D) images of individual cells in a living animal. The technique takes viewers deep under the subject's skin, showing molecular-level real-time details of lymph and blood vessels (Figure 1). The molecular imaging and characterization of tissue noninvasively at cellular resolution technique (dubbed *Mozart*, for short) promises to noninvasively allow physicians to detect tumors in places such as the skin, colon, or esophagus, as well as to view the abnormal blood vessels that signal the initial stages of macular degeneration, a leading cause of blindness.

"We developed Mozart over the past several years, and we are continuing to

apply it to study several different models of human disease," says graduate student Elliott SoRelle, a coresearcher on the project with Adam de la Zerda, an assistant professor of structural biology at Stanford. "We've come a long way with this project, including the ground-up development of biocompatible nanoparticle contrast agents and rigorous testing and validation of our image processing methods."

The new imaging approach builds on an existing technique for examining live tissue buried several millimeters under the skin. Yet that technology—optical coherence tomography (OCT)—isn't sensitive or selective enough to resolve images of individual cells. Mozart not only images

single cells but addresses a major challenge facing diagnosticians: differentiating between types of cells or tissues (such as detecting the cancerous cells that are beginning to multiply within an overall healthy tissue).

In conventional microscopy, scientists have access to agents that latch onto and highlight structures of interest and provide a view of where they are within a cell or body. OCT, however, lacks such helpful beacons. The Stanford researchers understood that tiny particles—gold

As the world looks for ways of making health care both better and more affordable, signal processing is playing an important role in the creation of a new generation of promising diagnostic and treatment technologies.

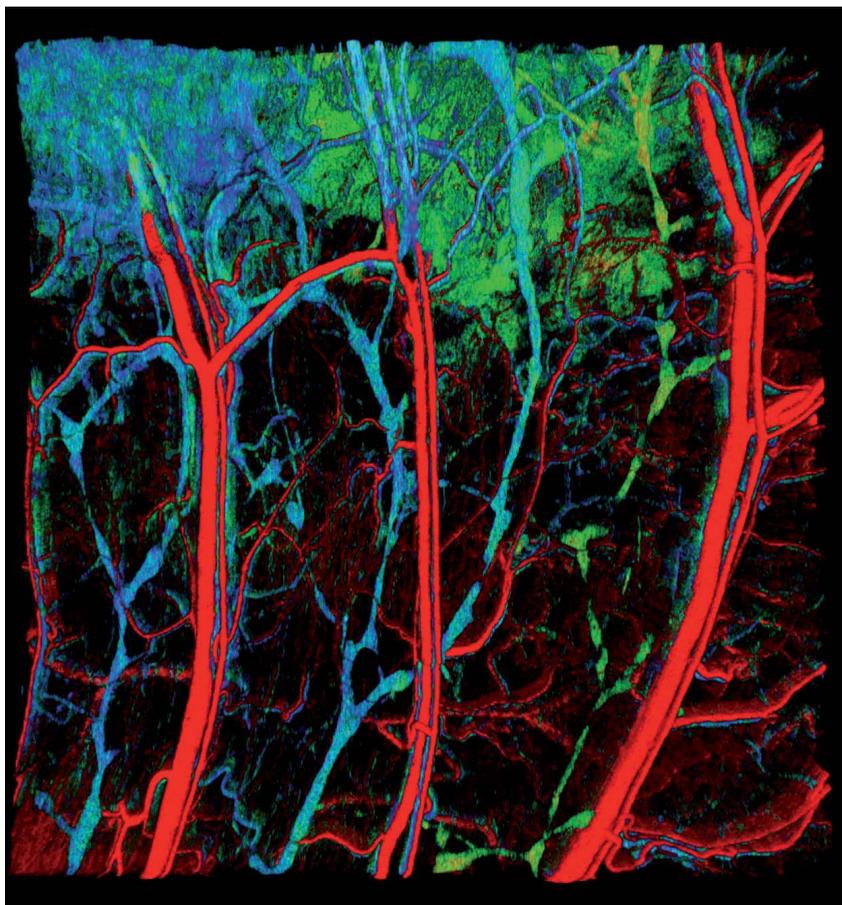


FIGURE 1. Stanford University researchers are using signal processing in a technology that creates real-time 3-D images of individual cells in a living animal. (Image courtesy of Stanford University.)

nanorods—possessed properties that would allow them to serve as visibility-enhancing contrast agents. Yet they also realized that commercially available nanorods didn't produce enough of a signal to be detected within a tissue.

The key to solving this problem, according to SoRelle, was developing a new class of large gold nanorods (LGNRs) that would vibrate at low frequencies within a structure of interest. With surrounding tissues vibrating at higher frequencies, the longer nanorods could perform as highly distinctive contrasting agents.

Turning this idea into reality meant that the researchers would need to find a way of filtering out the LGNRs' frequencies from the surrounding tissue. Addressing this challenge, coresearcher and graduate student Orly Liba developed algorithms specifically designed for this task. "Signal processing is essential to our research," Liba says.

"As a first step, reconstructing the OCT signal from the recorded spectrum to the spatial-domain image required performing a nonequispaced Fourier transform, which we implemented as matrix multiplication," Liba explains. "Next, the spectral analysis, which is used to detect the LGNRs in the tissue, is achieved by filtering the recorded spectrum with two separated Hann filters and converting the two results to the spatial domain separately." The researchers also implemented depth dependent gain to compensate for chromatic aberrations.

"The biggest signal processing challenge," Liba says, "was finding the best algorithm for detecting the tissue-embedded LGNRs, using their distinct spectral characteristics, while maintaining a high-resolution image and high detection sensitivity." "In order to do this, we experimented with several published algorithms and ended up using ideas from

these papers to create our own algorithm, which ended up achieving very good performance," Liba says. The researchers then tweaked the algorithm to make it even more robust to optical aberrations and more sensitive to detecting LGNRs.

Using SoRelle's LGNRs and Liba's algorithms, the researchers solved the challenge of detecting specific structures within 3-D images of living tissues. The team then tested their technology by looking inside the ear of a living mouse. They watched as the LGNRs were absorbed into the lymph system and transported through a network of valves. They were then able to distinguish between two different size LGNRs that resonated at different wavelengths in separate lymph vessels, and could also distinguish between the two nanorods inside the lymph system and the blood vessels.

Mozart combines the high resolution of OCT with molecular and functional information from contrast agents," SoRelle says. "Mozart's excellent spatial resolution and contrast agent detection sensitivity could potentially be used to improve guided surgery or to detect diseases in their earliest stages, which may be clinically advantageous," he adds.

After proving that the LGNRs can be viewed in living tissue, the researchers' next hope is to show that LGNRs can bind to specific kinds of cells, such as skin cancer or abnormal vessels in early stage macular degeneration. Such a technique could be used to learn more about how those diseases progress at the molecular level and to evaluate treatments in individual patients. "Now that we have demonstrated Mozart's capabilities, we are excited by its numerous future applications," SoRelle says.

One of the things that excites SoRelle most about Mozart is its potential for studying fundamental biological processes in vivo. "A tremendous amount of what we know about how cells function and communicate with each other has been discovered and characterized in cultured cells," he says. "Imagine if we could begin to explore those processes in real time in the context of living, breathing animals rather than on a microscope slide."

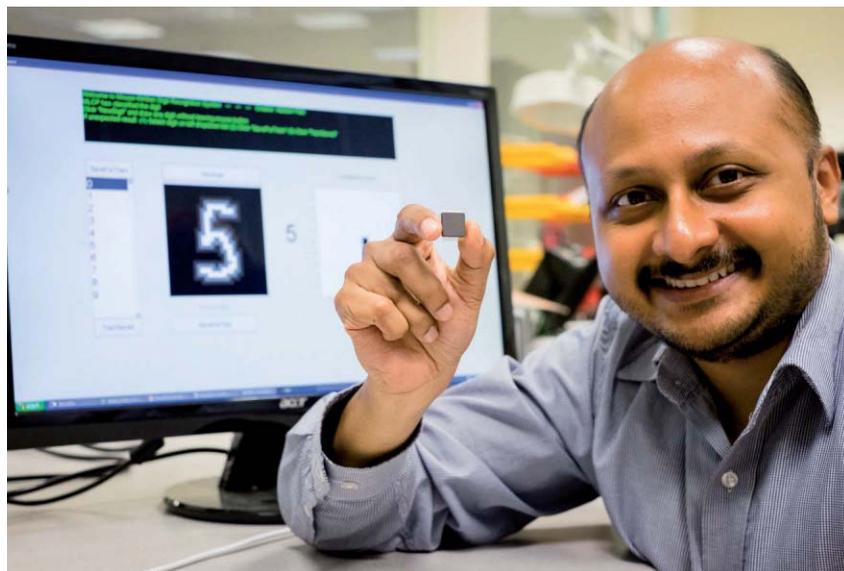


FIGURE 2. Arindam Basu, an assistant professor at NTU Singapore's School of Electrical and Electronic Engineering, holds a low-power chip that connects wirelessly to neural implants. (Image courtesy of NTU Singapore.)

Wireless neural implant connectivity

When embedded in a brain, neural implants can be used to alleviate the debilitating symptoms of Parkinson's disease, give paraplegics the ability to move their prosthetic limbs, or alert epileptics to an impending seizure. Researchers at Singapore's Nanyang Technological University (NTU) recently developed a chip that connects wirelessly to neural implants, enabling more efficient, accurate transmission of brain signals (Figure 2).

For prosthetic patients, a neural implant is typically connected via wires to a computer that decodes the brain signals and enables the artificial limb to move. Yet external wires are not only cumbersome but require permanent openings into the brain, increasing the risk of infections. The new chip is designed to resolve both of these problems. Project researcher Arindam Basu, an assistant professor at NTU's School of Electrical and Electronic Engineering, says that tests on data recorded from animal subjects showed that the chip could decode the brain's signal to the hand and fingers with 95% accuracy.

Raw neural data can easily amount to multiple megabits per second.

Transmitting this level of data reliably will drain an implant's battery within a matter of hours. Designed to be highly power efficient, the chip analyzes and decodes data at the implant site before compressing the results and sending it wirelessly to a small external receiver. "We can reduce the data rate to tens of kilobits per second, increasing the battery life a lot," Basu says.

"The chip consumes ten to 50 times less energy than off-the-shelf processors," Basu says. "We achieved this [capability] by adopting a mixed analog-digital processing approach as opposed to the fully digital processing used in commercial processors," he notes. Most of the processing is performed in analog to reduce energy while a small, but critical, fraction of computing is accomplished in digital to achieve accuracy. "The signal processing algorithm was also chosen carefully to suit this approach—a randomized neural network, like extreme learning machines, was selected where a large bulk of the weights or coefficients are random and fixed, which can be easily implemented in analog without loss of accuracy," Basu says.

The chip is also designed to analyze data patterns and detect any abnormal

mal or unusual patterns. "A pattern recognizer is quite general and can be trained to detect any pattern in a signal that may have diagnostic value," Basu says. "For example, for epileptic patients with medically incurable epilepsy, this chip can be trained to detect the electrical onset of epilepsy before the clinical onset kicks in." The capability can be used to trigger electrical stimulators to suppress the seizure and can also notify a caregiver about an impending seizure. The chip can also be used to track arrhythmia patterns in electrocardiograms, which are indicative of heart abnormalities.

The researchers opted for a two-layer neural network based on "random projection to higher dimension" in a first hidden layer followed by a trainable second layer. "As long as input dimension D is much larger than the number of classes C , the number of multiplications in the first layer ($D \times L$) is much larger than that in the second layer ($C \times L$) where L is the dimension of the hidden layer," Basu says.

"The key observation is that the random number multiplications in the first layer can be implemented in very low energy using analog circuit as opposed to traditionally used digital approaches (~10–100 times lower)," Basu explains. Analog circuits, however, suffer from non-idealities such as mismatches or statistical variations. "This is not a problem in the first stage where the weights are supposed to be random," Basu notes. "In fact, we exploit the mismatch to create the randomness." However, to achieve high accuracies, the researchers kept the second stage processing in the digital domain, enabling it to be separately trained for the particular mismatch distribution of every chip. "Thus by making a careful choice of algorithm and careful partitioning of analog and digital computing, we are able to achieve high accuracy as well as low energy," Basu says. "Project development and testing processes are ongoing," Basu continues.

Sweat monitor

Many medical professionals view wearable sensors as an affordable and highly

effective way of monitoring personal health. Continuously sampling human sweat, which is rich in physiological information, promises a convenient, informative, and noninvasive monitoring mechanism.

Existing noninvasive sweat biosensors either monitor only a single molecule at a time or lack the signal processing necessary to compensate for temperature effects or interactions among different molecules. But a new wearable monitoring device, developed by University of California, Berkeley, researchers, promises to give its wearers a fully integrated “perspiration analysis system” (Figure 3). Working with counterparts at the Stanford University School of Medicine, the researchers developed a monitor that binds to the skin and measures certain sweat metabolites and electrolytes. The device can also calibrate its readings based on skin temperature.

Sweat secretion is a complex process, notes project researcher Wei Gao, a Berkeley postdoctoral fellow. “Our platform is a powerful tool that can significantly advance large-scale and real-time physiological and clinical studies by facilitating the identification of informative biomarkers in sweat,” he says. Data read by the device can be used to alert wearers and caregivers to specific health concerns, such as dehydration, fatigue, and dangerously high body temperatures.

Gao feels that the new technology effectively spans the gap between signal transduction, signal processing, and wireless transmission in wearable biosensors by merging plastic-based skin sensors with silicon-integrated circuits consolidated onto a flexible circuit board for complex signal processing.

The new technology effectively spans the gap between signal transduction, signal processing, and wireless transmission in wearable biosensors by merging plastic-based skin sensors with silicon-integrated circuits consolidated onto a flexible circuit board for complex signal processing.



FIGURE 3. A new wearable monitoring device, developed by the University of California, Berkeley, researchers, promises to give its wearers a fully integrated “perspiration analysis system.” (Image courtesy of University of California, Berkeley.)

(Na and K ions), and skin temperature,” Gao says.

According to Gao, the measurement of Na and K levels is facilitated through the use of ion selective electrodes. The potential output of these sensors is linearly related to the logarithm of the concentration of target electrolytes. The resistance-based temperature sensor is based on metallic microwires; higher temperatures result in higher resistance. “Our device can real-time measure the electrical signals generated from all the sensors,” Gao says.

The device measures the electrical signals generated by all of the sensors simultaneously. “Signal processing is important to average the data and filter out noise and interference in the measurements,” Gao says. “More importantly, it allows us to do the real-time compensation and calibration to ensure

the accurate reading of the sensors,” he explains.

“By properly choosing different types of signal processing and introducing real-time system calibration, such as temperature compensation, our device can provide very stable readings,” Gao says. “In addition, the independent and selective operation of individual sensors is preserved by electrically decoupling the operating points of each sensor’s interface.”

Gao says that the project’s next step is performing population-level studies. “Sweat contains physiologically rich information about what’s happening in the body, with many different substances, from electrolyte ions and metabolites to protein molecules,” Gao notes. “The same platform can be exploited or reconfigured for in-situ analyses of other biomarkers within sweat and other human fluid samples to facilitate personalized and real-time physiological and clinical investigations.”

Author

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

SP

FROM THE GUEST EDITORS

John H.L. Hansen, Kazuya Takeda, Sanjeev M. Naik,
Mohan M. Trivedi, Gerhard U. Schmidt,
and Yingying (Jennifer) Chen

Signal Processing for Smart Vehicle Technologies

The invention of the automobile has transformed how people live, work, and interact in society. Today, with an ever-increasing number of in-vehicle options/activities, as well as the increasing demands being placed on the driver, vehicle platform, and transportation infrastructure, more is being asked of engineers, designers, scientists, and transportation specialists. Signal processing is playing an increasingly substantial role in this domain, including such general topics as monitoring driver distraction, vehicle lane/control detection/tracking, driver assistance through autonomous platforms, and vehicle infrastructure support and planning/monitoring. The diversity

of these problems requires a more collaborative effort from engineers and scientists in a diverse set of specialties. The impact to society is massive, including such broad aspects as 1) safety, 2) commerce (i.e., sales and support/maintenance of vehicles), 3) energy costs (i.e., fossil fuel consumption, etc.), and 4) population mobility for effective traffic management. How will signal processing advance today's vehicles into "smart" cars that are able to think and contribute to the task of operating a

vehicle? What safety concerns are there in moving from a 100% driver-controlled vehicle, to driver assistive technologies (e.g., cruise control, assistive braking, lane-departure monitoring, etc.), to full autonomous driving? Many new and emerging challenges arise and need to be addressed in collaborative ways.

This special issue provides a venue for summarizing, educating, and sharing the state of the art in signal processing applied to the domain of automotive systems. Due to the significance of this topic

from both an engineering/technology as well as a global society perspective, this special issue of *IEEE Signal Processing Magazine* will appear in two parts (part 1 is the current issue, and

part 2 is scheduled to be published in the spring of 2017). Highlighted below is the scope of topics addressed in varying degrees by the articles that are explored in both parts:

- digital signal processing technologies in adaptive automobiles, diagnosis, and maintenance
- speech, hands-free, and in-car communication algorithms and evaluation
- in-vehicle dialog systems and human-machine interfaces
- driver-status monitoring and distraction/stress detection
- computer vision methods for vehicle recognition and assisted driving

- multisensor fusion for driver identification and robust driver monitoring
- signal processing for position and velocity estimation and control
- signal processing for green vehicle-related energy management
- vehicle-to-vehicle and vehicle-to-infrastructure communications and networking
- autonomous, semiautonomous, and networked vehicular control
- human factors and cognitive science in enhancing vehicle and driver safety
- machine learning and data analytics associated with automotive systems
- issues regarding security and privacy aspects for smart vehicle systems.

In planning this special issue, we worked extensively to ensure a wide representation of the field. A large number of white papers were received, and the authors of a select set of white papers were invited to submit full papers that were then peer reviewed.

Six articles appearing in the current issue span a broad range of signal processing for vehicle systems. The first group contains three articles that address driver behavior and monitoring: "Driver-Behavior Modeling Using On-Road Driving Data," by Miyajima and Takeda, "Driver Status Monitoring Systems for Smart Vehicles Using Physiological Sensors" by Choi et al., and "Smart Driver Monitoring: When Signal Processing Meets Human Factors" by Aghaei et al. Next, Weng et al.'s article, "Conversational In-Vehicle Dialog

This special issue provides a venue for summarizing, educating, and sharing the state of the art in signal processing.

Systems,” explores past, present, and future trends. Since speech interaction and audio are underutilized modalities for driver interaction, this represents an important emerging trend in the field. Samarasinghe et al. focus on advancements in active noise control within car environments in their article “Recent Advances in Active Noise Control Inside Automobile Cabins.” Finally, Hult et al.’s article, “Coordination of Cooperative Autonomous Vehicles,” focuses on the ability to effectively coordinate autonomous vehicles within the transportation system.

We would like to encourage readers to explore these articles, as well as the field of signal processing for vehicle technologies, since the prospects for growth and impact on safety, legal, and social aspects are enormous. Finally, from a purely cognitive standpoint, we ask that all drivers be aware of the impact of cognitive load in employing any technologies during their driving tasks (e.g., please, no texting while driving). We look forward to bringing you the next installment of this special issue in the spring of 2017. Happy reading (and driving)!

Guest Editors



John H.L. Hansen, (john.hansen@utdallas.edu) is the associate dean for research and a professor of electrical engineering at the University of Texas at Dallas. He oversees the Center for Robust Speech Systems (CRSS)-UTDrive Lab and Speech-Speaker-Language Processing Lab. He has served as a technical program chair for the IEEE International Conference on Acoustics, Speech, and Signal Processing 2010, general chair for the International Speech Communication Association (ISCA) INTERSPEECH 2002, and organizer/contributor for the Biennial Workshop DSP for In-Vehicle Systems and Safety (2003–2015). He is an IEEE Fellow, ISCA fellow, and

has supervised 75 Ph.D./M.S. students and coauthored more than 600 papers in the fields of digital signal processing, speech processing, and driver distraction modeling/detection.



Kazuya Takeda (kazuya.takeda@nagoya-u.jp) is a professor at the Graduate School of Informatics and Green Mobility Collaborative Research Center, Nagoya University, Japan. His interest is understanding situated human behavior such as driving in terms of signal processing.



Sanjeev M. Naik (sanjeev.m.naik@gm.com) received the B.Tech., M.S.E.E., and Ph.D. degrees from IIT Bombay, University of Michigan, Ann Arbor, and the University of Illinois, Urbana–Champaign, respectively, all in electrical engineering, and an M.B.A. degree from the University of Michigan. He was employed by the Cummins Engine Company and joined General Motors in 1994, where he is the engineering group manager in advanced controls engineering, having led multiple technical teams developing control systems for propulsion, active safety, and electrification applications. He is an IEEE Senior Member and a member of SAE International. His research interests are in the application of advanced controls and signal processing techniques to develop safe, clean, and efficient fun-to-drive vehicles.



Mohan M. Trivedi (mtrivedi@ucsd.edu) is a Distinguished Professor of Electrical and Computer Engineering and the founding director of the Computer Vision and Robotics Research Laboratory and Laboratory for Intelligent and Safe Automobiles at the

University of California, San Diego. He serves as a senior editor of *IEEE Transactions on Intelligent Vehicles* and as a consultant to industry and government agencies in the United States and abroad. He is a Fellow of the IEEE, IAPR, and SPIE.



Gerhard U. Schmidt (gus@tf.uni-kiel.de) received the Dipl.-Ing. degree in 1996 and the Dr.-Ing. degree in 2001, both from the Darmstadt University of Technology, Germany. He has worked with the research groups of the Acoustic Signal Processing Department, Harman/Becker Automotive Systems, and at SVOX, Ulm, Germany. Parallel to his time at SVOX, he was a part-time professor with the Darmstadt University of Technology. Since 2010, he has been a full professor with Kiel University, Germany. His main research interests include adaptive methods for audio and medical signal processing.



Yingying (Jennifer) Chen (yingying.chen@stevens.edu) is a professor of electrical and computer engineering at Stevens Institute of Technology in Hoboken, New Jersey. Her research interests include cybersecurity and privacy, mobile and pervasive computing, and mobile health care. She received her Ph.D. degree in computer science from Rutgers University and has published more than 100 journal and conference papers. She has been recognized with numerous best paper awards at conferences and received the National Science Foundation CAREER Award and Google Faculty Research Award. She presently serves on the editorial boards of *IEEE Transactions on Mobile Computing*, *IEEE Transactions on Wireless Communications*, and *IEEE Network Magazine*.

SP

Chiyomi Miyajima and Kazuya Takeda

Driver-Behavior Modeling Using On-Road Driving Data

A new application for behavior signal processing

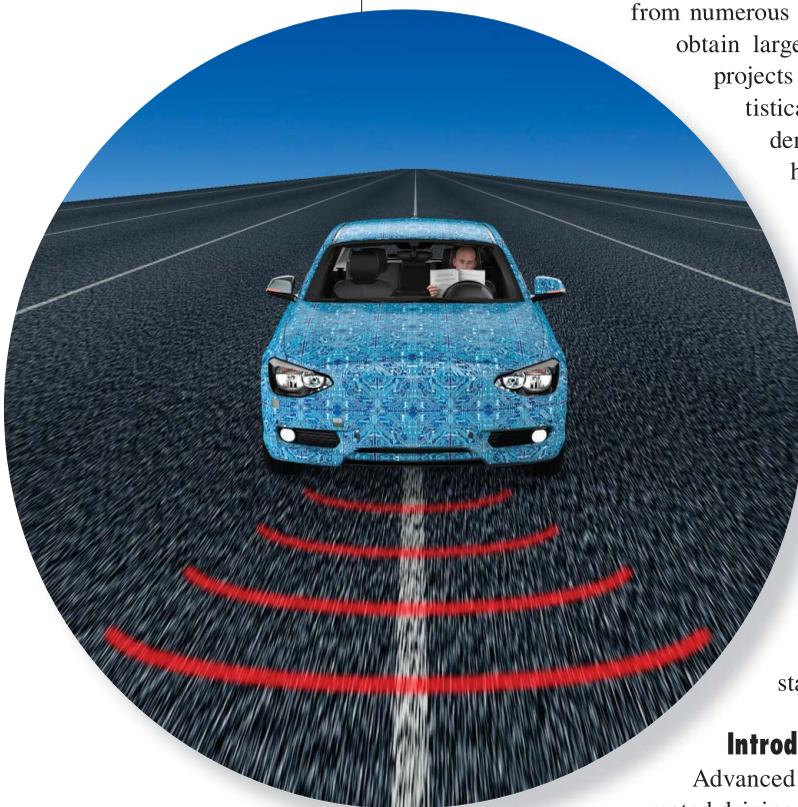
This article reviews data-centric approaches for statistical modeling of driver behavior. Modeling driver behavior is challenging due to its stochastic nature and the high degree of inter- and intradriver variability. One way to deal with the highly variable nature of driving behavior is to employ a data-centric approach that models driver behavior using large amounts of driving data collected from numerous drivers in a variety of traffic conditions. To obtain large amounts of realistic driving data, several projects have collected real-world driving data. Statistical machine-learning techniques, such as hidden Markov models (HMMs) and deep learning, have been successfully applied to model driver behavior using large amounts of driving data. We have also collected on-road data recording hundreds of drivers over more than 15 years. We have applied statistical signal processing and machine-learning techniques to this data to model various aspects of driver behavior, e.g., driver pedal-operation, car-following, and lane-change behaviors for predicting driver behavior and detecting risky driver behavior and driver frustration. By reviewing related studies and providing concrete examples of our own research, this article is intended to illustrate the usefulness of such data-centric approaches for statistical driver-behavior modeling.

Introduction

Advanced driver-assistance systems (ADASs) and automated driving systems have been under active development for many years. Such systems should provide support for drivers in a way that is in harmony with their own characteristic driver behavior. To develop assistance systems that are comfortable for humans to use, it is necessary to understand real human driving behavior by analyzing and modeling how human beings interact with their vehicles and the driving environment.

With recent advances in automotive electronics and sensor and communication technologies, we are now better able to collect driving data representing vehicle

SKY IMAGE LICENSED BY GRAPHIC STOCK
©ISTOCKPHOTO.COM/POSTERIORI



Digital Object Identifier 10.1109/MSP.2016.2602377
Date of publication: 4 November 2016

motion, driver behavior, and surrounding environmental factors during real-world driving. To collect a reasonable amount of realistic driving data for research purposes, many data-collection projects have been conducted around the world, including naturalistic driving studies (NDSs) and field operational tests (FOTs) [1]–[9].

We have also been conducting on-road driving-data collection since 1999 [10]. In addition to our own data-collection efforts, we have shared international driving data in collaboration with researchers in the United States and Europe [11] and have applied statistical signal processing and machine-learning techniques to the recorded driving data to analyze and model human driving behavior. Over the course of our research, we first applied cepstral analysis and Gaussian mixture models (GMMs) for modeling driver pedal-operation behavior [12]. We extended our GMM-based behavior-modeling framework to predict car-following behavior [13]–[15] and to detect risky driving behavior for a driver self-coaching system [16]. We then applied HMMs for modeling vehicle trajectories during lane-change maneuvers [17], and for integrated modeling of driver gaze and vehicle-operation behavior during lane changes [18]. We also employed a Bayesian network (BN) to detect driver frustration [19]. Other statistical frameworks were also used for detecting potential threats [20] and driver over-reliance on automated driver-assistance systems [21], [22].

Collection of on-road driving data

The first step that is necessary to analyze and model driver behavior is to collect a reasonable amount of realistic driving data. Table 1 shows examples of several projects around the world that have collected real-world or naturalistic driving data. The 100-car study is a landmark NDS that collected approximately 43,000 hours of driving data from 100 cars and 241 drivers [1]. This data has helped researchers understand the causes of crashes and allowed them to develop effective crash prevention measures [2]. The second Strategic Highway Research Program (SHRP2) also collected a huge amount of driving data from more than 3,000 drivers [3]. Satzoda et al. used the SHRP2 data to evaluate a technique for detecting critical events related to lanes and road boundaries [4]. The European Field Operational Test (euroFOT) collected data from more than 1,000 drivers to evaluate ADASs such as adaptive cruise control and lane departure warning systems [5]. UDRIVE, the first large-scale European NDS project, is also recording naturalistic driving data [6]. In addition, there are other research projects collecting large-scale corpora of real-world or naturalistic driving data [7]–[11]. Some of these data sets are publicly available for a fee, e.g., [8] and [9]. An EU-research project, FOT-Net Data [23], assists researchers in sharing data sets by providing information of available data sets and organizing meetings on sharing and reuse of FOT/NDS data.

One way to deal with the highly variable nature of driving behavior is to employ a data-centric approach that models driver behavior using large amounts of driving data collected from numerous drivers in a variety of traffic conditions.

Since 1999, Nagoya University has been collecting on-road driving behavior data using instrumented vehicles (Figure 1) [10], [11], which have included video, audio, vehicle sensor, and physiological data, from more than 800 drivers using these vehicles.

We have also developed a driving-data acquisition system using a smartphone and the vehicle's CAN-bus [24]. It transmits driving data obtained from the CAN-bus via a Bluetooth connection to the smartphone mounted on the windshield, and records the driving signals from the CAN-bus along with driving data obtained from the smartphone itself. Although this system cannot record the same range of driving data with the same quality as the instrumented vehicles, it allows us to easily collect large amounts of driving data from a much wider range of vehicles. So far, we have recorded driving data from more than 50 drivers operating different models of vehicles using this data acquisition system.

Statistical analysis and modeling of driver behavior

Statistical machine-learning techniques have been successfully applied to data-driven driver-behavior modeling. For example, Sakaguchi et al. modeled driver stopping behavior using a BN and successfully applied this model to detect deviant braking behavior [25]. Oliver et al. presented their driver-behavior recognition and prediction method based on dynamic graphical model, HMMs, and their extensions (coupled-HMM) trained using on-road driving data collected from

Table 1. Examples of on-road driving-data corpora.

Project (Period of Data Collection)	Area	Type of Vehicles	Number of Drivers or Vehicles
100-Car [1] (2003–2004)	United States	Passenger cars	241 drivers (100 vehicles)
SHRP2 [3] (2010–2013)	United States	Passenger cars	3,150 drivers
euroFOT [5] (2010–2011)	Europe	Passenger cars, trucks	1,100 drivers
UDRIVE [6] (2014–2016)	Europe	Passenger cars, trucks, powered two-wheelers	290 drivers
400-car [7] (2014)	Australia	Passenger cars	400 drivers
JSAE/TUAT [8] (2005–present)	Japan	Taxis	200 taxis
HQL [9], [25] (2001–2003)	Japan	Instrumented vehicles	97 drivers
NUDrive [10], [11] (1999–2010)	Japan	Instrumented vehicles	800 drivers

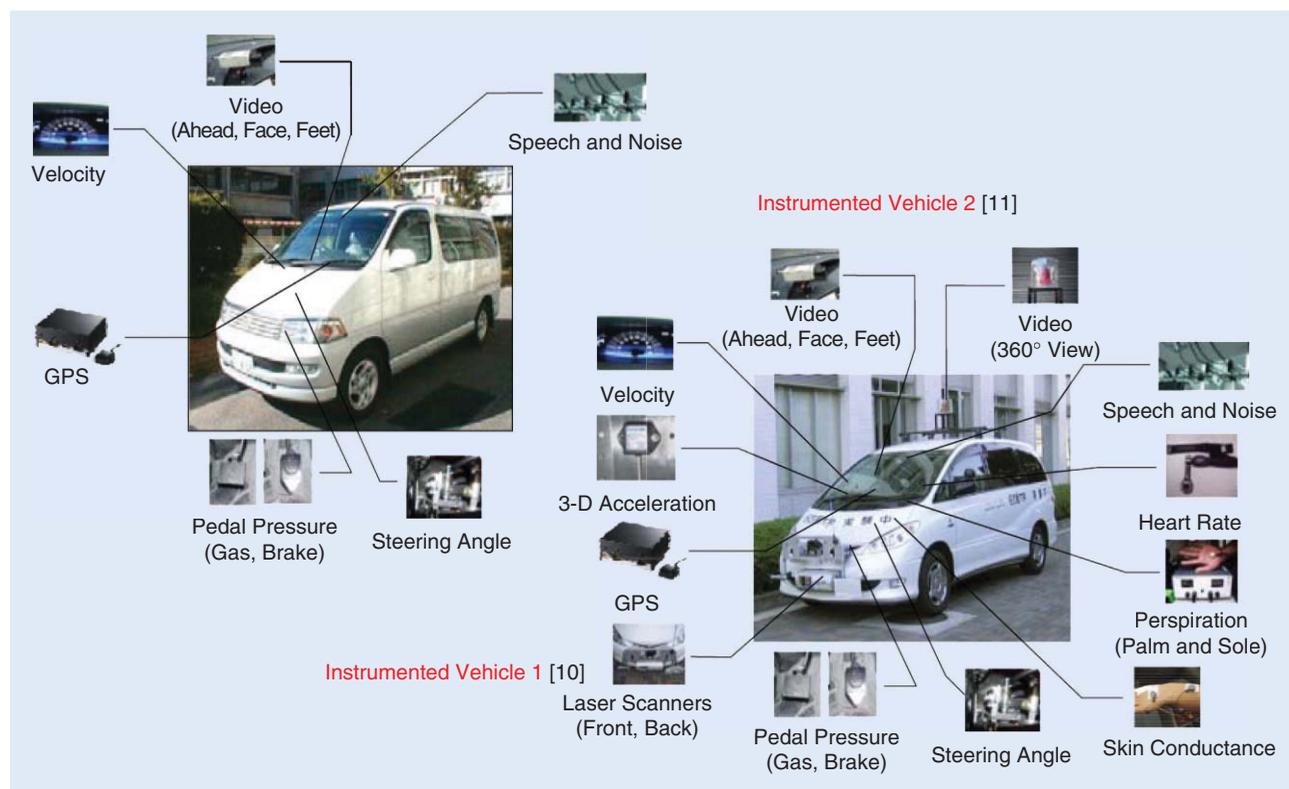


FIGURE 1. Instrumented vehicles used for on-road data collection.

70 drivers [26]. Pongtep et al. used GMMs to model car-following behavior, and demonstrated that GMM-based driver models can be easily adapted to individual drivers using adaptation techniques [14], [15]. Okuda et al. applied hybrid dynamical systems (HDSs) to model driver decision-making behavior during car following [27]. They showed that model parameters can be easily adapted online. Jain et al. applied a recurrent neural network (RNN) to predict driving behavior [28]. The RNN is a type of deep learning technique, and was used to integrate features of driver facial orientation, vehicle status, and road structure.

They compared other kinds of models based on support vector machines (SVMs), random forests, and HMMs, and showed that the RNN model achieved better prediction performance than the other models.

However, it is still difficult to utilize a petabyte-scale driving-data corpus effectively for further research due to the size and heterogeneity of the existing corpora. Taniguchi et al. [29] proposed a nonparametric method called a *double articulation analyzer (DAA)* for the symbolization of driving data. DAA is a hierarchical, bottom-up, data symbolization approach consisting of the nonparametric extension of an HMM, called a sticky *hierarchical Dirichlet process HMM (HDP-HMM)*, and an N-gram

The usefulness of the data-centric approach was demonstrated by reviewing the related studies, which analyzed and modeled driver behavior by applying statistical signal processing and machine-learning techniques to large amounts of driving data.

language model based on the Pitman-Yor process, called a *nested Pitman-Yor language model (NPYLM)*. DAA is an efficient method for compressing driving-data information when dealing with large amounts of data. Taniguchi et al. successfully applied the DAA to compress driving-data information, as well as to predict driving behavior and classify driving situations [29], [30]. The authors have also used a DAA framework to detect risky lane change behavior [31].

We have applied a variety of statistical signal processing methods in our own research to analyze and model various aspects of human driving behavior; we now review our own data-centric driver-modeling research [12]–[22].

Modeling driver pedal-operation behavior using cepstral analysis and GMMs

Figure 2 shows a comparison of the gas-pedal operation patterns of three different drivers while driving three different types of vehicles (Vehicle A: midsize sedan, Vehicle B: hatchback EV, and Vehicle C: minivan). Each driver in the study drove each vehicle. The graphs show 20-second samples of gas pedal position signals while they were driving at stable speeds of around 60 km/h. Examination of the recorded driving signals reveals that clear, individual patterns of

pedal-operation behavior can be observed. For example, Driver 1 operated the pedals more smoothly and maintained the same pedal positions longer, Driver 2 adjusted pedal positions frequently, while the pedal position patterns of Driver 3 exhibited rectangular pulses. These drivers maintained consistent pedal-operation patterns, regardless of the vehicle being driven.

To model such driver's individual characteristics of pedal-operation behavior, we applied cepstral analysis and GMMs to the recorded pedal-operation signals [12]. We assumed that the impulse signal of a driver's intention to depress the gas or brake pedal is filtered by driver characteristics, which can be represented as a spectral envelope. The output of this system is observed as a pedal-operation signal. GMMs were used to represent distributions of the spectral characteristics of pedal operation, extracted through cepstral analysis of the raw pedal-operation signals. The resulting driver-behavior model was evaluated through driver identification experiments using on-road driving data collected with an instrumented vehicle, as described previously [10]. Experimental results showed that the GMM-based behavior model achieved a driver identification rate of

76.8% for 276 drivers when using cepstral features extracted from pedal-operation signals, compared to 47.5% when raw pedal-operation signals were used.

Modeling car-following behavior using GMMs

We extended our GMM-based behavior modeling framework to predict car-following behavior [13]–[16]. This enabled us to generate car-following behavior patterns for individual drivers, allowing for the personalization of headway control, which represents a driver's preferred car-following distance. Gas- and brake-pedal operation patterns of the target driver were modeled using GMMs as a joint density with other driving signals. Gas- and brake-pedal operation signals while car following were generated from a joint probability distribution of the features modeled with GMMs so that conditional probability was maximized for the other observed driving signals. We confirmed that the GMM-based driver models could replicate driving signals representing the individual car-following patterns of different drivers [13]. We then proposed a driver-model adaptation scheme that allowed us to enhance the model's capability to represent the particular driving characteristics of individual drivers [14], [15].

To model such driver's individual characteristics of pedal-operation behavior, we applied cepstral analysis and GMMs to the recorded pedal-operation signals.

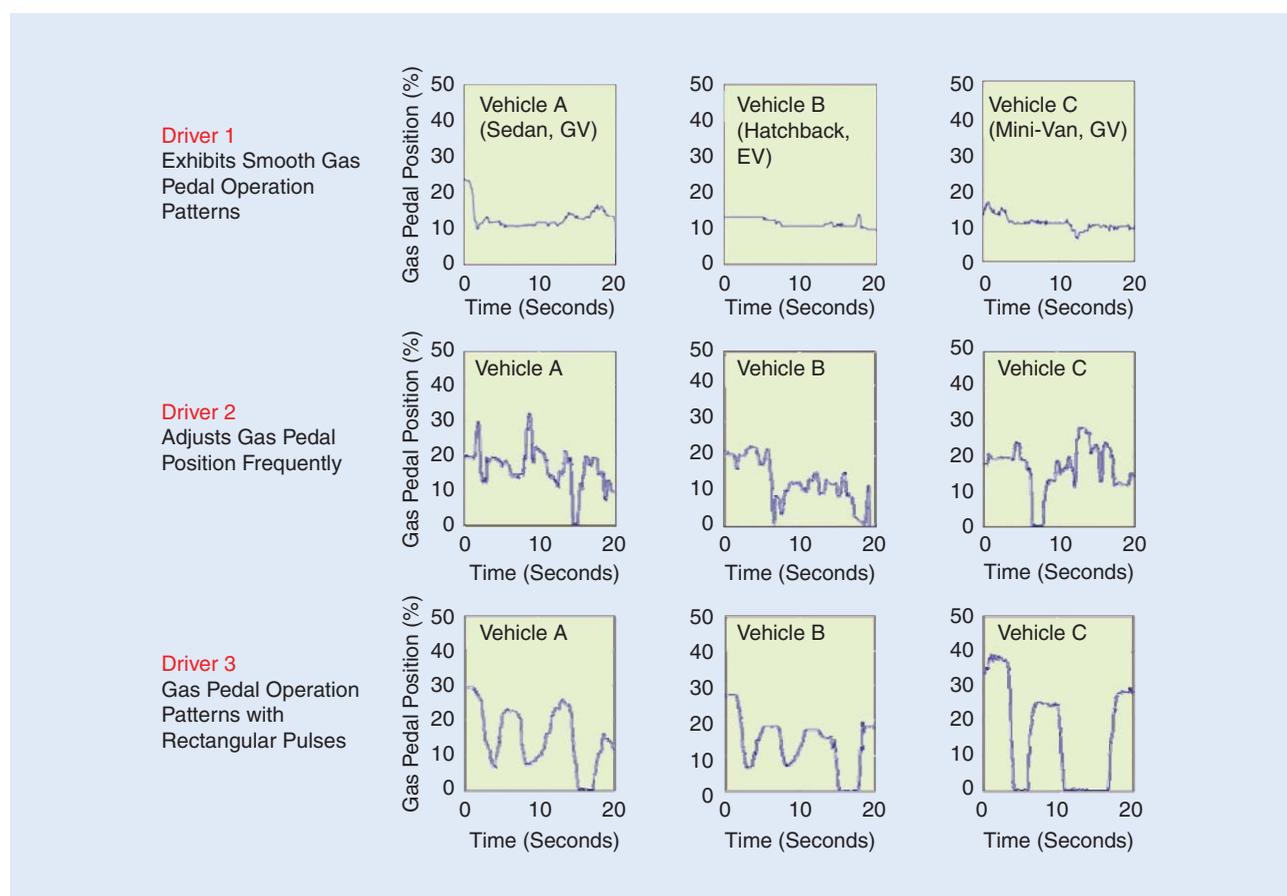


FIGURE 2. Individual differences in gas-pedal operation patterns.

We also used a GMM-based behavior-prediction model for our driver self-coaching system [16]. We calculated a prediction residual representing the degree to which observed driver-behavior signals deviated from the predicted value of the driver model. We used the prediction residual as a feature of abnormal driving, and then used it for detecting risky driving behavior. Drivers received feedback from a browser-based self-coaching system which allowed them to watch driving scenes showing their own risky driving behavior. Figure 3(a) and (b) compares the average number of risky driving scenes detected for drivers who did not receive feedback and those who did receive feedback, respectively. They drove the same routes twice (Trials 1 and 2) and drivers with feedback received feedback from the self-coaching system before Trial 2. We found that the number of risky driving scenes detected fell by more than 50% after drivers received feedback.

Modeling lane-change behavior using HMMs

A single-state model (GMM) can be extended into a multiple-state model (HMM). Using this approach, we

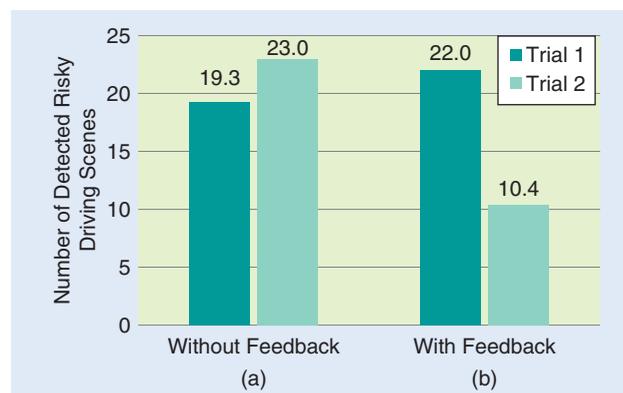


FIGURE 3. A comparison of number of detected risky driving scenes for drivers who (a) did not receive feedback and those who (b) received feedback.

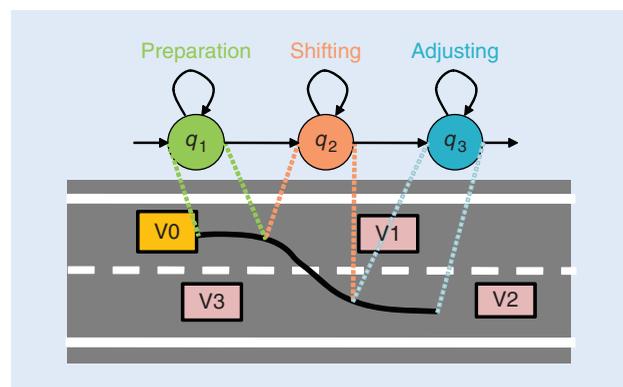


FIGURE 4. Modeling lane-change trajectory using a three-state HMM.

proposed a method of modeling lane change behavior that could predict the trajectory of a driver's vehicle during a lane change [17]. We assumed that a lane change consists of three different stages: preparation, shifting, and adjusting, as shown in Figure 4, and used three-state HMMs to model the lane-change trajectories of different drivers. Our goal was to predict a driver's lane-change trajectory for a period of about 20 seconds, given only the initial driving conditions. We employed a sampling algorithm to generate the most

probable trajectories. Then an optimal trajectory was selected from the generated trajectory candidates using a geometric function representing the driver's cognitive characteristics. Figure 5 shows an example of vehicle trajectory generated from the HMM. We confirmed that the model could generate reasonably accurate, personalized lane-change trajectory predictions.

Modeling lane-change behavior using multistream HMMs

Next, the multistate HMM was extended to a multistream model. We used a multistream HMM for the integrated modeling of driver gaze, vehicle-operation behavior, and vehicle motion during lane changes [18]. We first broke each target activity down into discrete acts (e.g., looking straight ahead, looking into the rear view mirror, etc.), and then jointly modeled sequences of these discrete acts using a multistream discrete HMM, as shown in Figure 6. A risky lane-change detection experiment was then conducted using driving data collected on an expressway from 11 drivers driving our instrumented vehicle [11]. Figure 7 compares risky lane-change detection performance (ROC curves) when using only gaze behavior, only driver vehicle-operation behavior (including vehicle motion), and both combined. The experimental results showed that our integrated models could detect risky lane-change behavior better than either gaze-only or vehicle operation behavior-only models.

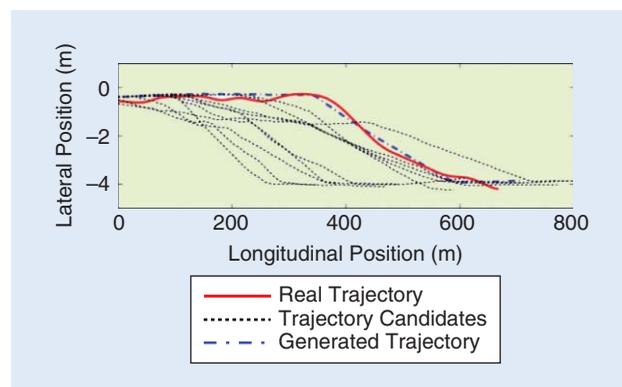


FIGURE 5. Lane-change trajectory generated from HMM.

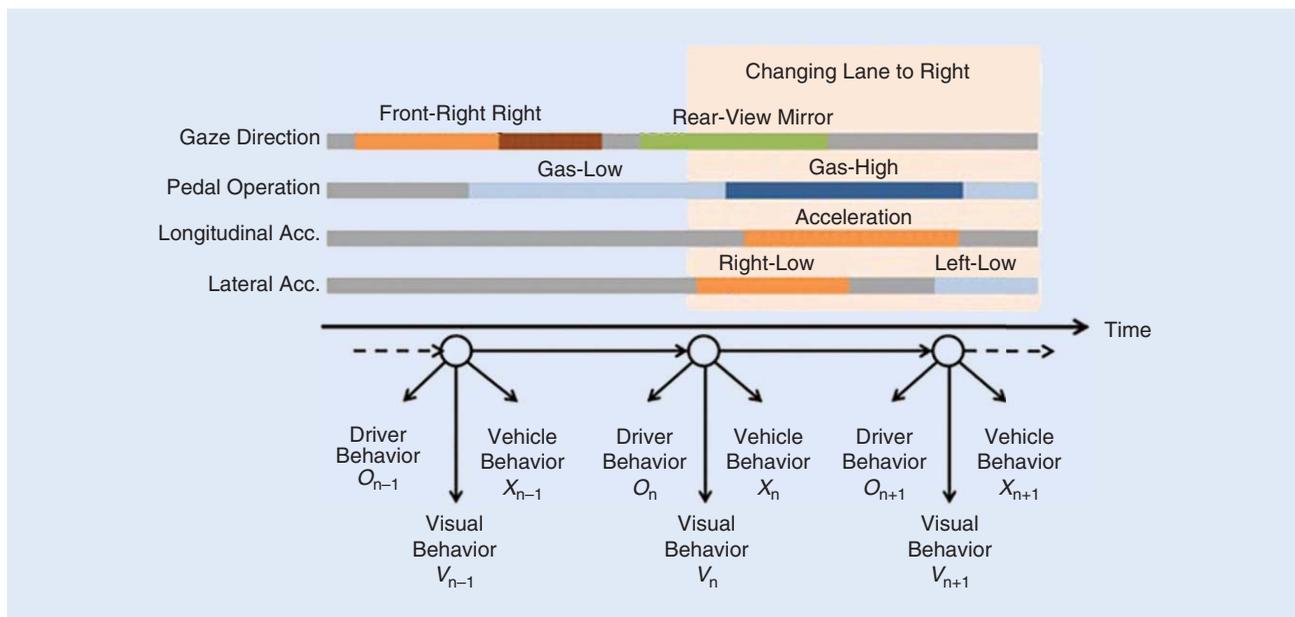


FIGURE 6. Modeling driver lane-change behavior using a multistream HMM.

Modeling driver frustration using BNs

In another study, we proposed a method for modeling driver frustration by integrating its causes and driver reactions, as shown in Figure 8 [19]. For example, a driver may feel frustrated due to congested traffic, parked vehicles blocking the roadway, and a long wait at a red traffic light. As a result, a driver’s physiological state, which is often expressed through facial expressions, can change, affecting the driving behavior. A BN was used to integrate different aspects of these observable features. BNs were then trained for each driver, and driver frustration was detected by calculating a posterior probability of frustration at the middle node of Figure 8 for the given features of the top and the bottom nodes. We conducted a driver frustration detection

experiment with 20 drivers, using driving data collected on city streets [11]. Figure 9(a) compares actual driver frustration reported by drivers themselves and (b) detected driver frustration, which is detected with a threshold of 0.5 for the posterior probability, shown in (c). The proposed method correctly detected driver frustration 80% of the time, with a false positive rate of 9%.

Detecting driver overreliance on ADAS

We investigated a possible method for detecting negative driver adaptation to an automated driving system by statistically analyzing the consistency of driver decision making and driver gaze behavior during automated driving based on a driver model which measures driver’s attentiveness [21],

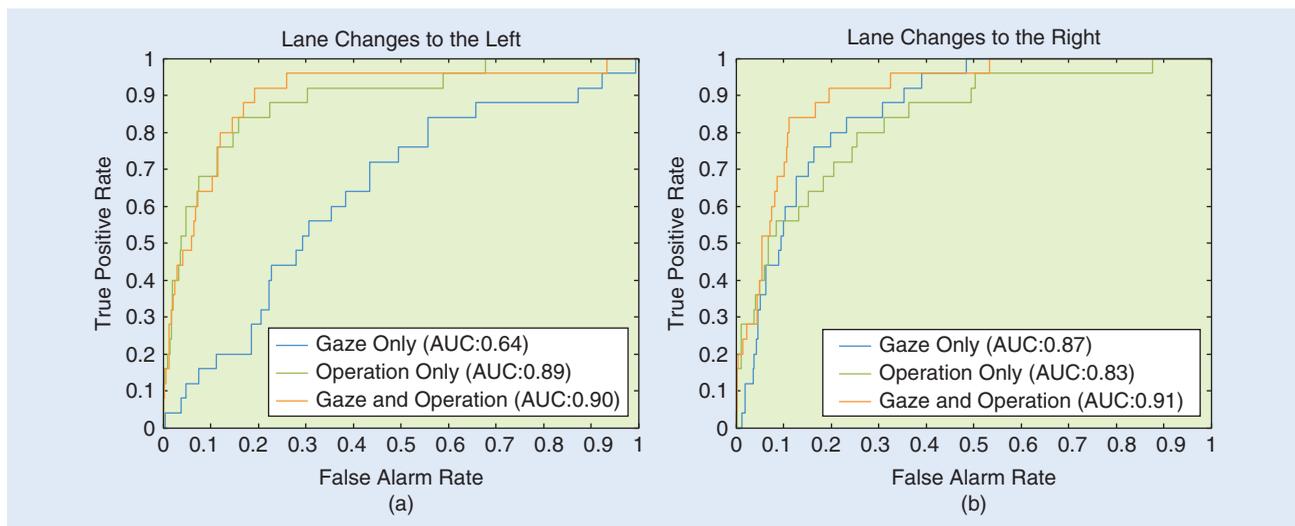


FIGURE 7. A comparison of risky lane-change detection performance (ROC curves) for (a) left lane changes and (b) right lane changes when using only driver gaze behavior (blue line), only driver vehicle-operation behavior (including vehicle motion) (green line), and both combined (orange line).

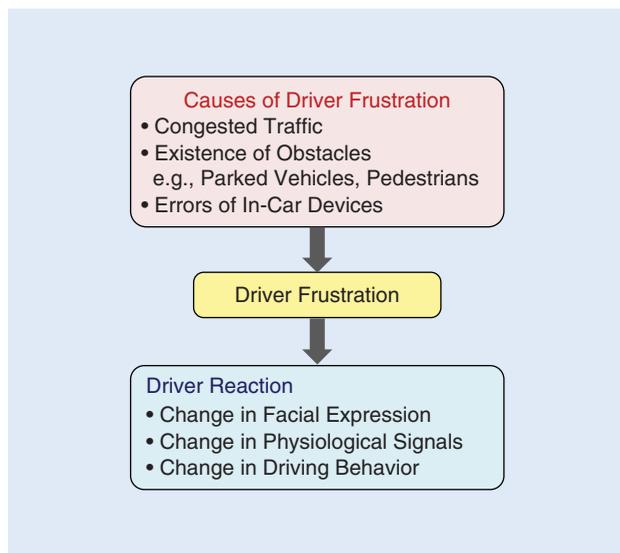


FIGURE 8. Modeling driver frustration using its causes and driver reactions.

[22]. We used logistic regression to analyze how drivers made lane change decisions (to either manually take control of the vehicle during the lane change or to allow the automated vehicle to make the lane change) in response to the risk level of the surrounding environment. Drivers were instructed to take manual control of the vehicle if they felt the situation was risky. Our experimental results suggested that drivers who paid less attention to the road ahead during automated

driving tended to be overdependent on the automated system. These drivers were less consistent when making decisions regarding lane changes and less sensitive to risk factors in the surrounding environment. For this experiment, we used a driving simulator due to the difficulty of conducting automated driving experiments on public roads at the present time. But soon large amounts of data from automated driving will also be available, and will be utilized for improving automated driving systems.

Conclusions and directions for future research

This article summarized data-centric approaches to modeling driver behavior and also surveyed some landmark projects for the collection of driving data on roads. We demonstrated the usefulness of the data-centric approach by reviewing the related studies, which analyzed and modeled driver behavior by applying statistical signal processing and machine-learning techniques to large amounts of driving data.

In the future, it will be possible to easily collect huge amounts of an even wider variety of real driving data via the Internet, from every connected vehicle or driver, everywhere and at any time. It will then be possible to more accurately model human driving behavior based on the huge number of driving-data samples being recorded in every possible driving environment and situation. Petabyte scale driving data and machine learning will also be effectively utilized for developing advanced driver-assistance and automated driving systems.

Acknowledgments

We would like to thank Prof. John Hansen and the anonymous reviewers for their helpful comments and suggestions. We would also like to thank our collaborators and students who contributed to our research.

Authors

Chiyomi Miyajima (miyajima@nagoya-u.jp) received her B.E., M.E., and Ph.D. degrees in computer science from Nagoya Institute of Technology, Japan, in 1996, 1998, and 2001, respectively. From 2001 to 2003, she was a research associate in the Department of Computer Science, Nagoya Institute of Technology. She was an assistant professor at the Graduate School of Information Science, Nagoya University, Japan, since 2003. She is currently a designated associate professor at the Institutes of Innovation for Future Society, Nagoya University. Her research interests include analysis and

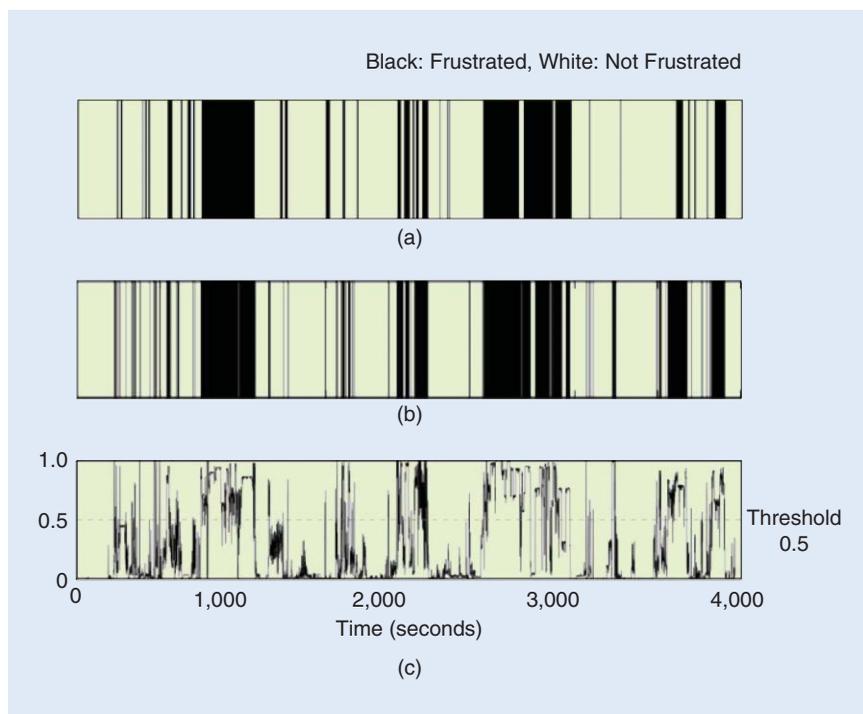


FIGURE 9. A comparison of actual and detected driver frustration using a BN. (a) Actual driver frustration reported by the drivers themselves. (b) Detected driver frustration using a BN. (c) Posterior probability at the "driver frustration" node of a BN.

modeling of driver behavior. She received the Best Conference Paper Award at the 2009 IEEE International Conference on Vehicular Electronics and Safety and the Best Paper Award at the Second International Symposium on Future Active Safety Technology Toward Zero Traffic Accidents.

Kazuya Takeda (kazuya.takeda@nagoya-u.jp) received his B.E.E., M.E.E., and Ph.D. degrees from Nagoya University, Japan, in 1983, 1985, and 1994, respectively. Since 1985, he has worked at Advanced Telecommunication Research Laboratories and at KDD R&D Laboratories, Japan. In 1995, he started a research group for signal processing applications at Nagoya University. His main focus is investigating driving behavior using data centric approaches, utilizing signal corpora of real driving behavior. He is currently a professor at the Institutes of Innovation for Future Society, Nagoya University. He is also a member of the Board of Governors of the IEEE Intelligent Transportation Systems Society. He is a Senior Member of the IEEE.

References

- [1] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzhaph, J. Jermeland, and R. R. Knipling, "The 100-car naturalistic driving study, Phase II: Results of the 100-car field experiment," National Highway Traffic Safety Administration, Washington, DC, Tech. Rep. DOT HS 810 593, Apr. 2006.
- [2] G. M. Fitch, S. E. Lee, S. G. Klauer, J. Hankey, J. Sudweeks, and T. A. Dingus, "Analysis of lane-change crashes and near-crashes," National Highway Traffic Safety Administration, Washington, DC, Tech. Rep. DOT HS 811 147, June 2009.
- [3] T. A. Dingus, J. M. Hankey, J. F. Antin, S. E. Lee, L. Eichelberger, K. E. Stulce, D. McGraw, M. Perez, and L. Stowe, "Naturalistic driving study: Technical coordination and quality control," Transportation Research Board, Washington, DC, SHRP2 Rep. S2-S06-RW-1, Mar. 2015.
- [4] R. K. Satzoda, P. Gunaratne, and M. M. Trivedi, "Drive analysis using lane semantics for data reduction in naturalistic driving studies," in *Proc. IEEE Intelligent Vehicles Symp.*, 2014, pp. 293–298.
- [5] M. Benmimoun, A. Pütz, A. Zlocki, and L. Eckstein, "euroFOT: Field operational test and impact assessment of advanced driver-assistance systems: Final results," in *Proc. FISITA World Automotive Congr. Lecture Notes in Electrical Engineering*, 2013, vol. 197, pp. 537–547.
- [6] R. Eenink, Y. Barnard, M. Baumann, X. Augros, and F. Utesch, "UDRIVE: The European naturalistic driving study," presented at the Transport Research Arena Conf., Apr. 2014.
- [7] M. A. Regan, A. M. Williamson, R. Grzebieta, J. L. Charlton, M. G. Lenné, B. Watson, N. L. Haworth, A. Rakotonirainy, J. E. Woolley, R. W. G. Anderson, T. M. Senserrick, and K. L. Young, "Australian 400-car Naturalistic Driving Study: Innovation in road safety research and policy," in *Proc. Australian Road Safety Research, Policing & Education Conf.*, 2013, pp. 1–13.
- [8] P. Raksincharoensak, "Drive recorder database for accident/incident study and its potential for active safety development," presented at the FOT-NET Workshop, Tokyo, Japan, Oct. 2013.
- [9] Research Institute of Human Engineering for Quality Life. Driving behavior database (in Japanese) [Online]. Available: <http://www.hqj.jp/database/drive/>
- [10] N. Kawaguchi, S. Matsubara, K. Takeda, and F. Itakura, "Multimedia data collection of in-car speech communication," in *Proc. Eurospeech*, 2001, pp. 2027–2030.
- [11] K. Takeda, J. H. Hansen, P. Boyraz, L. Malta, C. Miyajima, and H. Abut, "International large-scale vehicle corpora for research on driver behavior on the road," *IEEE Trans. Intell. Transport Syst.*, vol. 12, no. 4, pp. 1609–1623, Dec. 2011.
- [12] C. Miyajima, Y. Nishiwaki, K. Ozawa, T. Wakita, K. Itou, K. Takeda, and F. Itakura, "Driver modeling based on driving behavior and its evaluation in driver identification," *Proc. IEEE*, vol. 95, no. 2, pp. 427–437, Feb. 2007.
- [13] Y. Nishiwaki, C. Miyajima, N. Kitaoka, K. Itou, and K. Takeda, "Generation of pedal operation patterns of individual drivers in car-following for personalized cruise control," in *Proc. IEEE Intelligent Vehicles Symp.*, 2007, pp. 823–827.
- [14] P. Angkititrakul, C. Miyajima, and K. Takeda, "Modeling and adaptation of stochastic driver behavior model with application to car following," in *Proc. IEEE Intelligent Vehicles Symp.*, 2011, pp. 814–819.
- [15] P. Angkititrakul, C. Miyajima, and K. Takeda, "An improved driver-behavior model with combined individual and general drivin characteristics," in *Proc. IEEE Intelligent Vehicles Symp.*, 2012, pp. 426–431.
- [16] K. Takeda, C. Miyajima, T. Suzuki, P. Angkititrakul, K. Kurumida, Y. Kuroyanagi, H. Ishikawa, R. Terashima, T. Wakita, M. Oikawa, and Y. Komada, "Self-coaching system based on recorded driving data: Learning from one's experiences," *IEEE Trans. Intell. Transport. Syst.*, vol. 3, no. 4, pp. 1821–1831, Dec. 2012.
- [17] Y. Nishiwaki, C. Miyajima, N. Kitaoka, and K. Takeda, "A stochastic signal model for predicting the vehicle trajectory at lane change," in *Digital Signal Processing for In-Vehicle Systems and Safety, Part C*, J. H. L. Hansen, P. Boyraz, K. Takeda, and H. Abut, Eds. New York: Springer, 2012, ch. 19, pp. 271–282.
- [18] M. Mori, C. Miyajima, T. Hirayama, N. Kitaoka, and K. Takeda, "Integrated modeling of driver gaze and vehicle operation behavior to estimate risk level during lane changes," in *Proc. IEEE Int. Conf. Intelligent Transportation Systems*, 2013, pp. 2020–2025.
- [19] L. Malta, C. Miyajima, N. Kitaoka, and K. Takeda, "Analysis of real-world driver's frustration," *IEEE Trans. Intell. Transport. Syst.*, vol. 12, no. 1, pp. 109–118, Mar. 2011.
- [20] L. Malta, C. Miyajima, and K. Takeda, "A study of driver behavior under potential threats in vehicle traffic," *IEEE Trans. Intell. Transport. Syst.*, vol. 10, no. 2, pp. 201–210, June 2009.
- [21] H. Terai, H. Okuda, K. Hitomi, T. Bando, C. Miyajima, T. Hirayama, Y. Shinohara, M. Egawa, and K. Takeda, "An experimental study on the difference in drivers' decision-making behavior during manual and supported driving," in *Proc. 6th Int. Conf. Applied Human Factors and Ergonomics*, 2015, pp. 3136–3141.
- [22] C. Miyajima, S. Yamazaki, T. Bando, K. Hitomi, H. Terai, H. Okuda, T. Hirayama, M. Egawa, T. Suzuki, and K. Takeda, "Analyzing driver gaze behavior and consistency of decision making during automated driving," in *Proc. 2015 IEEE Intelligent Vehicles Symp.*, 2015, pp. 1293–1298.
- [23] FOT-Net Data. [Online]. Available: <http://fot-net.eu>
- [24] C. Miyajima, H. Ishikawa, M. Kaneko, N. Kitaoka, and K. Takeda, "Analysis of driving behavior signals recorded from different types of vehicles using CAN and Smartphone," presented at the 2nd Int. Symp. Future Active Safety Technology Toward Zero Traffic Accidents, Sept. 2013.
- [25] Y. Sakaguchi, M. Okuwa, K. Takiguchi, and M. Akamatsu, "Measuring and modeling of driver for detecting unusual behavior," in *Proc. 18th Int. Conf. on Enhanced Safety Vehicles*, 2003.
- [26] N. Oliver and A. P. Pentland, "Driver behavior recognition and prediction in a SmartCar," in *Proc. Int. Soc. Optical Engineering Aerosense Conf.*, Apr. 2000, pp. 280–290.
- [27] H. Okuda, N. Ikami, T. Suzuki, Y. Tazaki, and K. Takeda, "Modeling and analysis of driving behavior based on a probability-weighted ARX model," *IEEE Trans. Intell. Transport. Syst.*, vol. 14, no. 1, pp. 98–112, Aug. 2013.
- [28] A. Jain, A. Singh, H. S. Koppula, S. Soh, and A. Saxena, "Recurrent neural networks for driver activity anticipation via sensory-fusion architecture," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2016, pp. 3118–3125.
- [29] T. Taniguchi, S. Nagasaka, K. Hitomi, N. P. Chandrasiri, and T. Bando, "Semiotic prediction of driving behavior using unsupervised double articulation analyzer," in *Proc. IEEE Intelligent Vehicles Symp.*, 2012, pp. 849–853.
- [30] T. Bando, K. Takenaka, S. Nagasaka, and T. Taniguchi, "Unsupervised drive topic finding from driving behavioral data," in *Proc. IEEE Intelligent Vehicles Symp.*, 2013, pp. 177–182.
- [31] S. Yamazaki, C. Miyajima, E. Yurtsever, K. Takeda, M. Mori, K. Hitomi, and M. Egawa, "Integrating driving behavior and traffic context through signal symbolization," in *Proc. IEEE Intelligent Vehicles Symp.*, 2016, pp. 642–647.

Youjun Choi, Sang Ik Han, Seung-Hyun Kong,
and Hyunwoo Ko

Driver Status Monitoring Systems for Smart Vehicles Using Physiological Sensors

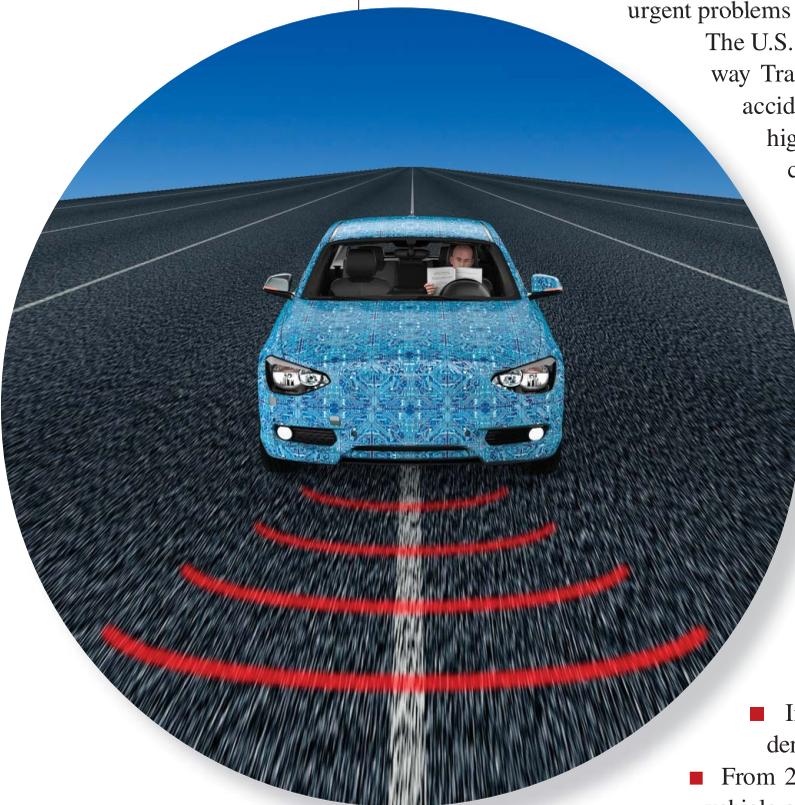
A safety enhancement system from automobile manufacturers

Automobiles provide a convenient form of transportation, and the number of automobiles in the world has been increasing rapidly, from 193 million in 1970 to more than 796 million in recent years [1]. However, automobiles have created a number of serious problems, such as accidents, traffic congestion, and air pollution, among which traffic accidents are one of the most serious and urgent problems threatening the safety of automobile users.

The U.S. Department of Transportation National Highway Traffic Safety Administration found that traffic accidents are mostly caused by drivers' inattention, high-speed driving, drunken driving, misperception, decision errors, and driver incapacitation (e.g., falling asleep or having a heart attack while driving). Among these causes, drunken driving and driver incapacitation account for about 25% of total traffic accidents [2]. In 2014, the American Automobile Association Foundation for Traffic Safety in the United States reported that an average of 328,000 traffic accidents annually involve a drowsy driver [3]. In Europe, 20–25% of total traffic accidents were due to drowsy drivers [4]:

- In France in 2011, there were 3,970 fatal accidents on the road, in which 732 cases occurred on straight roads; 85% of these accidents were due to drowsy drivers [4].
- In Germany, 25% of all fatal road traffic accidents were caused by drowsy drivers [4].
- From 2006 to 2010, in Finland, 17% of fatal motor vehicle accidents were related to fatigued drivers; they were responsible for 18% of deaths on the road [4].

Driver status monitoring (DSM) systems have emerged as an innovative technology to prevent traffic accidents from driver incapacitation. In recent developments of DSM systems, medical technologies used for patient diagnosis, including those utilizing electrocardiogram (ECG) and photoplethysmogram (PPG), are considered for the acquisition of driver's physiological signals, which is an effective approach that should be given a special attention.



SKY IMAGE LICENSED BY GRAPHIC STOCK
(ISTOCKPHOTO.COM/POSTERIORI)

Digital Object Identifier 10.1109/MSP.2016.2602095
Date of publication: 4 November 2016

Overview

To reduce the occurrence of serious traffic accidents caused by driver incapacitation due to fatigue and drowsiness, and to protect drivers from fatal accidents, increasing attention is being paid to DSM systems equipped with vision sensors, steering angle sensors (SASs), and physiological sensors. These types of DSM systems has been developed mostly by major automobile manufacturers such as Ford, Toyota, BMW, among others, since the early 2000s [5], and some important technologies are already commercialized or considered for smart vehicles. This article introduces the overall design of DSM systems, as applied to smart vehicles, being developed by the major automobile manufacturers. In particular, we focus on DSM systems utilizing physiological signals such as ECG and PPG, rather than the conventional DSM systems using vision sensors or SAS.

The section “Signal Processing Algorithms for DSM Systems Using Physiological Signals” introduces the signal processing techniques and noise reduction algorithms that are tested and used by the automobile manufacturers for the acquisition and enhancement of physiological signals in noise to extract the parameters related to the driver’s status. Electromagnetic compatibility (EMC) issues that need to be addressed for the application of DSM technologies to actual smart vehicles are introduced briefly in the section “EMC Issues in DSM Systems Using Physiological Signals.” The purpose of this article is to present a guideline for the design of DSM systems for readers who would like to develop the system for commercialization, by introducing the performance of DSM systems as obtained through actual vehicle tests.

Various DSM technologies

In general, DSM systems utilize various sensors to measure the steering angle, ECG, and PPG to monitor the driver’s status and analyze driver’s physiological signals and movements in the vehicle. The conventional DSM systems are equipped with a camera [charge-coupled device (CCD), infrared (IR), stereo, etc.] installed on the steering column and/or infrared-light emitting diodes (IR-LEDs) mounted inside vehicles so that the system can measure the driver’s eye blink rate, head location, and the driver’s facial direction to detect the driver’s status. Because the detection and recognition performance of the conventional DSM systems largely depend on image and vision processing algorithms, the conventional DSM systems require computationally expensive vision processing algorithms [5], [6], in general.

Some automobile manufacturers have developed DSM systems using driver’s driving patterns so that the performance of the system can be more reliable [4]. In these systems, once the driver’s driving patterns are obtained from a precise SAS, the patterns are stored in the in-vehicle database (DB) and used to estimate the driver’s status by comparing vehicle’s current movements with the DB. However, since a driver’s driving pattern

often depends on the driver’s intention (e.g., driver’s intentional eye blinking does not indicate a driver’s driving pattern), it requires additional information to increase the accuracy in the driver’s status detection [5], [6].

Recently, physiological signals such as ECG and PPG have been used in DSM systems, where the level of detection accuracy is greater when compared to the DSM systems using vision sensors or driving patterns, because the physiological signals are highly correlated to the driver’s physical status [7], [8]. Table 1 shows the features of DSM systems developed by major automobile manufacturers.

Signal processing algorithms for DSM systems using physiological signals

As mentioned in the section “Various Driver Status Monitoring Technologies,” the conventional DSM systems may not be able to detect a driver’s abnormal status correctly; for example, when a driver’s eye blinking and reckless driving are intentional. Therefore, there has been a strong demand for new DSM technologies [22], and among the new technologies introduced, DSM systems using driver’s physiological signals have gained increasing attention. In general, physiological signals are classified into ECG and PPG, which are measured from the

driver’s heart rate and pulse, respectively. In recent realizations of DSM systems using physiological signals, ECGs are more reliable in noisy in-vehicle environments. Therefore, this section focuses on DSM systems that utilize ECG and PPG for the primary and secondary observations, respectively. In addition, we will later introduce DSM algorithms used in the DSM systems.

Overall process of DSM systems using physiological signals

A single cycle of the ECG consists of a P signal, a QRS complex, and a T signal (refer to [23, Fig. 1]). Among them, the QRS complex has a relatively higher amplitude and signal-to-noise ratio (SNR) in comparison with the P and T signals, so it is utilized to monitor driver’s physical status. On the other hand, a single cycle of the PPG is composed of systolic and diastolic peaks [24]. Since the systolic peak has a relatively higher amplitude than the diastolic peak, it is used to detect the driver’s physical status.

In the system configuration, DSM systems using physiological signals consist of a physiological signal measurement block and a DSM algorithm. The physiological signal measurement block employs an analog signal processing technique for ECG acquisition, which is classified into a contact-based ECG acquisition technique that obtains ECG through the driver’s two hands on the steering wheel, and a noncontact-based ECG acquisition technique that acquires ECG through a capacitive connection between the driver and the electrodes on the driver’s seat. The contact-based ECG acquisition technique uses two electrodes installed on the steering wheel to have

Driver status monitoring systems have emerged as an innovative technology to prevent traffic accidents from driver incapacitation.

direct contact to both of the driver's hands. Since at least three measurement points on the body, i.e., two points on the left and right sides of the heart and a ground point, are necessary to stably measure the ECG, the analog signal processing technique utilizes the ground point on the analog circuit as a virtual ground point. The noncontact-based ECG acquisition technique uses pairs of electrodes for a stable ECG measurement, where a pair of electrodes is considered as a single channel to measure the ECG. The noncontact-based ECG acquisition technique improves the quality and sensitivity of the ECG measurement by utilizing simultaneous measurements obtained from multiple channels. In practice, both the contact-based and noncontact-based ECG acquisition techniques utilize an instrumental amplifier (INA) that has a high common mode rejection ratio (CMRR) (generally 80~120 dB) to clearly acquire and amplify weak ECG in mV units under very noisy environments. On the other hand, the PPG is measured by illuminating a point on the hand with a light and detecting the reflecting light; a photo detector is used to detect reflected lights from the tissue, blood vessels, and bone.

The analog signal processing technique for PPG (in other words, the contact-based PPG acquisition technique) uses

a nanometer green LED and current to voltage converter to acquire and amplify the weak PPG. In practice, high-pass filter (HPF), low-pass filter (LPF), and notch filter are employed to reduce the amplified noise and to enhance the acquired ECG and PPG [22].

The DSM algorithm consists of three functions performing QRS detection, signal enhancement, and driver status analysis. The QRS detection function extracts the QRS complex from ECG. Generally, threshold-based detection is widely used for easy implementation and low computational complexity, and depending on the threshold determination, there are fixed threshold (FT) and adaptive threshold

(AT) methods [22] used in practice. The FT method can be used effectively in a stationary ECG, however, the AT method can improve the accuracy of the QRS complex detection by setting thresholds adaptively, even though it may not provide a universal solution. In addition, there are other methods used for the QRS detection, for example, based on the neural networks and hidden Markov models, but due to disadvantages such as implementation difficulty and high computational cost, other methods are used only in medical applications that require high reliability.

In the system configuration, DSM systems using physiological signals consist of a physiological signal measurement block and a DSM algorithm.

Table 1. Classification of various DSM techniques.

Manufacturer	Development Status	Classification	Method
BMW [9], [10]	R&D	ECG	Monitoring heart rate using a steering wheel with a skin-resistance sensor
Mercedes Benz [11], [12]	Commercialized	Steering pattern	Analyzing of driving pattern with 70 parameters by a SAS and other drivers' behaviors, such as audio/air conditioner/window switch use
	R&D	ECG, PPG	Fusing PPGs and ECGs
Volkswagen [13]	Commercialized	Steering pattern	Monitoring counter-steering patterns
	R&D	Vision	Tracking head position, eye and face with a camera, light source, and image processor
Volvo [14]	Commercialized	Steering pattern	Monitoring steering patterns near the lane using a front camera and an SAS
	R&D	Vision	Comparing the threshold and current value by monitoring gaze, eyelids, eye blink rate, and head (or face) angle using a dashboard-mounted IR sensor and camera
Ford [15], [16]	R&D	ECG	Monitoring heart rate using noncontact-based ECG sensor implemented in a car seat
Toyota and Denso [8], [17]–[19]	Commercialized	Vision	Monitoring gaze, eyelids, eye blink rate, and head position using a dashboard-mounted IR-LED and a complementary metal-oxide-semiconductor camera
	R&D	ECG, PPG	Monitoring ECG and PPG using a steering wheel with different electrodes and 525 nanometer green LEDs
Denso [20]	R&D	ECG	Monitoring heart rate using noncontact-based ECG sensor implemented in a car seat
Nissan [21]	Commercialized	Steering pattern	Monitoring the constructed baseline of a given driving path with the driver's steering patterns using a SAS
	R&D	EEG	Predicting the driver's next driving actions beyond estimating the state and cognition of the driver

The signal enhancement function distinguishes the QRS complex from a distorted QRS complex by variability, abnormalities, low SNR, and artifacts [22]. This function makes use of the amplitude and slope of the R-peak as characteristic features, and a combination of Hilbert transformation (HT), mathematical morphology (MM), empirical mode decomposition (EMD), and filter banks (FBs) can be used; HT is used to extract the signal envelope, which is a characteristic feature of the ECG, MM shows excellent performance in reducing the impact of motion artifacts and line drifts, and EMD and FBs convert the QRS signal into a frequency domain and decompose the entire spectrum into subband spectrum.

The driver status analysis function measures the heartbeat interval (the interval between R-peaks, also known as the *RR interval*). The heart rate variability (HRV) technique [7] uses variations of the measured RR intervals to analyze a driver's autonomic nervous activity in either the time or frequency domain [23]. The analysis in the time domain utilizes mean RR (seconds), standard deviation of the RR interval (SDNN) (milliseconds), mean heart rate (HR) (beats/minute), standard deviation of the HR (Std RR) (milliseconds), and the root mean square of successive heartbeat interval differences (RMSSDs) (milliseconds) to estimate the effect of the autonomic nervous system on HR [7]. On the other hand, the frequency-domain analysis is classified into parametric and nonparametric methods [7]. The parametric and nonparametric methods estimate the power spectrum density (PSD) by means of the autoregressive (AR) model and fast Fourier transform (FFT) [25], respectively. The parametric method computes the power in low frequency (LF) and that in high frequency (HF) ($LF_{AR}[\text{ms}^2]$, $HF_{AR}[\text{ms}^2]$), and the percentage of the power in the low- and high-frequency segments ($LF_{AR}[\%]$ and $HF_{AR}[\%]$), respectively) in the parameterized FFT spectrum. Similarly, the nonparametric method estimates the power in low and high frequencies ($LF_{FFT}[\text{ms}^2]$, $HF_{FFT}[\text{ms}^2]$), and the percentage of the power in low- and high-frequency segments ($LF_{FFT}[\%]$, $HF_{FFT}[\%]$) in the FFT spectrum [7]. The frequency ranges of 0.04–0.15 Hz and 0.15–0.4 Hz are used for low- and high-frequency segments, respectively.

Slow-acting sympathetic activity may increase the HR, while fast-acting parasympathetic activity decreases the HR. A sympathetic activity influences the power in both low and high frequencies, whereas a parasympathetic activity makes an effect on the power only in high frequency. The balance between the effects of the sympathetic and parasympathetic activities is referred to as the *sympathovagal balance*, which can be measured by the ratio of powers in LF to that in HF [23]. It is found that when the sympathovagal balance changes to increase a sympathetic activity, the RR interval increases, and the driver status analysis function detects that the driver workload decreases (driver falls asleep). On the other hand, when the sympathovagal

balance changes to increase a parasympathetic activity, the RR interval decreases and the driver workload increases (driver wakes up).

Analog signal processing for DSM systems using physiological signals

Automobile manufacturers such as Toyota, Ford, BMW, and Daimler AG have been developing DSM systems using the ECG and PPG of the driver since the late 2000s [6], and some of these DSM systems are applied to actual vehicles. In 2010, BMW jointly developed a steering wheel with Technische Universitaet Muenchen as part of the Fit4Age Project. The developed steering wheel features contact-based signal acquisition using a strip-type skin-resistance sensor and a punctiform-type reflective pulse oximetry PPG sensor, which are installed in the BMW 730d [10]. In the system shown in

Figure 1, two conductive strip electrodes are attached all around the steering wheel to measure the driver's skin resistance. Because the resistance acquired from the driver's skin depends on how the driver holds the steering wheel, an electrodermal activity (EDA) circuit [26] that applies an automatic bias control using two operational amplifiers is applied to increase the

dynamic range of sensors [10]. A real driving test of a BMW 730d equipped with the developed steering wheel is conducted by 21 test drivers of average age 65. These participants drove the vehicle three times during 10 minutes on a 16-km long preselected route (highways, state roads, and urban areas) to evaluate the actual performance of the developed steering wheel [9]. From the driving test, it is found that approximately 81% of the meaningful measurements of skin resistance were obtained from the strip-type sensor, while 44% of the valid measurements of PPG were observed from the punctiform-type sensor. One study [9] reported that employing a strip-type sensor is suitable for commercializing the developed steering wheel by BMW.

From 2008 to 2011, Denso, in collaboration with Toyota and Nippon Medical University, developed a special steering wheel that can measure (acquire) ECG and PPG [17]–[19]. For ECG acquisition, positive and negative electrodes are attached to the right and left sides of the steering wheel, respectively, and chrome-coated metal electrodes with high input impedance (90,000 Ω) are used. From the steering wheel, 1–5 mV ECG is measured through the electrodes and amplified about 1,700 times. For PPG acquisition, a 525-nm green LED is attached on the steering wheel, which is widely used to reduce the surface reflection from the skin. The baseline wander, considered as an artifact caused by perspiration, respiration, body movements, and unstable contact with electrodes, is minimized after the acquired ECG and PPG pass through an HPF ($f_c = 0.3$ Hz for both ECG and PPG) [18]. Similarly, noise is minimized using an LPF ($f_c = 35$ Hz for ECG and $f_c = 30$ Hz for PPG), and then, the measured ECG and PPG are digitalized by an analog-digital converter (ADC)

Slow-acting sympathetic activity may increase the HR, while fast-acting parasympathetic activity decreases the HR.

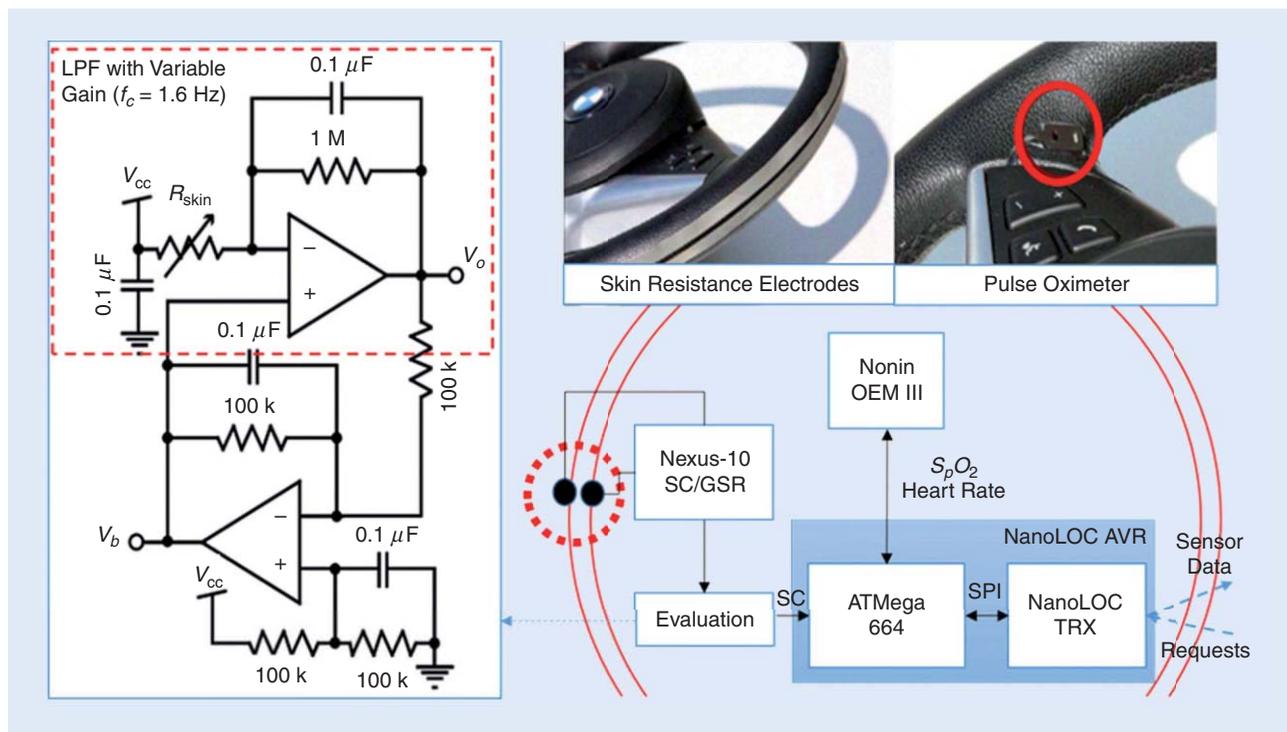


FIGURE 1. A contact-based physiological signal measurement technique developed by BMW [10], [26].

at a sampling rate of 100 Hz and a resolution of 10 bits, as shown in Figure 2. Since the contact area between the steering wheel and the driver's hands varies irregularly while turning the steering wheel on a curvy road and driving on an uneven road, the acquired ECG and PPG include a large amount of noise. Therefore, SNR improvement is required when ECG and PPG are measured through the contact on the steering wheel in real driving environments.

From the real driving tests on several roads, Denso found that the SNRs of ECG and PPG required more than 10 dB and 40 dB, respectively, to extract clear ECG and PPG. It is demonstrated that the SNR of PPG varies from 34 dB (stationary vehicle) to 15 dB (driving vehicle) during real driving tests at the speed of 80 km/hour and under acceleration of 0.65 m/s^2 . The SNR of PPG can be deteriorated by 19 dB, while the SNR

of ECG is degraded by 8 dB under the identical driving test condition [18]. To improve the SNR of PPG by enhancing the physical contact, a spring was mounted under the bottom of the PPG sensor to maintain a constant contact area between the driver's hands and the PPG sensor [18]. In addition, an ECG-triggered ensemble-averaging (EATEA) signal processing technique [18] that utilizes the peak detection with the ECG is applied to improve the SNR. As a result, Denso found that the SNR of PPG could be improved by 27 dB from 15 dB (without a spring and an EATEA) to 42 dB (with both a spring and an EATEA) during real driving tests at the speed of 80 km/hour and under acceleration of 0.65 m/s^2 .

ECG cannot be acquired with one hand because the contact-based technique requires that at least both hands are continuously in contact with the steering wheel.

In practice, ECG cannot be acquired with one hand because the contact-based technique requires that at least both hands are continuously in contact with the steering wheel. Therefore, some automobile manufacturers attempt to overcome this problem by developing noncontact-based techniques.

In 2013, Denso [20] developed a noncontact-based ECG acquisition technique utilizing a capacitance coupling between the driver and electrodes installed under the seat cover of the backrest. Denso's noncontact-based ECG acquisition technique utilizes the impedance, generated on the seat cover and the driver's clothing, which varies depending on the driver's physique, seating posture, and the type of clothing. Denso measured the impedance of the leather seat, T-shirt, shirt, and jacket and developed an equivalent circuit of the measured impedance of

each clothing in the form of parallel RC circuit. The resistance (R) components of the input impedances Z_1 and Z_2 have large values (in units of tera-ohms) in the sensor head [20], where Z_1 and Z_2 represent the combined impedance of the seat cover and clothing. The value of the circuit input resistance R_I is set to $1 \text{ G}\Omega$ to prevent a gain reduction due to a resistance-based voltage division and to maintain a dc bias value. In addition, the driven right leg (DRL) [27] circuit is applied to provide a feedback to the driver and to reduce the effect of static electricity resulting from the common mode noise. In [20], it is shown that the noncontact-based ECG acquisition technique achieves a target CMRR of

120 dB, and a higher CMRR is obtained by providing feedback from an electrode on the steering wheel gripped by the driver's hands, which is illustrated in Figure 3. The noise in the ECG is minimized by HPF ($f_c = 7$ Hz) and LPF ($f_c = 70$ Hz), and digitized by an ADC with a sampling rate of 1 kHz. The values of Z_1 and Z_2 will be close (balanced) to each other when the driver's back is in a good contact with the seat. However, if Z_1 and Z_2 lose a balanced state due to the driver's motion, the CMRR may decrease. Therefore, a multichannel noncontact-based ECG acquisition technique using a multiplexer is developed to achieve the CMRR value of larger than 120 dB. The performance of Denso's noncontact-based ECG acquisition technique is verified in a series of driving tests on straight roads and curves at a speed of about 70 km/hour [20]. From the real driving tests, the SNR (the ratio of the R-peak in the QRS complex to the baseline noise) of the measured ECG in cruise control is found to be 6.4 dB, and that under acceleration and deceleration is 5.9 dB.

Ford developed a noncontact-based ECG technique employing six insulated electrodes (tin-coated copper plates of $8\text{ cm} \times 5\text{ cm}$ in size) as ECG measurement sensors attached on the backrest of the seat [15], [16]. The technique was jointly developed with Aachen University in Germany and the Philips Chair for Medical Information Technology in 2011, and is applied to Ford's S-Max [16]. Ford utilizes the capacitive ECG (cECG) method [15] to obtain ECG measurements, even when the driver wears thick clothing. The impedance depends on the distance between the driver and insulated electrodes, which varies irregularly due to the driver's motion. In the technique, the impedance must be kept high, therefore, the active electrode technique [28] is applied to maintain a high impedance. Moreover, since the noncontact-based ECG acquisition technique is sensitive to noise; an active shielding technique [29] is needed to reduce the effect of noise. In the equivalent circuit of the active electrode, as depicted in Figure 4 [15], the overall coupling impedance Z_{couple} between the driver and the insulated electrode can be computed using $Z_{\text{couple}} = 1 / ((1/R_{\text{cloth}}) + j\omega C_{\text{cloth}}) + 1 / ((1/R_{\text{ins}}) + j\omega C_{\text{ins}})$, where, R_{ins} and R_{cloth} are the resistance of the electrode insulation and the driver's clothing, respectively, while C_{cloth} and C_{ins} represent the capacitive behavior of the overall contact impedance [15]. For a sufficient voltage drop at the input of the active electrode, a high-input impedance Z_B is required in the front end of the operational amplifier. The impedances Z_{couple} and Z_B perform high-pass filtering, therefore, noise components at the low frequencies of the measured ECG are removed sufficiently. However, when using the cECG method, static charges stored in the coupling capacitance and on the clothing should be removed, for which the bias resistance (R_{Bias}) is connected to Z_B in parallel. Since Z_B is large, the overall amplifier input impedance (Z_E) becomes $Z_E = (R_{\text{bias}} Z_B / R_{\text{bias}} + Z_B) \cong R_{\text{Bias}}$ [15]. Since coupled power-line interference occurring from electrodes exists in the ECG measurements [30], an INA with CMRR of about 115 dB is applied to remove the interference. In addition, LPF is used to remove the harmonics of the interference that is 50 Hz or

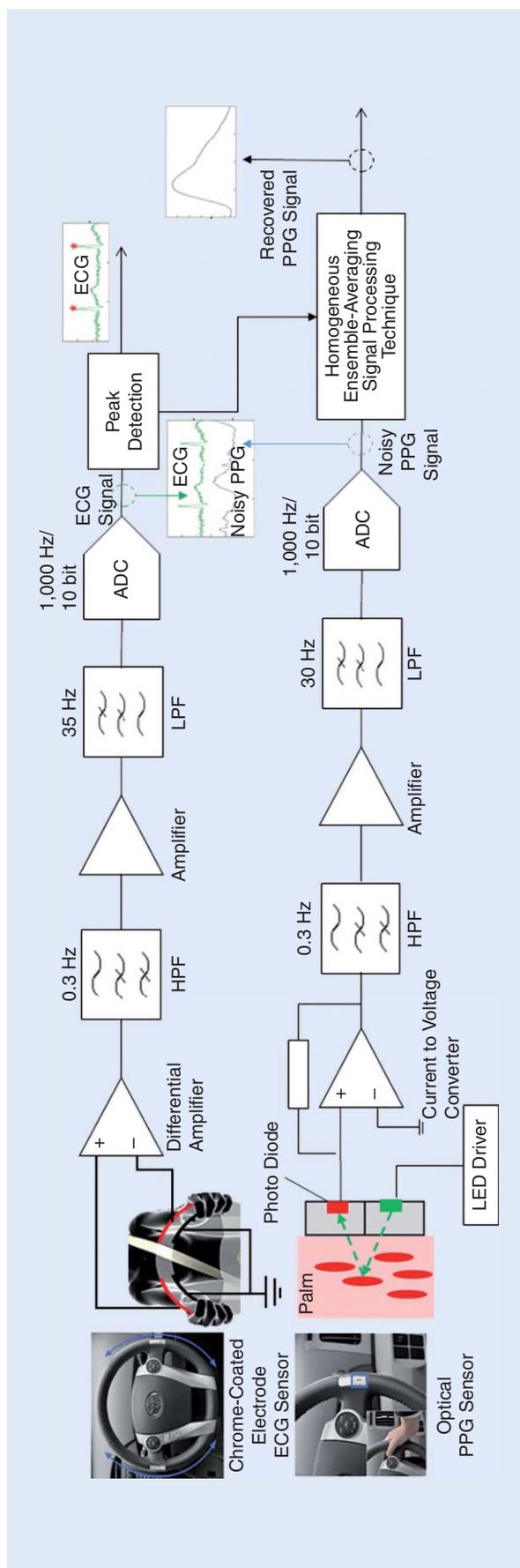


FIGURE 2. Analog signal processing for contact-based ECG and PPG acquisition by Toyota and Denso [18].

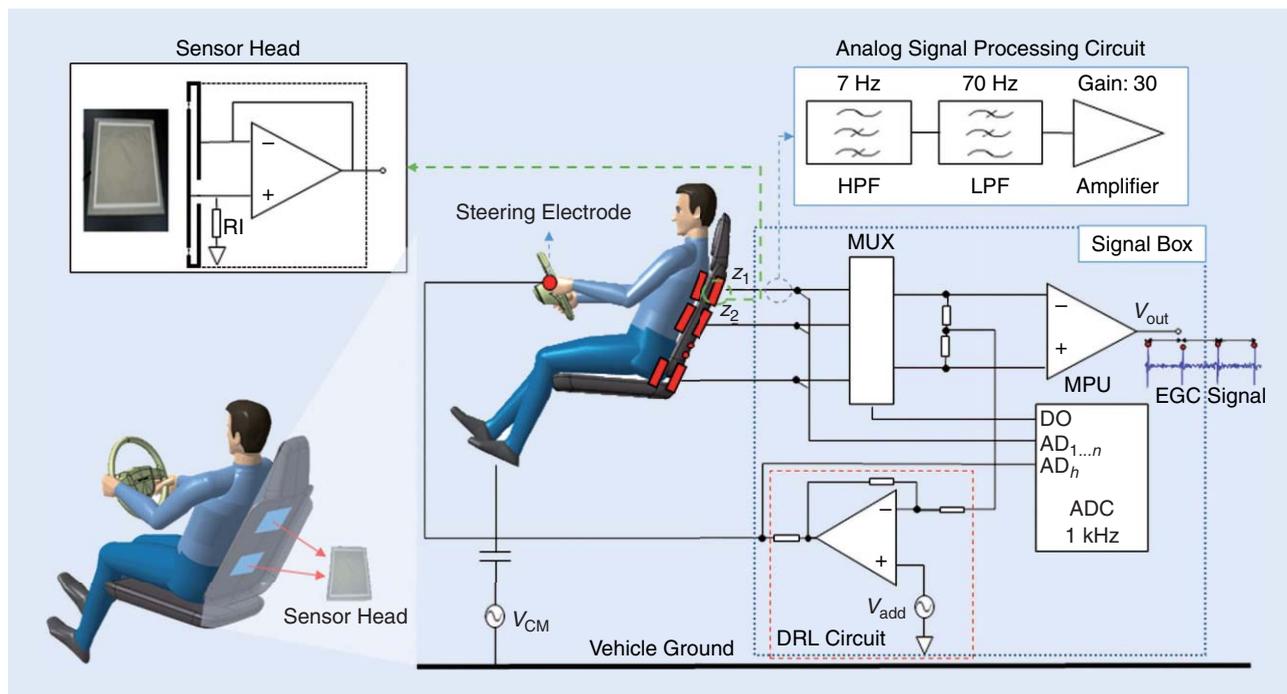


FIGURE 3. Analog signal processing for noncontact-based ECG acquisition by Denso [20].

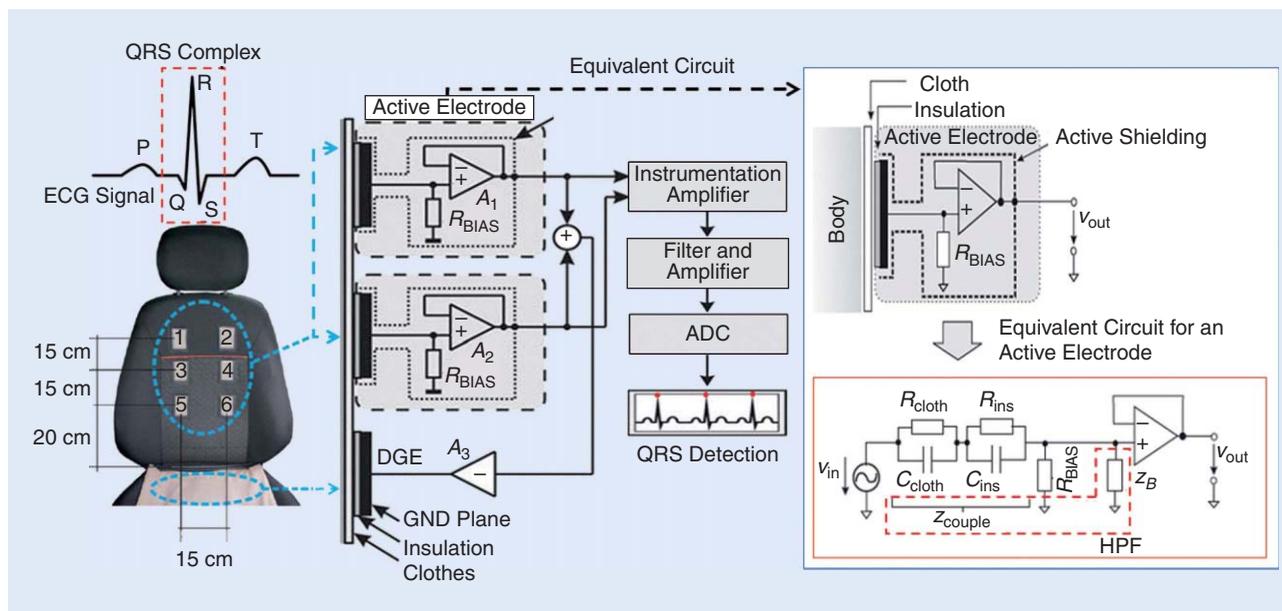


FIGURE 4. Analog signal processing for noncontact-based ECG acquisition by Ford [15], [16].

higher, and other noise components at a high frequency (typically >100 Hz). A notch filter ($f_c = 50$ Hz) is also used to further suppress noise components in the LPF output, while a HPF ($f_c = 0.3$ Hz) is employed to remove the dc noise component and to minimize the baseline drift. In [30], the driven ground electrode (DGE) method is used to reduce the interference on the cECG measurements additionally. Ford performed real driving tests with five subjects. The percentage of time (T_{icp}), during which at least four consecutive RR intervals is found

in the ECG measurements, is calculated; as a result, it is demonstrated that T_{icp} in the highway driving ($T_{icp} = 65\% \sim 98\%$) is much higher than that in the city driving ($T_{icp} = 6\% \sim 63\%$) [16]. This proves that Ford's noncontact ECG acquisition technique is suitable to measure the ECG while driving on the highway.

Daimler AG [10], [31] utilizes both contact-based and noncontact-based ECG acquisition techniques. As shown in Figure 5, five electrodes, made of brass, are attached

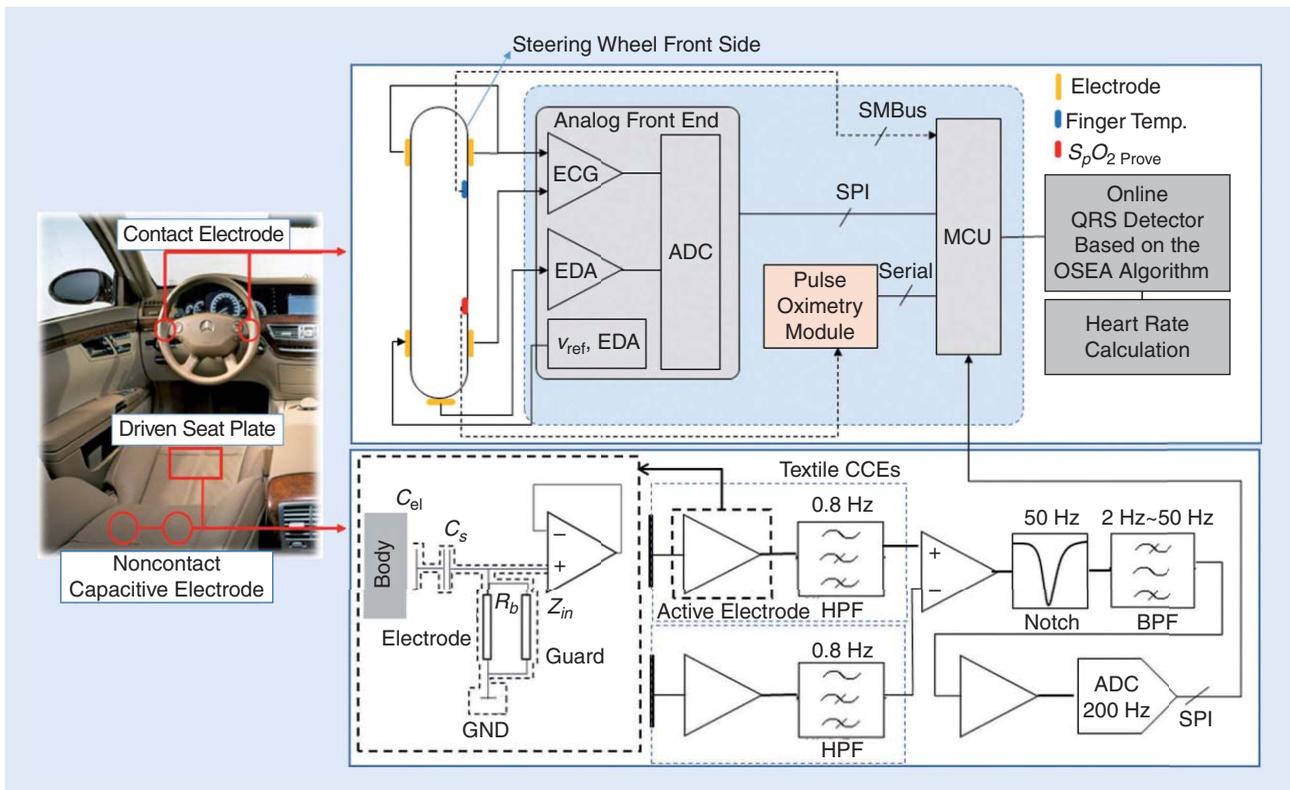


FIGURE 5. Analog signal processing for ECG acquisition technique by Daimler AG [10], [31].

to the steering wheel for a contact-based ECG acquisition, while capacitively coupled electrodes are attached to the seat and on the backrest of the seat for a noncontact-based ECG acquisition. Daimler AG employs an active electrode technique to get a high input impedance, similarly to the study by Ford. In the study [10], [31], Daimler AG can successfully reduce the variation of the capacitance between the driver and electrodes by connecting a capacitor (C_s) to an amplifier input in series [31]. An HPF ($f_c = 0.8$ Hz) is applied to the analog front end to reduce a baseline drift and dc offsets. Noise generated from the power baseline is removed by a notch filter ($f_c = 50$ Hz) and a bandpass filter ($f_{cl} = 2$ Hz and $f_{ch} = 50$ Hz).

DSM algorithms processing physiological signals

The DSM systems utilizing physiological signals need to remove artifacts in the signal measurements and to detect a driver's status while driving [5]. Therefore, DSM algorithms focus on removing artifacts, but there are few studies introduced in the literature. This section introduces DSM algorithms for contact-based ECG acquisition and noncontact-based ECG acquisition techniques developed by Toyota and Ford, respectively.

The signal detection function in the Toyota's DSM algorithm [8] calculates the correlation coefficient r successively by matching the QRS complex candidate found in the signal measurements with a QRS complex template. When r is larger than the threshold 0.7, the prompt QRS complex candidate is assumed a true QRS complex, as explained in the

QRS detection in Figure 6(a). The RR interval is measured between the two consecutive QRS complexes, and stored in the RR interval vector \vec{I} . Note that the detected R-peak positions are used to measure the RR intervals, however, there are true R-peaks and false R-peaks (outlier) due to the motion artifacts by the driver in the detected R-peaks.

The signal enhancement function in Figure 6(b) is performed to eliminate the R-peak by motion artifacts in the following steps:

- Calculate the mean ($\mu\vec{I}$) and the standard deviation ($\sigma\vec{I}$) of \vec{I} .
- Exclude outliers and store the RR intervals remaining after the exclusion of outliers to \vec{J} .
- Calculate the standard deviation ($\sigma\vec{J}$) and the median ($\text{Med}\vec{J}$) of \vec{J} of a set of 11 consecutive intervals in \vec{J} for each time index.
- Search for and remove outliers using the $\text{Med}\vec{J}$ and $\sigma\vec{J}$.
- Calculate the smoothed RR interval ($\overline{\text{RRI}}$) by replacing the detected outliers with $\text{Med}\vec{J}$.

The driver status analysis function in Figure 6(c) consists of signal analysis and detection rules subfunctions. In the signal analysis subfunction, $\overline{\text{RRI}}$ is transformed into the frequency domain to find the frequency parameter that is related to the sympathetic and parasympathetic nerves using the following steps:

- Sample $\overline{\text{RRI}}$ at 8 Hz.
- Divide sampled $\overline{\text{RRI}}$ data in five epoch intervals to obtain confident spectral estimation and to apply

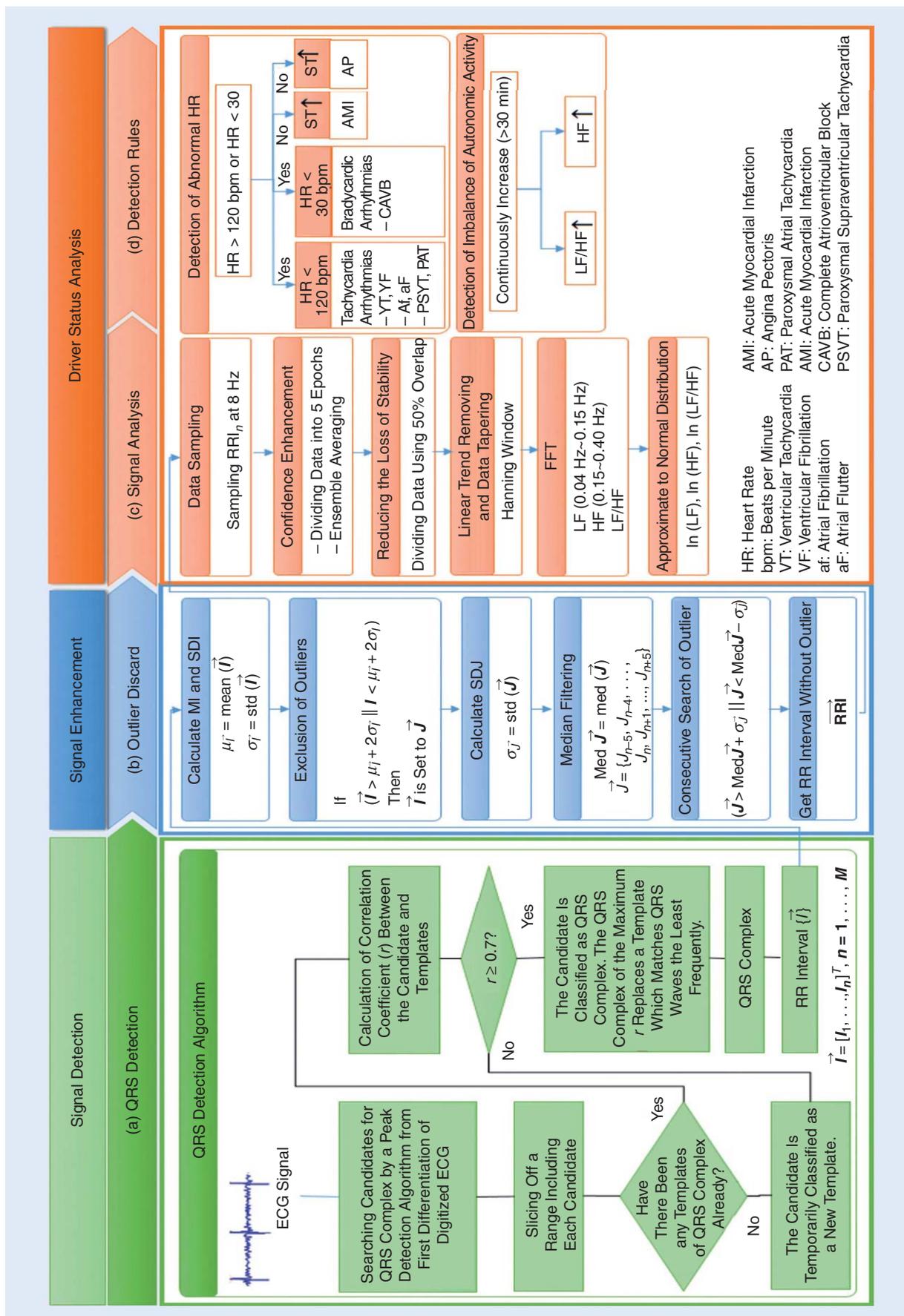


FIGURE 6. Toyota's DSM algorithm.

ensemble averaging, where the data length of each epoch is 64 seconds. The data is divided using 50% overlap to reduce the loss of stability.

- Apply the Hanning window to reduce linear trends in the data.
- Calculate the power of LF components that is a parameter of combined sympathetic and parasympathetic activities, the power of HF components that is a parameter of parasympathetic activity, and the ratio LF/HF as a parameter of sympathetic activity for each epoch.
- Calculate the natural logarithms of LF, HF, and LF/HF to approximate the distributions of LF, HF, and LF/HF to the normal distribution.

Toyota attached a chest ECG detection patch directly to the driver's chest [8], [19] to compare the ECG obtained from the patch with that measured from the DSM system. The mutual information technique (MIT) [8], which measures the similarity between two signals, is used to compare the two ECGs. The value of MIT higher than 0.047 indicates that two signals are strongly correlated, and in Toyota's study [19], the MIT value of the two ECGs is much higher than 0.047; MIT values for HR, $\ln(\text{LF})$, $\ln(\text{HF})$, and $\ln(\text{LF}/\text{HF})$ are 0.225, 0.223, 0.209, 0.184, respectively, with 95% confidence interval.

When the measured ECG includes too many R-peaks with motion artifacts, it is impossible to continuously detect the R-peaks successfully. Therefore, in practice, PPG is acquired to detect the lost RR intervals in the ECG so that the missing RR intervals can be replaced with the PPG peak intervals.

The detection rules subfunction analyzes the status of the driver using the $\ln(\text{LF})$, $\ln(\text{HF})$, and $\ln(\text{LF}/\text{HF})$ calculated in the signal analysis subfunction and the HR obtained from RR intervals [8], [19]. As explained in Figure 6(d), Toyota defines the detection rule of abnormal physiological status based on the HR, and demonstrates that the driver is drowsy if the LF/HF or HF increases steadily for more than 30 minutes.

Ford introduces another DSM algorithm utilizing a signal processing technique to remove outliers [16] and the HRV technique [7] to detect the driver workload. QRS complex is extracted from ECG by open source ECG analysis (OSEA) [32] as described in Figure 7(a), and the outliers of the extracted QRS complex are removed as explained in Figure 7(b)–(d). In the algorithm, a signal enhancement function using the quality indices QI is applied [16]. The basic quality indices used in the algorithm are quality index amplitude ($QI_{\text{Amplitude}}$), quality index standard deviation (QI_{Std}), and quality index saturation (QI_{sat}), as defined in Figure 7(b), that are calculated using \vec{x}_m and $\vec{x}_{m,\text{in}}$, where \vec{x}_m and $\vec{x}_{m,\text{in}}$ are the vectors of the signal measurements within ± 200 milliseconds and ± 50 milliseconds around the prompt R-peak of the QRS complex, respectively. The quality index amplitude $QI_{\text{Amplitude}}$ is the peak-to-peak amplitude ratio of $\vec{x}_{m,\text{in}}$ and

\vec{x}_m and used to discard strong but short baseline shift in \vec{x}_m , QI_{Std} is the standard deviation ratio of $\vec{x}_{m,\text{in}}$ and \vec{x}_m , and $QI_{\text{Std},m}$ indicates how much the R-peak of \vec{x}_m is separated from the noise around the R-peak. The quality index saturation QI_{sat} is a binary-valued index used to detect samples, in \vec{x}_m , that belong to the upper or lower saturation region of the ECG amplifier. The quality indices multiplication QI_m is calculated by multiplying $QI_{\text{Amplitude}}$, QI_{Std} , and QI_{sat} , and is used to remove outliers and to generate a training data before applying to the principal component analysis (PCA). The PCA and Hotelling's T-squared value [16] are used to calculate the probabilistic distance from the QRS complex to the

center of the training data, and to calculate the confidence measure (CM). A vector with a small deviation to the training data will result in a small Hotelling's T-squared value and, hence, is very likely to be a true QRS complex [16].

Since Ford's noncontact-based ECG acquisition technique measures the ECGs from multiple pairs of electrodes, the peaks from different channels need to be aligned before evaluating each QRS complex.

Referring to the ANSI/AAMI EC 57 norm

[32], all R-peaks from the multiple [three in Figure 7(d)] channels within a time interval of ± 75 ms can be assumed one single R-peak from the same heartbeat, and a missing R-peak in a channel is removed for the later processing. This leads to triplets of CMs based on Hotelling's T-squared values. The number of CMs whose values are smaller than a specific threshold determines if the R-peak is from a true QRS complex [16].

In [7], Ford found the relationship between the driver's HR and driver workload from the real driving tests. In the study, two men and two women (average age of 29.5 ± 2.08 years) are selected as the subjects, and the authors collected the ECGs while driving between Aachen and Brussels. The analysis of the driver's status is performed in both time and frequency domains. This study proves that the driver's HR and driver workload have a strong relationship, as shown in Figure 7(e), while the percentage of the power in the LF segments, $LF_{AR}[\%]$, is not very related to the driver workload. In particular, the HRs of all subjects are about 75~86 beats/minute in the city, about 72~83 beats/minute on highway I, and about 68~84 beats/minute on highway II. The traffic congestion in the city results in a higher HR, which indicates that the driver workload in the city is greater than that on the highway.

EMC issues in DSM systems using physiological signals

As the smart vehicles offer more driver-friendly functions, the number of electrical control units (ECUs) is growing. As a result, the number of problems caused by the electromagnetic waves from the ECUs is continuously increasing. Since the physiological signals, such as ECG, that is typically around $100\mu\text{V}\sim 3\text{mV}$ are very sensitive to electromagnetic waves, it

It is essential to design an analog signal processing circuit with a consideration of EMC to protect the DSM system from noise and interference from the ECUs.

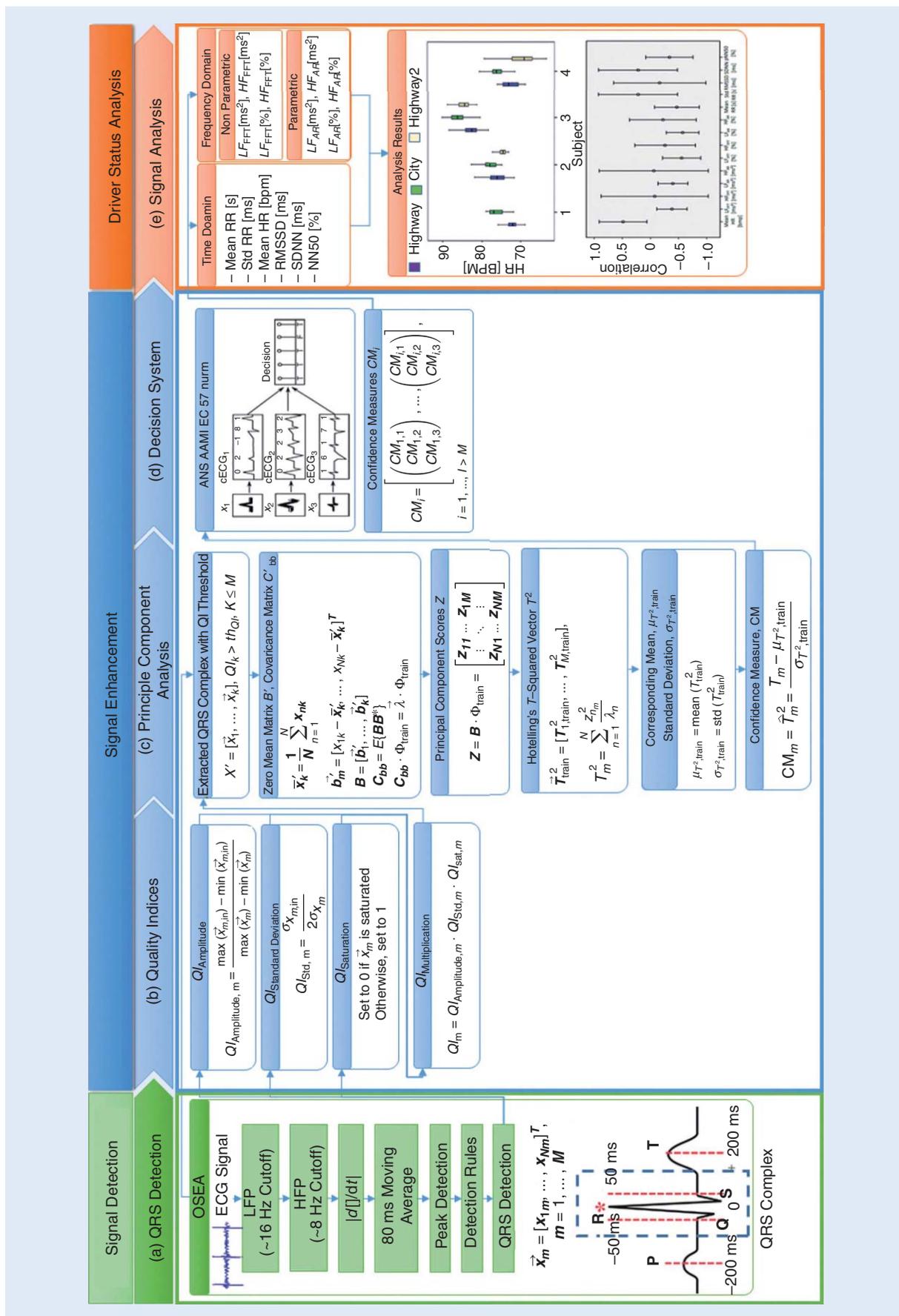


FIGURE 7. Ford's DSM algorithm.

may cause a malfunction of the DSM system. Therefore, it is essential to design an analog signal processing circuit with a consideration of EMC to protect the DSM system from noise and interference from the ECUs.

Therefore, in this section, we investigate the four EMCs, i.e., electromagnetic interference (EMI), electromagnetic susceptibility (EMS), surge, electrostatic discharge (ESD), that need to be considered when applying a DSM system using physiological signals to the vehicles.

EMI is the electromagnetic wave emitted from the ECU. As the electromagnetic waves can cause errors in other electronic devices, ECUs should be designed not to emit electromagnetic waves over the levels approved by the standard. In general, EMI is classified into conducted emission (CE) and radiated emission (RE); CE is the electromagnetic noise that is generated at frequencies of 30 MHz or lower from an ECU and penetrates into other ECUs through signal lines and power lines. Meanwhile, RE is the electromagnetic noise that is created at frequencies of 30 MHz or higher and propagates through the air.

The EMS indicates the impact of electromagnetic interference from outside a vehicle. EMS is classified into bulk current injection (BCI) and radiated susceptibility (RS). BCI is defined as the endurance of electromagnetic interference released through power connectors, while RS is the endurance of electromagnetic interference transmitted through the air. Therefore, in particular, the noise floor of the DSM systems for HR monitoring should be set to satisfy a noise level equivalent to 25 μV peak to peak over the bandwidth of 0.5–40 Hz [34].

A surge noise occurs when a sudden voltage is applied or current change occurs, for example, turning on/off of adjacent ECUs. This may supply a large current or voltage to the DSM system instantaneously. In general, a human's heart is the most sensitive to an electric current in a frequency range between 50–60 Hz so that a small current of 35 μA root mean square (rms) can damage the heart and cause a life-threatening event. Therefore, the DSM systems using the ECG should be carefully designed by employing a resistance or current limiting to allow a current of less than 10 μA rms to flow through the system [34].

Electrostatic discharge (ESD) is the sudden flow of electricity between two electrically charged objects by contact; it occurs when the driver touches the electrodes. The DSM system needs to limit the current that can be discharged through the contact of the driver's hand, so it requires a built-in protection circuitry to mitigate the ESD [34].

As a result, the DSM systems using physiological signals must be tested based on EMC standards. The EMC tests should be conducted with the CISPR (International Special Committee on Radio Interference) standards for emissions and the International Organization for Standardization (ISO) standards for susceptibility. EMI, EMS, surge noise, and ESD are

assessed by the CISPR25, ISO11452, ISO7637, and ISO10605 standards, respectively.

Conclusions

Traffic accidents are one of the most serious problems threatening the safety of automobile users. As reported in [4], the driver's incapacitation due to drowsiness and fatigue is one of the major causes resulting in fatal traffic accidents.

The DSM system is an innovative solution that can dramatically reduce traffic accidents caused by the driver's incapacitation [35]. Therefore, it is essential to develop DSM systems using vision sensors and SASs or using the driver's physiological signals such as ECG or PPG. The DSM systems using physiological signals have recently been developed by automo-

bile manufacturers because physiological signals contain an accurate information of the HR, which is very useful to determine the driver's physical status. In this study, we have introduced analog signal processing techniques employed in the DSM systems to acquire the physiological signals and DSM algorithms to detect the driver's status. However, DSM systems using physiological signals require further research on analog signal processing techniques, since physiological signals are very weak and vulnerable to noise and interference in the in-vehicle environments. Studies have shown that automobile manufacturers, universities, and research institutions are making efforts to improve the performance of the DSM systems using physiological signals and to apply the system to real vehicles. We expect that employing the DSM system will reduce traffic accidents due to the driver's incapacitation and, therefore, enhance the safety of automobile users.

Acknowledgments

Seung-Hyun Kong is the corresponding author of this article. This work (2013R1A2A2A01067863) was supported by the Midcareer Researcher Program through a National Research Foundation grant funded by the Korean government (MEST).

Authors

Yujun Choi (uchoi@kaist.ac.kr) received his B.S. degree in electrical engineering from Chonbuk National University in 2006, his M.S. degree in mechatronics engineering from the Gwangju Institute of Science and Technology in 2008, and he is currently a Ph.D. candidate with the Cho Chun Shik Graduate School for Green Transportation at the Korea Advanced Institute of Science and Technology, South Korea. Since 2011, he has been a senior researcher at the Korea Automotive Technology Institute. His research interests include mobility integration service and driver status monitoring systems for automotive.

Sang Ik Han (sangik.han@kaist.ac.kr) received his B.S. degree in electrical and electronics engineering from

The DSM system is an innovative solution that can dramatically reduce traffic accidents caused by the driver's incapacitation.

Chung-Ang University, Seoul, South Korea, in 2006; his M.S. degree in optics from the University of Central Florida, Orlando, in 2009; and his Ph.D. degree in electrical engineering from the University of Texas at Dallas, Richardson, in 2014. He is currently a research associate with the Cho Chun Shik Graduate School for Green Transportation at the Korea Advanced Institute of Science and Technology, South Korea.

Seung-Hyun Kong (skong@kaist.ac.kr) received his B.S. degree in electronics engineering from Sogang University, Seoul, South Korea, in 1992; his M.S. degree in electrical and computer engineering from Polytechnic University, New York, in 1994; and his Ph.D. degree in aeronautics and astronautics from Stanford University, Palo Alto, California, in 2005. His research interests include vehicular localization and navigation techniques, vehicular communication systems, cooperative vehicle systems, and advanced signal processing techniques for sensors and vehicular applications. He is currently an associate professor in the Cho Chun Shik Graduate School for Green Transportation at the Korea Advanced Institute of Science and Technology, South Korea, where he has been a faculty member since 2010. He is a Senior Member of the IEEE.

Hyunwoo Ko (kohyunwoo@kaist.ac.kr) received his B.S. degree in mechanical and control engineering from Handong Global University, Pohang, in 2014, and his M.S. degree from the Cho Chun Shik Graduate School for Green Transportation at the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2016.

References

- [1] S. C. Davis, S. W. Diegel, and R. G. Boundy, "Transportation Energy Data Book: Edition 34," Office of Energy Efficiency and Renewable, Oak Ridge National Laboratory, report ORNL-6991, 2015.
- [2] J. C. Fell and M. Freedman, "The relative frequency of unsafe driving acts in serious traffic crashes," National Highway Traffic Safety Administration (NHTSA), report DOT-HS-809-206, Washington, DC, 2001.
- [3] B. C. Tefft, "Prevalence of motor vehicle crashes involving drowsy drivers, United States, 2009–2013," Report Prepared for the American Automobile Association (AAA) Foundation for Traffic Safety, Washington, DC, 2014.
- [4] T. Akerstedt, C. Bassetti, F. Cirignotta, D. Garcia-borreguero, M. Goncalves, J. Horne, D. Leger, M. Partinen, T. Penzel, P. Philip, and J. C. Verster, "Sleepiness at the wheel," Institut National du Sommeil et de la Vigilanc, European Sleep Research Society, 2013.
- [5] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, 2011.
- [6] S. Boverie, "General driver monitoring module definition SoA," Public Rep. DESERVE No. D3.2.1, 2013.
- [7] B. Eilebrecht, S. Wolter, J. Lem, H. J. Lindner, R. Vogt, M. Malter, and S. Leonhardt, "The relevance of HRV parameters for driver workload detection in real world driving," presented at the IEEE Computing in Cardiology Conf., Krakow, Poland, Sept. 2012.
- [8] M. Osaka. (2012). Customized heart check system by using integrated information of electrocardiogram and plethysmogram outside the driver's awareness from an automobile steering wheel [Online]. INTECH Open Access Publisher. Available: <http://cdn.intechopen.com/pdfs-wm/27021.pdf>
- [9] L. D'Angelo and T. C. Luth, "Integrated systems for distraction-free vital signs measurement in vehicles," *ATZ Worldwide eMag.*, vol. 113, no. 11, pp. 52–56, 2011.
- [10] L. T. D'Angelo, J. Parlow, W. Spiessl, S. Hoch, and T. C. Luth, "A system for unobtrusive in-car vital parameter acquisition and processing," presented at the 4th Int. Conf. IEEE Pervasive Computing Technologies for Healthcare, Munich, Mar. 2010.
- [11] S. Heuer, B. Chamadiya, A. Gharbi, C. Kunze, and M. Wagner, "Unobtrusive in-vehicle biosignal instrumentation for advanced driver assistance and active safety," presented at the IEEE EMBS Conf., Kuala Lumpur, Malaysia, 2010.
- [12] Mercedes-Benz. (2016). Attention assist system [Online]. Available: https://techcenter.mercedes-benz.com/en_CA/attention_assist/detail.html
- [13] T. Von Jan, T. Karnahl, K. Seifert, J. Hilgenstock, and R. Zobel, "Don't sleep and drive-VW's fatigue detection technology," in *Proc. 19th Int. Conf. Enhanced Safety of Vehicles*, 2005, pp. 1–12.
- [14] Volvo Technology Corp., "Method for interpreting a drivers' head and eye activity," Eur. Patent EP2204118B1, July 23, 2014.
- [15] M. Walter, B. Eilebrecht, T. Wartzek, and S. Leonhardt, "The smart car seat: Personalized monitoring of vital signs in automotive applications," *Pers. Ubiquit. Comput.*, vol. 15, no. 7, pp. 707–715, 2011.
- [16] T. Wartzek, B. Eilebrecht, J. Lem, H.-J. Linder, S. Leonhardt, and M. Walter, "ECG on the road: Robust and unobtrusive estimation of heart rate," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 11, pp. 3112–3120, 2011.
- [17] K. Futatsuyama, T. Kawachi, and T. Nakagawa, "Cuffless blood pressure monitoring using steering wheel sensor system," *Trans. Jpn. Soc. Med. Biol. Eng.*, vol. 52, Supplement, pp. OS-67–OS-68, Oct. 2014.
- [18] K. Futatsuyama, N. Mitsumoto, T. Kawachi, and T. Nakagawa, "Noise robust optical sensor for driver's vital signs," SAE Technical Paper 2011-01-1024, 2011, doi:10.4271/2011-01-1024.
- [19] M. Osaka, H. Murata, Y. Fuwamoto, S. Nanba, K. Sakai, and T. Katoh, "Application of heart rate variability analysis to electrocardiogram recorded outside the driver's awareness from an automobile steering wheel," *Circulation J.*, vol. 72, no. 11, pp. 1867–1873, 2008.
- [20] K. Sakai, K. Yanai, S. Okada, and K. Nishii, "Design of seat mounted ECG sensor system for vehicle application," *SAE Int. J. Passeng. Cars-Electron. Electr. Syst.* 6(1):342–348, 2013, doi:10.4271/2013-01-1339.
- [21] Nissan Motor Co. Ltd., "Driver's fatigue degree estimation apparatus and driver's fatigue degree estimation method," U.S. Patent 0 164 400 A1, June 18, 2015.
- [22] M. Elgendi, B. Eskofier, S. Dokos, and D. Abbott, "Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems," *PLoS One*, vol. 9, no. 1, pp. e845101, 2014.
- [23] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, 2003.
- [24] D. McDuff, S. Gontarek, and R. W. Picard, "Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 12, pp. 2948–2954, 2014.
- [25] B. Kim, S.-H. Kong, and S. Kim, "Low computational enhancement of STFT-based parameter estimation," *IEEE J. Select. Topics Signal Process.*, vol. 9, no. 8, pp. 1610–1619, 2015.
- [26] M. Z. Poh, N. C. Swenson, and R. W. Picard, "A wearable sensor for unobtrusive, long-term assessment of electrodermal activity," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 5, pp. 1243–1252, 2010.
- [27] B. Winter Bruce and J. G. Webster, "Driven-right-leg circuit design," *IEEE Trans. Biomed. Eng.*, vol. BME-30, no. 1, pp. 62–6, Jan. 1983.
- [28] S. Leonhardt and A. Aleksandrivic, "Non-contact ECG monitoring for automotive application," in *Proc. 5th Int. Symp. Medical Devices and Biosensors*, Hong Kong, China, 2008, pp. 183–185.
- [29] R. J. Prance, A. Debray, T. D. Clark, H. Prance, M. Nock, C. J. Harland, and A. J. Clippingdale, "An ultra-low-noise electrical-potential probe for human-body scanning," *Meas. Sci. Technol.*, vol. 11, no. 3, pp. 291–297, 2000.
- [30] A. Schommartz, B. Eilbrecht, T. Wartzek, M. Walter, and S. Leonhardt, "Advances in modern capacitive ECG systems for continuous cardiovascular monitoring," *Acta Polytech.*, vol. 51, no. 5, pp. 100–105, 2011.
- [31] B. Chamadiya, S. Heuer, U. G. Hofmann, and M. Wagner, "Towards a capacitively coupled electrocardiography system for car seat integration," in *Proc. 4th European Conf. Int. Federation for Medical and Biological Engineering*, Berlin, Germany, 2009, vol. 22, pp. 1217–1221.
- [32] P. Hamilton, "Open source ECG analysis," in *Proc. IEEE Computers in Cardiology Conf.*, 2002, pp. 101–104.
- [33] "Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms," ANSI/Association for the Advancement of Medical Instrumentation, Tech. Rep. EC57, 1998.
- [34] B. Crone, "Mitigation strategies for ECG design challenges," Analog Devices Inc., Tech. Rep. MS-2160, 2011.
- [35] I. Akahashi, T. N. H. Kanamori, T. Tanaka, S. Kato, Y. Ninomiya, E. Takeuchi, M. Makiguchi, and H. Aoki, "Automated safety vehicle stop system for cardiac emergencies," presented at the IEEE Int. Conf. Emerging Technologies and Innovative Business Practices for the Transformation of Societies, Balaclava, Mauritius, 2016.

Amirhossein S. Aghaei, Birsen Donmez,
Cheng Chen Liu, Dengbo He, George Liu,
Konstantinos N. Plataniotis,
Huei-Yen Winnie Chen, and Zohreh Sojoudi

Smart Driver Monitoring: When Signal Processing Meets Human Factors

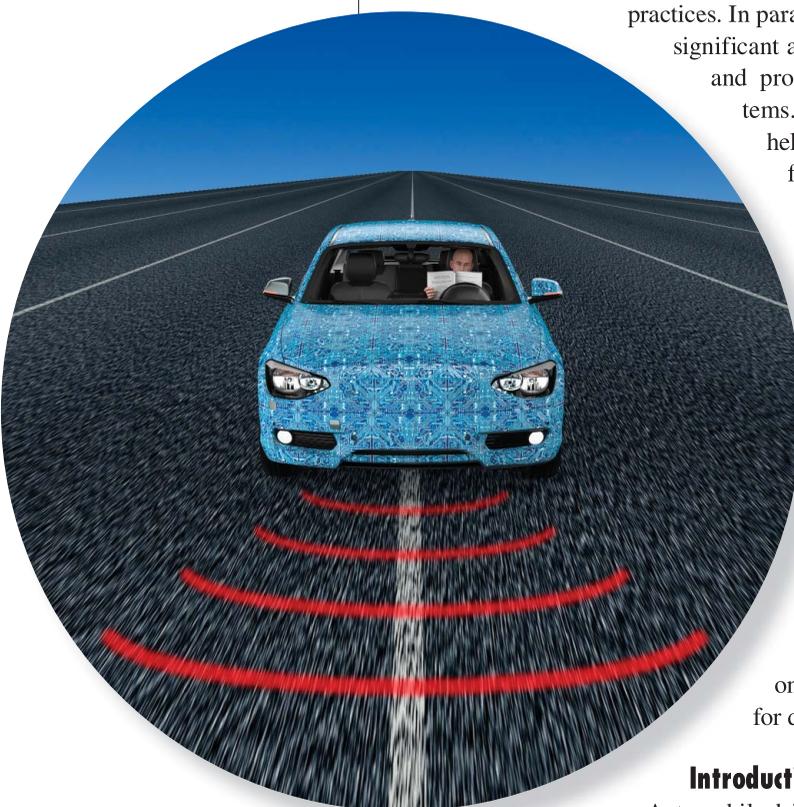
In the driver's seat

This article provides an interdisciplinary perspective on driver monitoring systems by discussing state-of-the-art signal processing solutions in the context of road safety issues identified in human factors research. Recently, the human factors community has made significant progress in understanding driver behaviors and assessed the efficacy of various interventions for unsafe driving practices. In parallel, the signal processing community has had significant advancements in developing signal acquisition and processing methods for driver monitoring systems. This article aims to bridge these efforts and help initiate new collaborations across the two fields. Toward this end, we discuss how vehicle measures, facial/body expressions, and physiological signals can assist in improving driving safety through adaptive interactions with the driver, based on the driver's state and driving environment. Moreover, by highlighting the current human factors research in road safety, we provide insights for building feedback and mitigation technologies, which can act both in real time and postdrive. We provide insights into areas with great potential to improve driver monitoring systems, which have not yet been extensively studied in the literature, such as affect recognition and data fusion. Finally, a high-level discussion is given on the challenges and possible future directions for driver monitoring systems.

Introduction

Automobile driving is a demanding activity in which the drivers simultaneously control the vehicle (laterally and longitudinally), manage hazards, and make decisions about navigation and route planning. Driving can also be a dangerous activity, with driver error a major crash risk source. The World Health Organization estimates that motor vehicle crashes kill 3,000 people a day [1]. In 2013 alone, the United States reported 32,719 fatalities and estimated 2,313,000 injuries from motor vehicle crashes [2]. The majority of crashes is attributed to driver error. An analysis of crashes between 2005 and 2007

SKY IMAGE LICENSED BY GRAPHIC STOCK
©ISTOCKPHOTO.COM/POSTERIORI



Digital Object Identifier 10.1109/MSP.2016.2602379
Date of publication: 4 November 2016

in the United States found the driver as the critical reason in 94% of crashes, citing errors such as distraction, sleep, excessive speed, and false assumption of others' actions [3].

Certain driver populations are at an elevated risk. Older drivers are affected by age-related declines in perceptual, cognitive, and motor abilities [4], while younger and particularly novice drivers tend to lack sufficient skills to recognize or anticipate road hazards [5]. Younger drivers may also be risk unaware and engage in potentially risky activities, such as speeding. For example, among U.S. drivers involved in fatal crashes in 2013, males aged 15–24 years old were the most likely to be speeding at the time of the crash [2]. In addition to the driver's age and gender, there are social-psychological factors that can also contribute to crash risk. For example, self-reported sensation-seeking tendency, which is a personality trait, has been associated with risky driving (e.g., records of impaired driving and self-reported speeding [6]). Attitudes (evaluations of the potential outcomes) and perceived social norms (how other people behave, or what societal expectations there are) about unsafe driving behaviors can also play a role in vehicular crashes. For example, [7] found that drivers who drove faster than others also had a more positive attitude toward speeding. The same study also found that drivers who perceived others to drive at excessive speeds were more likely to drive fast compared to those who perceived others to comply with the limits. Medical conditions, such as sleep apnea [8], are also known to increase crash risk. Finally, more transient driver characteristics, i.e., driver state, such as fatigue, drowsiness, distraction, mental workload, and mood, can also affect safety by impairing a driver's information-processing abilities and altering his or her risk-taking tendencies.

Infotainment systems and smart personal devices (e.g., smartphones) can add to the already demanding activity of driving and may increase crash risk. Digital interactions have become an integral part of many drivers' daily life such that they expect digital content to be available even while driving. Furthermore, it is inevitable that a growing array of technologies find their way into the vehicle given their large economic value. It should be noted that although certain types of technology lead to distraction, technology can also enhance safety. Automobile manufacturers are constantly seeking innovative solutions that leverage emerging sensor technology and increased computational power to support the driver. In the past decade, there has been a rapidly growing interest in both the signal processing and the human factors communities to develop smart driver monitoring systems that can sense and monitor the driver's state, vehicle operation, and changes in the environment to provide drivers with useful and appropriately timed feedback.

A smart driver monitoring system can provide immediate driving assistance (e.g., automated lane-departure warnings), support drivers in maintaining situation awareness (SA) while

driving (e.g., alerts of incidents ahead), and foster positive behavioral changes in the long run (e.g., by providing aggregated information after a trip to help drivers understand potential consequences of unsafe driving behaviors). A smart driver monitoring system can also address some of the risk factors associated with certain driver characteristics. For example, the 2016 Chevy Malibu has a feature tailored to teen drivers. This feature warns teen drivers when they exceed a predetermined speed limit, blocks the sound from the stereo until the front seat belts are buckled, and generates a report card on driving safety metrics (e.g., number of over-the-speed-limit warnings and forward collision alerts). In addition to real-time feedback on speed and seat belt use, the system provides an opportunity for parents to discuss

safer driving behaviors with teen drivers in a postdrive feedback form, thus calibrating or reinforcing a more accurate mental model of safe driving.

While the Malibu's teen driver monitoring system provides useful feedback, it does not directly address the more transient characteristics of the driver. The driver's state, such as fatigue, anger, or mental overload (i.e., when the driver's cognitive capacity limit has been reached), can influence driving behavior, both in terms of information processing capabilities (e.g., reduced reaction times to hazards due to drowsiness) and decisions to perform risky behaviors (e.g., following a vehicle too closely while in a state of aggravation). The ability to monitor driver state can ensure that factors such as current mood, mental workload, and vigilance level of the driver are taken into account and feedback is provided in a timely manner (e.g., based on early detection of drowsiness rather than detection of unsafe driving performance). However, there are many challenges associated with monitoring driver state. The remainder of this article focuses on addressing these challenges, as well as reviewing important design considerations regarding driver feedback. Overall, we aim to 1) provide readers with an overview of current and emerging signal acquisition and processing techniques for driver state monitoring systems, and 2) draw insights from human factors research on driver behaviors and intervention strategies to facilitate discussions about existing techniques for smart driver monitoring systems.

Driver state monitoring

The state of a driver can have a significant impact on her driving performance, a relationship best captured by the well-known Yerkes–Dodson law [9]. The Yerkes–Dodson law, when adapted to driver state [10], suggests an inverted-U function between driver's arousal and driving performance. As illustrated in Figure 1, both low arousal and high arousal are associated with performance decrements, and optimal performance occurs when there is an appropriate amount of arousal to keep the driver attentive, but not stressed. On one hand, a driver who is fatigued or has lost vigilance due to monotonous

The human factors community has made significant progress in understanding driver behaviors, and assessed the efficacy of various interventions for unsafe driving practices.

conditions (e.g., prolonged driving on a straight road with little traffic) is at an increased risk of losing control of the vehicle or failing to respond to hazards in time. On the other hand, strenuous driving conditions (e.g., reduced visibility or heavy traffic) or additional workload from a secondary task (e.g., having to manipulate the navigation system or engaging in a heated phone conversation) can increase stress, overload the driver, and drastically limit the cognitive resources available to drive safely.

A smart driver monitoring system should provide appropriate feedback or interventions that take into consideration the current condition of the driving environment (e.g., poor weather), the behaviors exhibited by the driver (e.g., hands off the steering wheel), as well as the driver's physiological state (e.g., excitement or drowsiness). Smart monitoring can also be extended to monitor other in-vehicle activities, including the driver's manual or auditory interaction with in-vehicle devices, driver's conversation on the phone or with passengers, and conversation among passengers. Data about the in-vehicle environment and activities may help interpret driver state more accurately. However, due to limited space, this article focuses on driver state monitoring.

While the exact state of a driver may not be directly measurable, it can often be inferred from its manifestations, such as driver's facial and body expressions (e.g., eye gaze, eye closure, yawning, and seating posture), how the driver controls the vehicle [e.g., speed, lane deviation (LD), and headway distance], and/or driver's physiological signals [e.g., heart rate (HR) and brain activity]. Figure 2 provides a schematic overview of some examples for these three measurement categories. The remainder of this section describes the aforementioned measurement categories and the technologies associated with them, and concludes with remarks about the current algorithms for fusing the information obtained from these measures.

Vehicle-based measures

Advanced driver monitoring technologies based on vehicle measures are becoming widely available on the market and are provided by both car manufacturers and aftermarket retrofits. The following list categorizes these measures into three major groups and briefly summarizes some of the recent findings in the literature regarding these measures and their potential relationship with driver's arousal level. It should be noted that the reported relationships are based on experimental studies. For detailed information about the experimental procedures and participant demographics, please refer to the corresponding references.

■ Driver input to the vehicle:

- **Steering:** Drowsiness or low arousal may decrease the frequency of steering reversals, deteriorate steering performance, increase the amplitude of steering-wheel movements, and increase the standard deviation of steering angle. Visual distraction may cause steering neglect and overcompensation, whereas cognitive distraction or overload may lead to undercompensation (e.g., see [11]).

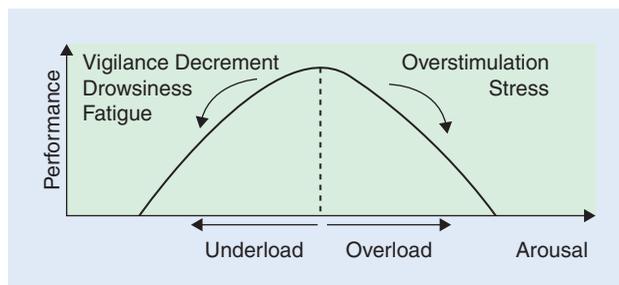


FIGURE 1. The inverted-U model for driver performance as a function of arousal.

- **Braking:** Both low and high arousal may cause changes in the brake and acceleration patterns (e.g., see [12]). Cognitive distraction may cause hard (high intensity) braking (e.g., see [11]), and high arousal may increase response time when braking is required (e.g., see [13]).
- **Vehicle response to driver input:**
 - **Velocity/acceleration:** Fatigue or low arousal may increase the standard deviation of speed. Visual distraction due to working with infotainment systems may decrease speed, increase speed variance, and result in unintended speed changes (e.g., see [11] and [13]).
 - **Jerk:** Both low and high arousal may cause no steering correction for a prolonged period of time followed by a jerky motion to correct steering (e.g., see [14]).
- **Vehicle state relative to traffic and the driving environment:**
 - **Headway distance:** Distractions may both increase or decrease headway distance, depending on the type of distraction and the overall driving demands. Some studies have shown that working with email systems or using iPods lead to increased headway distance, whereas some studies have shown watching DVD players can result in decreased headway distance (e.g., see [13]).
 - **LD:** Drowsiness or low arousal may increase standard deviation of lane position (e.g., see [13]). Distraction due to secondary tasks may also increase LD (e.g., see [15] and [16]).
 - **TLC:** Drowsiness or low arousal may lead to irregularities in the vehicle's tracking and increase the range of deviations. Both low and high arousal may cause driving patterns with lower TLC values (e.g., see [17]).

The first two categories can be directly measured by sensors installed inside the vehicle, whereas the third category requires information regarding the driving environment or the road geometry. In the latter case, the extra information can be obtained from cameras or radars installed on the vehicle, or from the information provided by smart infrastructures [vehicle-to-infrastructure (V2I) communication] or other vehicles [vehicle-to-vehicle communication (V2V)]. Figure 3 presents data collected in a driving simulator experiment conducted by the authors, which illustrates the changes in a variety of vehicle state measures with respect to driver sleepiness and demonstrates how measures like TLC and LD can be investigated in terms of range (min, max), mean, and standard deviation. It can be seen that when the driver's sleepiness increases, a

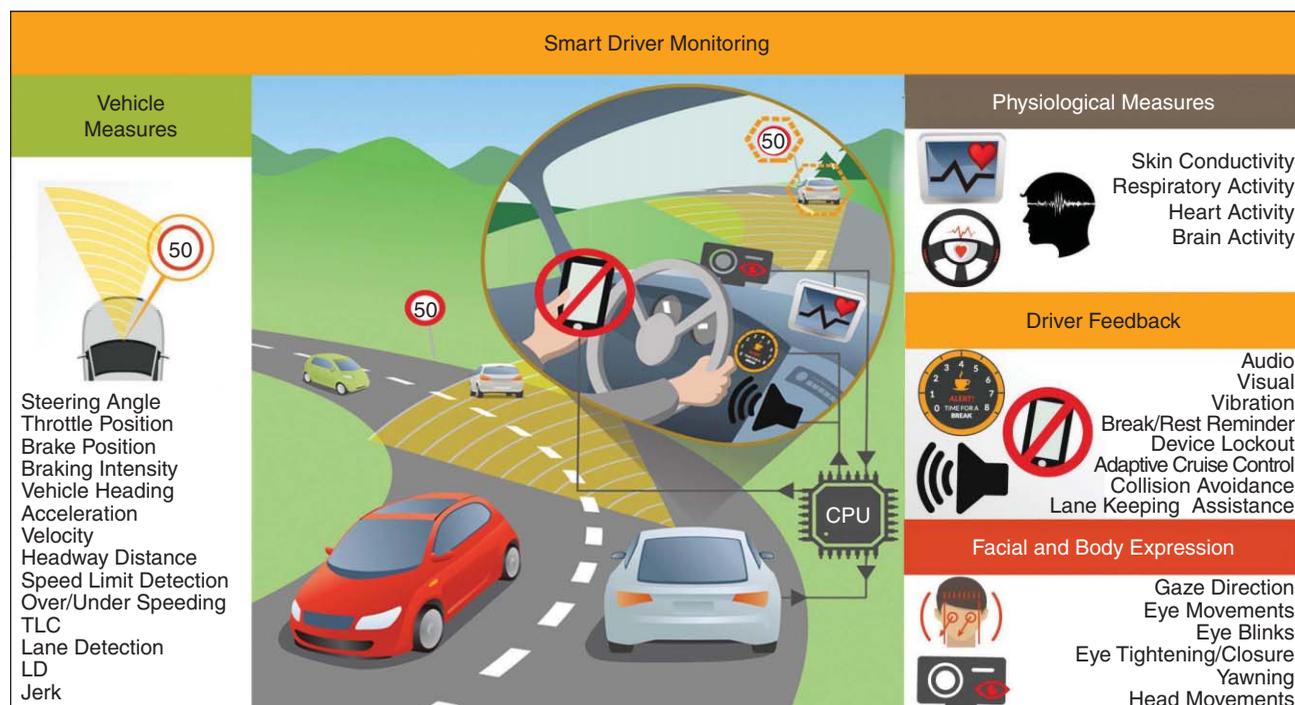


FIGURE 2. An illustration of sensors and feedback technologies that can be used in smart driver monitoring systems. TLC: time-to-lane crossing.

number of vehicle measures are affected. Over/underspeeding happens more frequently at higher levels of sleepiness, and so does large LD variance. Moreover, during interval G and at the end of interval I, the driver changes lanes very frequently, which suggests a degradation in lane keeping ability for this particular driving scenario. Finally, at higher levels of sleepiness, TLC tends to get closer to zero, which is an indicator of unsafe driving. This effect is reflected in the graphs that indicate the number/fraction of TLC measure going below a certain threshold. Note that the experimental setup used for the study in Figure 3 is described in Figure 4. A comprehensive review of the literature on vehicle measures is provided in [13] and [19].

Almost all major automobile manufacturers have equipped their newer high-end models with warning technologies, such as lane departure (which tracks the vehicle's position in the lane and alerts the driver when the vehicle starts to drift laterally) and collision warning (which uses a forward-looking camera or a radar to measure the distance to the lead vehicle and alerts the driver when this distance is less than a safe threshold). Furthermore, many manufacturers such as Tesla, Mercedes-Benz, BMW, Lexus, Infinity, and Honda have equipped some of their models with vehicle control intervention technologies. Some examples include lane-keeping assistance (which actively applies corrective torque to the steering wheel to keep the vehicle in the lane) and collision avoidance (which activates the brakes if the driver does not respond to the collision warning).

Aside from the aforementioned systems, some manufacturers have introduced smart technologies that utilize vehicle

measures to monitor driver states, such as drowsiness. For example, Mercedes-Benz's Attention Assist (http://m.mercedes-benz.ca/en_CA/attention_assist/detail.html) warns the driver if signs of drowsiness are detected. This system is mainly based on tracking steering wheel movements, while assessing parameters such as speed, longitudinal/lateral acceleration, and indicator and pedal usage. It also takes into account external factors such as crosswinds and the unevenness of the road. Another example is Volkswagen's Driver Fatigue Detection System (http://www.volkswagen.com.au/en/technology_and_service/technical-glossary/fatigue-detection.html), which continually measures steering wheel movements together with other vehicle signals as side information, and warns the driver if signs of fatigue are detected.

Facial and body expression measures

Observable cues from driver's facial and body expressions can be used to infer about a driver's arousal level. For example, eye blinks, percentage of eye closure (PERCLOS) [20], and yawning are particularly useful for detecting fatigue and drowsiness, whereas eye movements, gaze direction, and head movements provide information on drivers' visual attention (e.g., on the road or infotainment system). These measures are heavily reliant on video processing, and thus have been traditionally difficult to implement due to computational limitations. However, recent advancements in computational power and algorithm design have made such implementations possible in real time.

Figure 4 illustrates an example of using video cameras [in (a), a dashboard-mounted camera is shown and in (b) the head-mounted eye tracker is shown] in a research setting

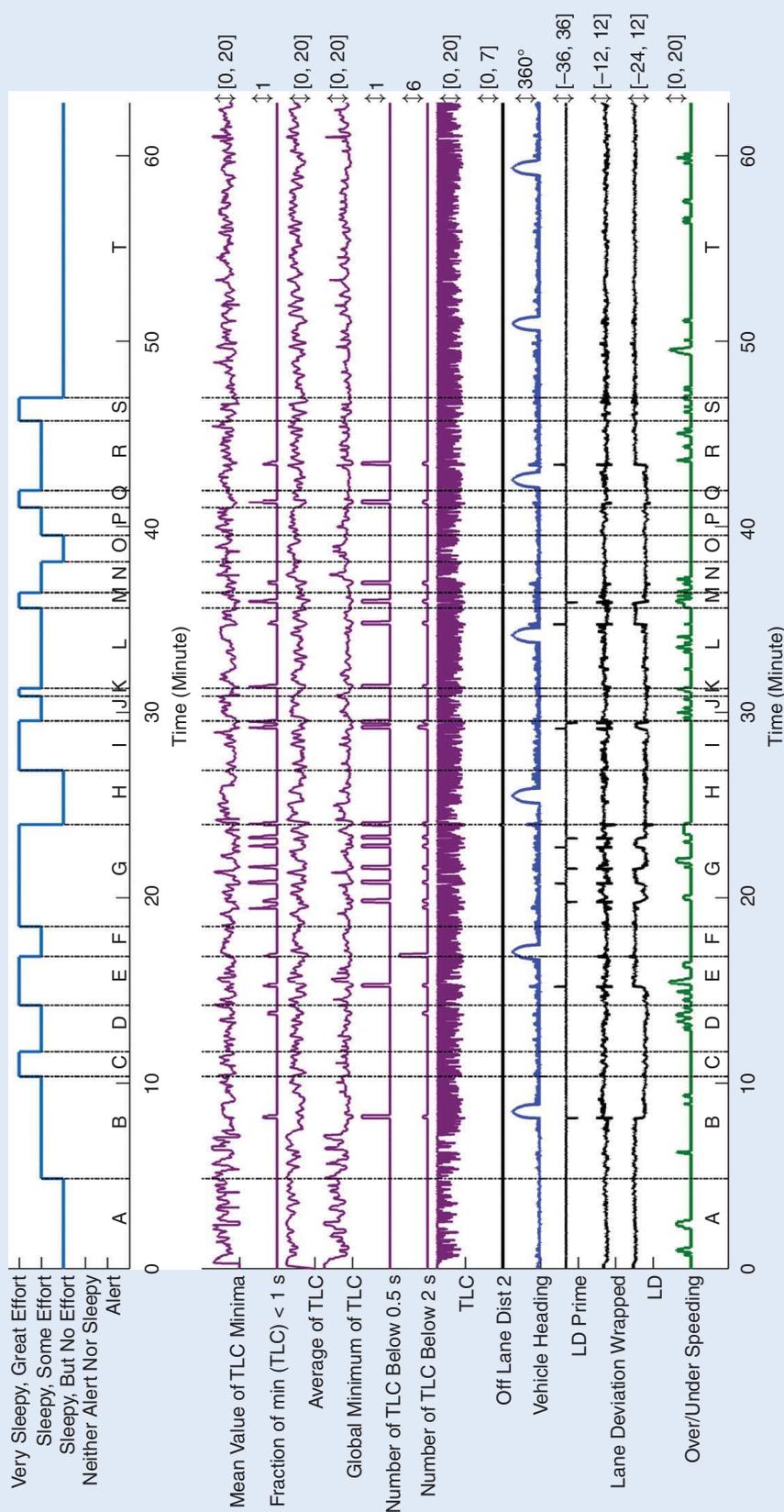


FIGURE 3. An example data of vehicle measures from a simulator study on driver sleepiness we conducted, which is currently under analysis. Data shown here represent a participant's one-hour drive, observing a 60-mi/hour speed limit on a two-lane track with no traffic. At the top, sleepiness is illustrated using a five-level subjective scale (measured based on [18]), and at the bottom, the changes in various vehicle measures are illustrated. The vertical dotted lines separate the time intervals during which the sleepiness has a constant scale value. The range for each vehicle measure is illustrated on the right side.

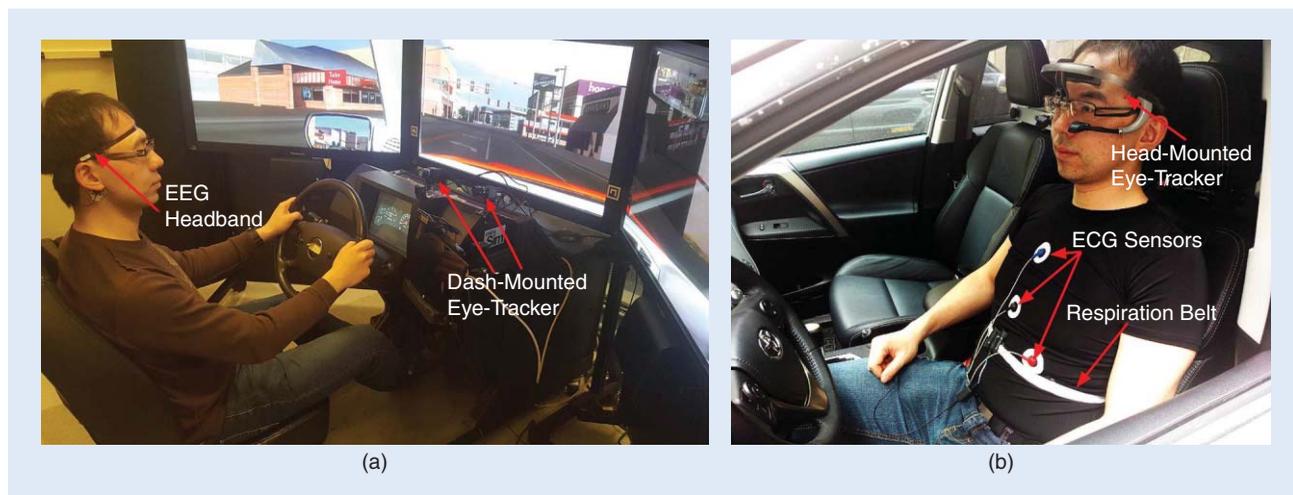


FIGURE 4. Example research setup at the University of Toronto for (a) a driving simulator and (b) an instrumented vehicle. The simulator shown is a fixed-based simulator, paired with a dashboard-mounted eye tracker. Both the driving simulator and the instrumented vehicle are set up to provide common vehicle measures, including speed, lane position, brake/accelerator pedal position, and headway distance. In the instrumented vehicle, the driver is shown wearing a head-mounted eye tracker. Electrocardiogram (ECG), galvanic skin response (GSR), and respiration sensors are also attached to driver's clothing to illustrate potential sensor placement on the participant. For actual data collection, physiological sensors are to be placed directly on the skin.

for driver state monitoring. A commercial example is the Lexus Driver Monitoring System, which was implemented in some Lexus vehicles beginning in 2006 (e.g., Model GS 450h) and in some Toyota vehicles beginning in 2008 (e.g., Crown Hybrid), and which utilizes six built-in near-infrared (NIR) light-emitting diode (LED) sensors as well as a charge-coupled device camera facing toward the driver. It performs eye tracking, eyelid detection, and head motion detection to detect the onset of sleepiness from facial expressions and warns the driver.

A typical system for analyzing facial and body expressions is illustrated in Figure 5. First, visual input is acquired using one or more cameras. A camera system with an active lighting source (NIR) is often used to alleviate the problem of extremely large illumination variations during driving. Then, the region of interest or facial/body landmarks need to be located by either direct detection from each frame, or with the help of tracking algorithms. At the next stage, metafeatures, such as PERCLOS, head pose, and gaze direction (and different statistics associated with them), are extracted. Finally, the system makes a decision by looking for particular patterns in the time series of the metafeatures, which often involves fusion with other signals as will be discussed later in the “Data Fusion” section. A comprehensive review of associated algorithms and techniques used in the literature is provided in [11] and [21].

The first fundamental task for this system is to successfully locate desired tracking features in each frame in real time, which can be the face, eyes, or more precise facial landmarks. The ensuing software detection algorithm is highly dependent on the image acquisition hardware used. In earlier studies, structured illumination/camera eye-tracking systems

with an NIR source were widely used to exploit the special optical property of human pupils known as the *dark/bright* pupil effect. When the illumination is coaxial with gaze path, pupils reflect light back to the source and appear bright in the acquired images, while an offset of illumination placement would not create such an effect and would result in dark pupils [22]. For this system, pupils are located by simple subtraction of the alternating dark pupil frame and the bright pupil frame, a method that is computationally cheap [22]. Although this eye-tracking technique is accurate in the laboratory environment, its heavy reliance on the hardware setup degrades its robustness for realistic driving scenarios. Based on advancements in computational power and machine-learning algorithms, it is possible to develop intelligent computer vision solutions with large video data input for real-time applications that do not rely heavily on image acquisition

hardware configurations. For example, the boosted cascade classifier, the underlying idea behind the Viola–Jones object recognition algorithm [23], has been incorporated in many recent camera-based driver face or eye detection algorithms (e.g., see [24]). Furthermore, tracking algorithms are commonly utilized for enhancing the

processing speed of the overall system. These algorithms allow us to replace the costly processing of a full frame of the input image by the lower cost searching for the target in a more constrained manner. Algorithms such as Kalman filter or particle filter have been commonly used in both eye trackers with structured lighting [25] and computer vision-based systems [24].

Next, discriminant metafeatures can be calculated based on the obtained lower-level features. This step can be seen as higher-level feature extraction. Many of the visual indicators

Observable cues from driver's facial and body expressions can be used to infer about a driver's arousal level.

for fatigue detection focus on the eye area, using detailed geometric measurements of iris, pupils, eyelids, and sometimes the specular point on the iris. These measures are also used for gaze estimation, in which head pose estimation is also necessary. In general, there are two approaches for head movement tracking: the rigid three-dimensional (3-D) head model [24] and front face orientation estimation using facial landmarks [25]. Metafeatures can also be calculated based on statistics of the aforementioned visual cues, for example frequency, average intensity, or standard deviation in a certain time window. Using multiple metafeatures often improves the system performance, where a fuzzy inference system [22] or Bayesian network [25] can be used for data fusion. Detection can also be performed by learning special patterns from the temporal signals using algorithms such as support vector machines [26]. Data fusion and inference methods will be further discussed the “Data Fusion” section. An alternative approach is the use of holistic feature extraction algorithms that do not require any handcrafted features and extract statistically discriminant features/patterns from the image frame using holistic machine-learning and pattern recognition techniques. This approach is widely used in other areas (e.g., face recognition, emotion recognition) and has been shown to have great potential; however, it has not yet been implemented in driver monitoring.

Today’s cameras are empowered with abilities beyond human vision, and thus can capture extremely subtle changes, such as HR using thermal imaging [27]. These technologies potentially enable nonintrusive evaluation of driver physiology. Another example is the use of pupillary dilation as an indicator of cognitive load. In addition to facial features, body

expressions such as seating posture, speaking, and head gestures have been studied in the literature to infer human states [24], [28]. For example, activities related to distracted driving such as answering a cell phone, eating, and smoking can be detected by hand tracking [29]. Moreover, vision systems looking at feet and hand positions can directly monitor driver’s interaction with the vehicle, and use this information for maneuver prediction [30].

It is worth mentionin that monitoring a driver’s facial and body expressions also provides possible means for determining the driver’s emotions. These measures are successfully used in other contexts such as human-computer interaction studies for affect recognition. There is great potential to use facial expressions for affect recognition in the context of driver monitoring, which has not been explored yet. Finally, it is noteworthy that facial/body expressions do not necessarily need to be measured using video recorders. For example, eye movement, gaze direction, and blinks can all be monitored by attaching a set of electrical sensors to the face around the eyes and measuring the electrooculogram (EOG) signals. However, this method is highly invasive and is not suitable for realistic driving conditions. Another example of monitoring body expression without video recorders is the use of sensors embedded in the driver’s seat to measure the body pressure distribution on the seat, from which certain body postures can be detected [11].

Physiological measures

Physiological measures can recognize the changes in driver’s arousal level [31]. Arousal is largely affected by the autonomous nervous system (ANS), which includes sympathetic and parasympathetic branches. The sympathetic branch

Today’s cameras are empowered with abilities beyond human vision, and thus can capture extremely subtle changes, such as HR using thermal imaging.

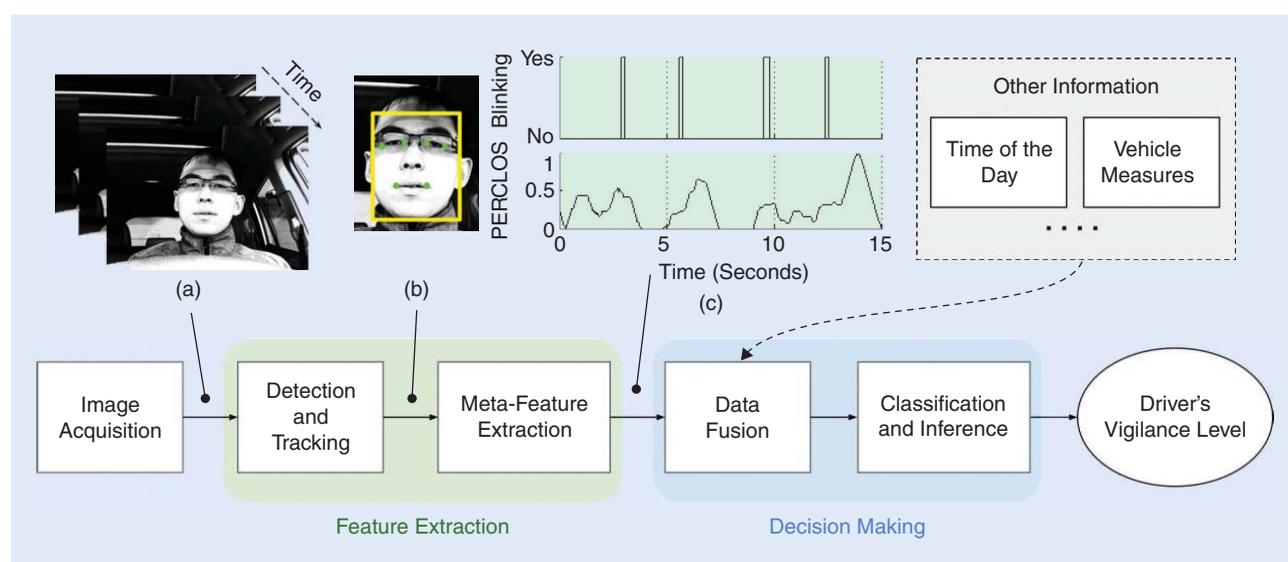


FIGURE 5. An example of driver monitoring pipeline using facial and body expression measures, depicting typical processing steps and intermediate outputs from them: (a) video frames acquired, (b) tracking of driver’s face (yellow rectangular box) and facial landmarks (green dots), and (c) example time series plot of two metafeatures: blinking (binary variable) and PERCLOS (scales from 0 to 1 representing eyes open to close).

generates an alerting response in stressful situations, such as a fight-or-flight response in an extreme case. This alert state can be recognized by increased breathing rate, accelerated HR, sweaty palms, and dilated pupils [32]. On the other hand, parasympathetic branch is mostly activated during relaxed situations, such as sleep periods. Parasympathetic activation leads to decreased breathing rate, HR, and blood pressure. Therefore, breathing rate, HR, and skin moisture can be used as indicators of the ANS's activity, which in turn can indicate the driver's arousal level and alertness. These indicators/responses can be measured using breathing, ECG, and GSR sensors. Two other commonly used measures in driver monitoring literature are EOG, which was explained in the previous section, and electroencephalogram (EEG), which monitors the brain's electrical activities and can be used to detect vigilance or mental workload level. Figure 4 illustrates some example physiological sensors.

Physiological measures will be particularly important with higher levels of automation in emerging self-driving vehicles. When a vehicle takes control (i.e., automated driving), there will be no driving performance measurement from the driver. However, physiological measures can still provide information on whether or not the driver is alert. This information together with facial/body expressions and measures of driver's current activity can be used as a deciding factor on how to transfer back the control from the vehicle to the driver. A short review of physiological measures is given next.

Skin moisture/conductance

GSR, also known as *electrodermal response*, has been commonly used to estimate driver's physiological state [33]. GSR is a measurement of the skin conductance controlled by the sympathetic nervous system through skin's sweat glands [34]. Highly demanding mental tasks can increase skin moisture and, hence, skin conductance. Electrodermal activity consists of two components: general tonic activity, which corresponds to the slower activities and background signal; and phasic activity, which corresponds to the faster changing signal elements. Parameters such as frequency, amplitude, latency, rise time, and recovery time of the electrodermal activities are among the measures that have been used in the literature. A comprehensive review of GSR measures is provided in [35].

Respiratory activities

The respiratory frequency is generally between 9–21 cycles/minute. During relaxation, this rate can become as low as six cycles/minute, whereas performing highly demanding mental tasks cause higher than normal frequencies [36]. Respiratory activities can be measured either through sensors that are connected to the person's mouth/nose, which are not suitable for driving, or using a breathing belt which measures

the changes in thoracic or abdominal circumference during respiration (Figure 4).

Eye movements

EOG monitors eye movements, hence eye gaze, by placing sensors around the eyes to measure the corneo-retinal potential between the front and the back of the eye. Furthermore, several eye-blink indicators such as blink rate, duration, and latency can be derived from EOG to analyze visual attention, mental workload, and drowsiness [31]. Despite EOG's promising results, current measurement techniques are highly intrusive during driving and cannot be used in practice.

Heart activities

HR and HR variability (HRV) can be used to monitor the effect of ANS on the heart. Heightened levels of task difficulty result in increased HR, whereas fatigue/drowsiness decreases HR. HRV is defined as the variation in the time interval between heartbeats, i.e., the beat-to-beat interval. In the context of driver monitoring, HRV is usually calculated by taking the Fourier transform of the HR signal. High frequencies of HRV's power spectrum (0.15–0.40 Hz) have been found to reflect parasympathetic activities, whereas

low frequencies (0.04–0.15 Hz) reflect both sympathetic and parasympathetic activities (see [37] and references therein). Studies have shown that changes in mental workload may affect HRV: decreases in midfrequency bands (0.07–0.14 Hz) have been associated with increases in mental effort (e.g., [39]). Another study also reported reduction in the high-frequency component of HRV and an increase in the low- to high-frequency ratio in more stressful situations [38]. HR can be easily measured using a single sensor that tracks heartbeats from the artery pulsation, or it can be extracted from ECG, which is a recording of the projection of the electric activities of the heart on the body surface. In a standard medical ECG device, ten electrodes (conductive pads) are attached to the person's chest, which is rather impractical and invasive for realistic driving conditions. Therefore, recently, there has been a growing interest to measure ECG using sensors embedded on the steering wheel (Texas Instruments; <http://www.ti.com/lit/pdf/tidu479>) or the driver's seat (www.medtees.com/content/ecg_seat_fact_sheet_2.pdf).

Brain activities

There exist two noninvasive portable measurement methods to monitor brain activities: functional NIR spectroscopy (fNIR) and EEG. In the fNIR method, the concentration of (de)oxygenated hemoglobin in different parts of the brain cortex is measured using NIR electromagnetic waves, based on which, active parts of the cortex are determined. fNIR is particularly helpful in monitoring increased arousal due to mental workload [40]. A review of the features that can be extracted from fNIR signals, and the required processing steps, is provided in [41].

It is worth mentioning that monitoring a driver's facial and body expressions also provides possible means for determining the driver's emotions.

In the EEG method, two (or three including the active ground) to more than 100 electrodes are located on the scalp, each measuring an aggregate of electric voltage fields from millions of neighboring neurons. EEG signals can be expressed in terms of a number of rhythmic activities, which are usually divided into the following frequency bands: Delta (< 4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–30 Hz), and Gamma (> 30 Hz) [42]. When no major cognitive or motor task is performed, large populations of neurons are synchronized and result in steady rhythmic activities. In contrast, during cognitive or motor tasks, the synchronization of these populations usually decreases (in some cases increases), which results in a decrease (or increase) in the power of corresponding oscillatory rhythms.

The waking EEG is mostly characterized by a desynchronized pattern, causing rapid, high-frequency waves in the beta range. When people are awake in a quiet relaxed state, the increased synchrony of underlying neural activity in nonstimulated brain regions results in patterns of alpha waves particularly at the posterior part of the brain. In driver monitoring systems, one of the most difficult tasks is to detect the onset of drowsiness, often referred to as *nonrapid eye movement (NREM)* sleep stage N1 [43]. However, in EEG signals, this transition can be easily identified by the changes of alpha and theta waves. If the driver passes this stage and enters the next NREM sleep stage (N2) [43], the eye movements will cease, the HR slows down, and body temperature decreases, preparing the body to enter deep sleep. At this stage, background EEG oscillations decrease below 5 Hz. Furthermore, these slow oscillations will be superimposed by periodic, transient EEG patterns called *sleep spindles* and *K-complexes*. These synchronization changes at different regions of the brain and their corresponding effects on the EEG characteristics in time/frequency/space domains can be used for driver state monitoring. Some examples of commonly used methods for extraction of discriminative features from EEG signals include: parametric spectral estimation (e.g., autoregressive/moving-average), nonparametric spectral estimation (e.g., Fourier/wavelet transform), bandpass filtering, and spatial filtering (e.g., common spatial patterns, independent component analysis, surface Laplacian filtering).

Figure 6 provides an example of the changes in an EEG feature (the ratio between powers in Alpha and Theta bands for T7-O1 bipolar EEG channel) and the driver's sleepiness level recorded in a simulator experiment. In Figure 6, both graphs in (a) and (b) illustrate the relationship between the changes in sleepiness level and the EEG signal. Note that the sleepiness scale is a subjective scale with limited granularity (five discrete levels) whereas EEG provides a nonsubjective continuous measure of sleepiness. As shown in the lower plot of Figure 6(a), corresponding to the 3-minute moving average window (dashed line), the brain wave power ratio shows an increasing trend with increasing sleepiness level, and vice

versa (except a few locations, such as “I” and “J”). A comprehensive review of various EEG processing techniques used to monitor driver's mental workload, fatigue, and drowsiness is provided in [31].

Finally, it is noteworthy that any facial activity or eye movement can cause interference in the EEG signals, which are commonly referred to as *artifacts*. For example, the blue highlighted parts in Figure 7 are the artifacts caused by eye blinks. The common practice in EEG analysis is to remove these artifacts to study the underlying brain activity. However, in the context of driver monitoring, these artifacts can provide side information regarding the facial activities, in particular eye movements. Therefore, one area with great potential for future research is to exploit EEG artifacts for extraction of features, such as blink duration/frequency, eye closures, and even eye gaze direction.

There is great potential for the development of new fusion techniques for driver monitoring.

Data fusion

Each of the three measurement categories discussed in the previous sections has its own pros and cons, which have to be taken into account in the design of smart driver monitoring systems. The main limitation of both vehicle-based measures and facial/body expressions is that they depend on factors that tend to manifest themselves at the late stages of fatigue or drowsiness (low arousal) or mental overload (high arousal), whereas physiological measures can provide early indicators of changes in arousal. However, since current technologies for physiological measurement require direct attachment of sensors to driver's body, these measures might be more intrusive than other categories. Similarly, some drivers may have concerns with having a camera analyzing their facial/body expressions while driving. If the driver's concerns regarding privacy, ease of use, and nonintrusiveness are taken into account in the design of new sensing technologies such that drivers embrace them, the combination of these three categories can provide a multifaceted driver monitoring system that can benefit from the advantages of each of these categories and provide higher accuracy and performance. Furthermore, information provided by intelligent transportation systems [44] as well as information about driver's interactions with his surroundings, such as conversation with other passengers, can also be utilized. Combining multiple sources of information calls for the development of new data fusion techniques. For example, [45] uses a fuzzy Bayesian framework to combine the information obtained from facial features, ECG, photoplethysmography, temperature, and a three-axis accelerometer to monitor driver drowsiness. Still, fusion techniques for driver monitoring systems are in their early stages and are not fully explored yet. Potential reasons include: 1) relatively recent success in computationally efficient real-time analysis of video data, 2) relatively recent advancements in nonintrusive measurement techniques for physiological signals, and 3) the complex nature of sensing technologies in driver monitoring that usually requires multidisciplinary

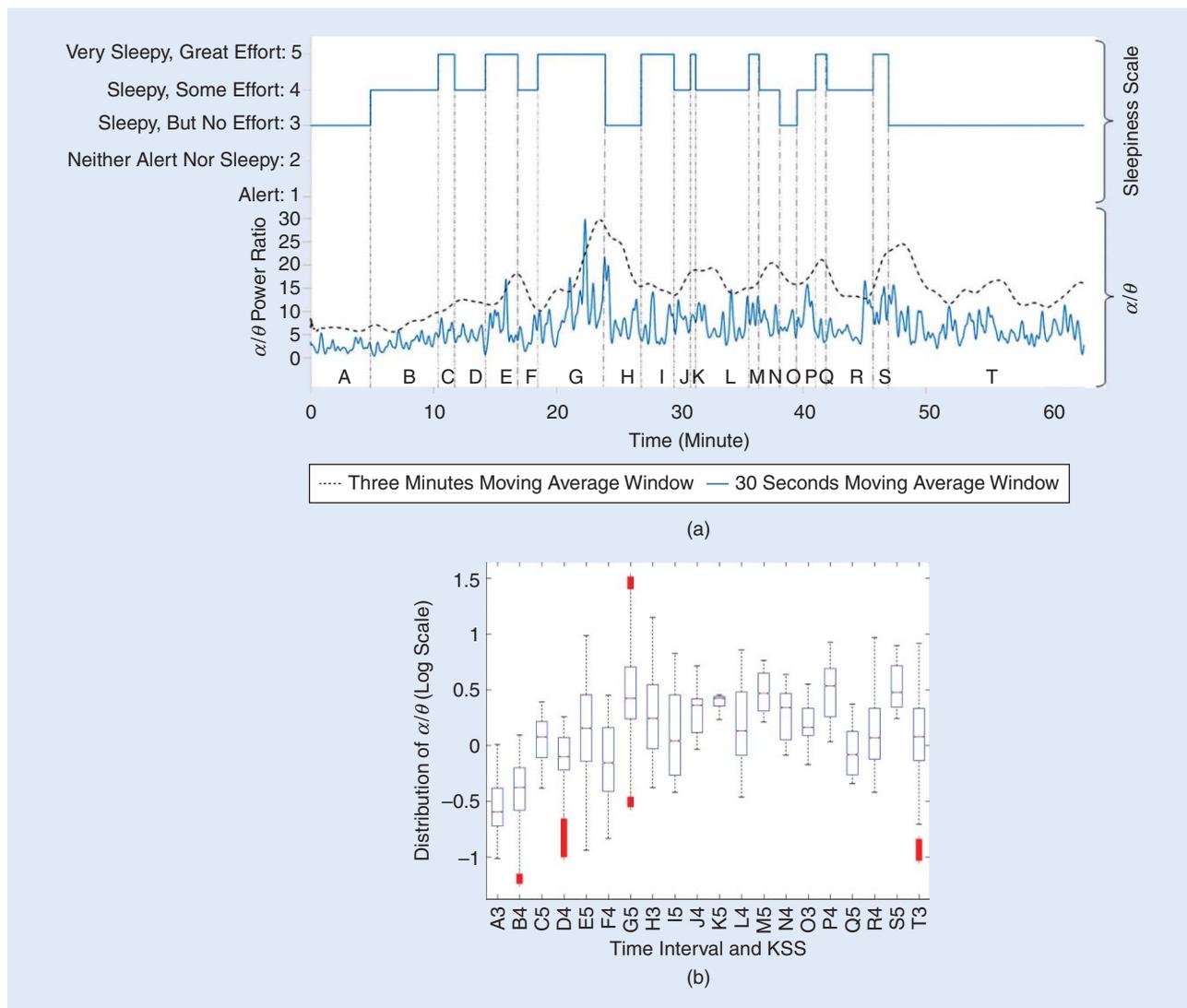


FIGURE 6. An example of changes in brain activity with respect to sleepiness during a one-hour driving simulator experiment: (a) At the top, a driver's sleepiness using a five-level scale is illustrated, and at the bottom, the changes in the ratio of Alpha band power over Theta band power is illustrated. The vertical dotted lines separate the time intervals during which the level of sleepiness is constant. (b) The box-plot representation of changes in the distribution of Alpha/Theta power ratio for different time intervals. The labels on the horizontal axis denote the time interval and its corresponding sleepiness scale value. KSS: Karolinska sleepiness scale.

collaboration between experts from several different areas for development of effective fusion techniques. Considering the extensive body of research on data fusion in the signal processing community, there is great potential for the development of new fusion techniques for driver monitoring.

Feedback mechanisms

This section presents a discussion of how different feedback timings can support overall safer driving, and highlights some design considerations for the various types of feedback.

Real-time feedback

Real-time feedback is provided during driving. For example, a driver may have a habit of driving too closely behind another vehicle (i.e., tailgating), may receive feedback about an unsafe

headway (e.g., warning through a collision avoidance system, verbal comment made by a passenger), and may adjust her headway accordingly. Studies have shown that real-time feedback is useful in notifying the driver about a potentially hazardous situation or an improper action that may lead to dangerous events (e.g., rear-end collision warning using time headway [46], and driver distraction feedback based on drivers' off-road eye glances [47]). However, humans have limited attentional resources and feedback provided during driving, like any other secondary task, may consume a certain amount of resources available to driving. For thorough discussions of the important role of attention in everyday driving, readers are referred to [48]. Also, [49] provides a detailed depiction of driver-feedback interactions, discussing how feedback can

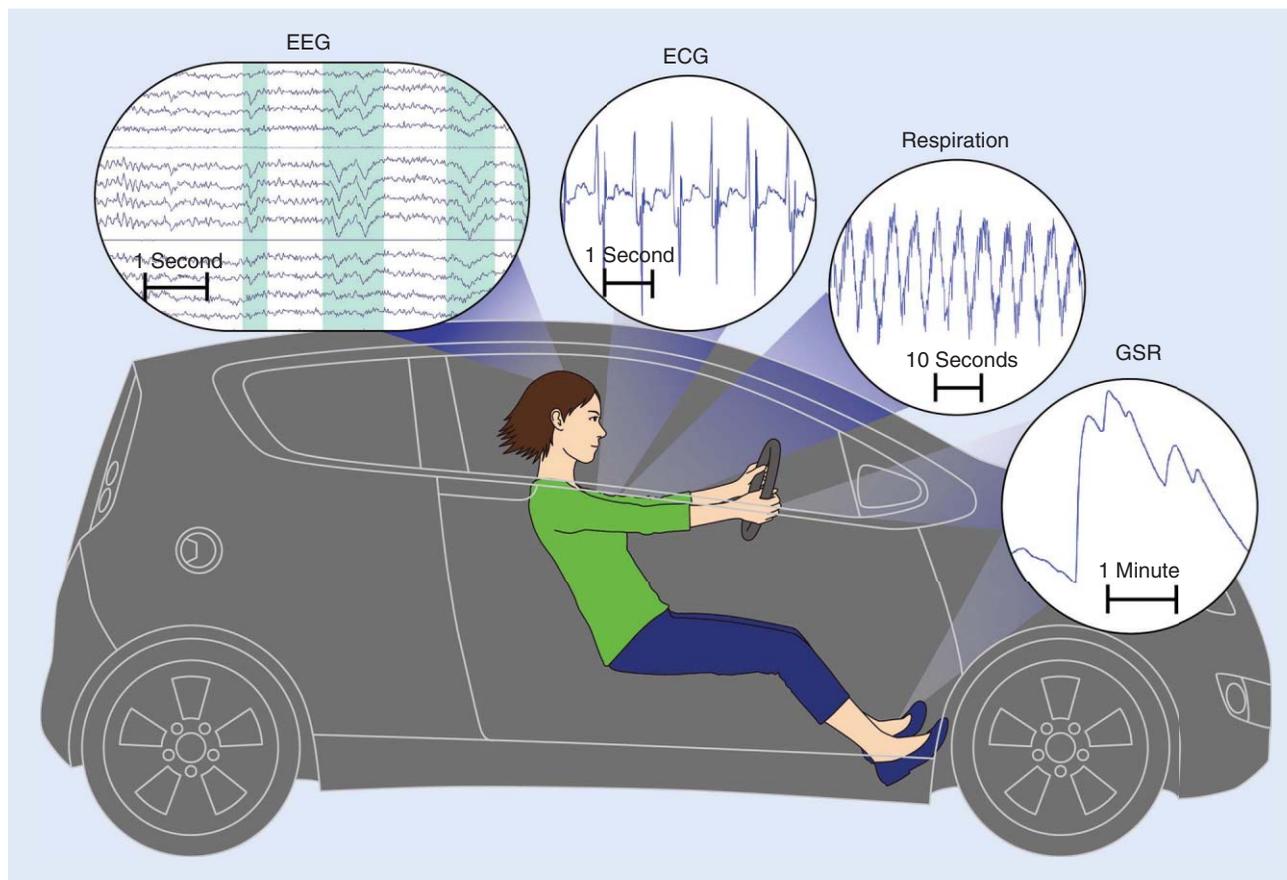


FIGURE 7. Example physiological measures recorded in a simulator study (setup as shown in Figure 4): EEG collected using a 14-channel wireless headset; Respiration activities recorded using an inductive respiratory belt; ECG recorded using three disposable gel-electrodes located on the driver's chest; and GSR recorded using two disposable gel-electrodes located on the sole of the driver's left foot.

both help and interfere with the allocation of limited attentional resources.

Overall, real-time feedback design needs to be concise, to communicate a salient message to the driver, and be delivered through an appropriate modality that will be perceived by the driver with minimal additional demand. For example, an auditory alert is likely to capture a driver's attention immediately even when the driving task is demanding [46]. However, if the driver is already cognitively overloaded, such feedback may become detrimental by interfering with the ongoing driving task. In such cases, delaying feedback, even for a few seconds, can mitigate the potential disruption while still providing some positive effects on driving.

Postdrive feedback

Postdrive feedback is provided after a trip. A driver's mental model about safe driving can guide her decisions and intentional behaviors. For example, if a driver perceives a significant increase in driving demands, she may allocate more attention to driving by pausing her conversation with a passenger. While the mental model of safe driving may be in part guided by formal learning, it is also updated continually according to

driving experience, feedback, and perceived social norms. For example, the memory of previous driving experiences may influence a driver's future behaviors. However, a study has estimated that 80% of near crashes are forgotten after only two weeks [50]. Thus, postdrive feedback provided after driving may facilitate long-term memory formation of critical driving incidents to motivate behavioral changes. The Chevy Malibu teen driver feature is an example for this type of feedback. Another example would be to use video replays of driving incidents (e.g., near crashes) to help strengthen the memory of such events, and thus influencing driver's mental model of safe driving.

SA and anticipatory driving

SA is the perception of critical elements, the comprehension of their meaning, and the projection of their status [51]. A high level of SA has been associated with expert operators in complex and dynamic environments, including aviation, healthcare, and automobile driving [51], [52]. For example, operating at a lower level of SA, a driver may not register the object(s) he sees, e.g., a stalled truck, to be a potential hazard, while another driver, operating at a high level of SA, may not only perceive and understand the stalled truck

as a potential hazard, but actively anticipate other drivers' behaviors around the stalled truck. Furthermore, SA in driving includes an internal awareness of the driver's own state and behaviors. For example, a driver aware of her drowsiness may decide to take a break from driving. SA may be best supported through real-time feedback, helping drivers track and understand critical elements on the road. A smart driver monitoring system that detects low arousal state of the driver may also encourage drivers to take a rest break or seek out some stimulation, such as music.

A high level of SA can be demonstrated through "anticipatory driving," a cognitive competence in identifying traffic situations through perception of cues, such that the vehicle can be positioned efficiently for potential upcoming changes in traffic [53]. Feedback intended to support anticipatory driving can focus on highlighting potential traffic conflicts ahead, and may be particularly useful for the younger inexperienced drivers, who have been systematically found to exhibit less efficient visual scanning patterns compared to experienced drivers [54], and are less proficient in interpreting the situation correctly [53], [55]. For more experienced drivers, a feedback system may still be helpful in directing their attention back to critical elements in the driving scene if they are distracted. A number of aids have been proposed in the literature to support anticipatory driving, e.g., highlighting of important cues [56], visual depiction of stopping distance [57], and visual feedback using an LED array for modulating brake pedal control [58]. However, interface designs in these studies remain largely at the conceptual stage, and have been evaluated in limited settings such as in driving simulators.

Concluding remarks and future directions

Recent developments in sensing technologies as well as advancements in our understanding of human factors affecting driving performance provide opportunities for the development of smart driver monitoring systems. These systems can use vehicle measures, facial/body expressions, and physiological signals to continuously monitor driver state and also sense the environment to provide feedback to drivers or take vehicle control if there is a need. In practice, however, only vehicle measures are widely used in commercially available driver monitoring systems. The use of facial/body expressions is limited due to technical challenges in acquisition and real-time processing of these signals. For physiological measures, the most important challenge has been nonintrusive signal acquisition, which has not been possible for signals such as EEG and ECG until recently. To fully exploit the information provided by these three measurement categories, the information extracted from different measures can be further processed using data fusion techniques. However, considering the fact that simultaneous real-time

acquisition and analysis of vehicle measures, facial/body expressions, and physiological data has not been feasible until recently, data fusion techniques have not yet been studied much in the context of smart driver monitoring.

As vehicles are becoming increasingly automated, it is becoming important for the vehicle to be able to monitor the driver's state to safely take control from the driver and transfer it back to the driver when there is a need. (For a comprehensive discussion of various types and levels of automation, refer to [59] and the Preliminary Statement of Policy by the National Highway Traffic Safety Administration, available at www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf.) In the former scenario, i.e., the transfer of control from the driver to the vehicle, driver state monitoring can play a crucial role in the timely and accurate engagement of automatic safety systems.

In such a scenario, an understanding of driver's intentions becomes paramount. For example, the activation of the turn signal can temporarily disable lane departure warnings. However, cases such as intentional tailgating are difficult or impossible to detect. Therefore, an open question is: Should there be situations in which autonomous systems are given the authority to override driver input? It is clear that vehicles should warn drivers of drowsiness or distraction, but it is unclear whether safety systems should prevent intentionally risky driving behaviors. In the latter scenario, i.e., the transfer of control from the vehicle to the driver, the role of driver monitoring systems is as important but less obvious. Once the vehicle gains control (e.g., due to driver drowsiness, or at driver's request), the driver monitoring system can actively monitor the driver by measuring facial/body expressions and physiological signals. Such continuous monitoring of the driver allows the system to determine when it is safe for the driver to regain vehicle control.

A major challenge in the research and development of smart driver monitoring systems is the unique interdisciplinary nature of this area, which requires a close collaboration between researchers and practitioners in both signal processing and human factors communities. The recent advancements in human factors research on driving safety can provide insights for incorporating efficient mitigation technologies into smart monitoring systems, using both real-time and postdrive feedback. This article presented a high-level interdisciplinary overview of the smart driver monitoring systems to facilitate such collaborations.

Acknowledgment

We acknowledge support by the Natural Science and Engineering Research Council through the strategic partnership grant titled DREAMS: Enhancing Driver Interaction with Digital Media through Cognitive Monitoring.

As vehicles are becoming increasingly automated, it is becoming important for the vehicle to be able to monitor the driver's state to safely take control from the driver and transfer it back to the driver when there is a need.

Authors

Amirhossein S. Aghaei (aghaei@ece.utoronto.ca) received his master's and Ph.D. degrees in electrical and computer engineering from the University of Toronto in 2008 and 2014, respectively. He was an industrial postdoctoral researcher at the University of Toronto and Qualcomm Canada Inc. from 2014 to 2016, after which he joined Microsoft. His fields of expertise include driver state monitoring, physiological signal processing, brain-computer interfacing, machine learning, and big data analytics.

Birsen Donmez (donmez@mie.utoronto.ca) received her B.S. degree in mechanical engineering from Bogazici University in 2001. She received an M.S. degree in 2004 and her Ph.D. degree in 2007, both in industrial engineering, and an M.S. degree in statistics in 2007, all from the University of Iowa. She is currently an associate professor at the University of Toronto in the Mechanical and Industrial Engineering Department. She is the Canada Research Chair in Human Factors and Transportation. Her research focuses on improving human behavior and performance in multitask situations. She received the inaugural Stephanie Binder Young Professional Award from the Human Factors and Ergonomics Society in 2014 and an Ontario Early Researcher Award in 2015. She serves on multiple Transportation Research Board committees and is an associate editor of *IEEE Transactions on Human-Machine Systems*.

Cheng Chen Liu (cheng.liu@mail.utoronto.ca) received her B.A.Sc. degree in computer engineering from the University of Toronto, where she is currently an M.A.Sc. student with the Electrical and Computer Engineering Department. Her research interests include computer vision and machine learning, focusing on video-based driver monitoring.

Dengbo He (dengbo.he@mail.utoronto.ca) received his bachelor's degree in mechanical engineering from Hunan University, Changsha, China, in 2012 and his M.S. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2016. He is currently a Ph.D. student at the University of Toronto in the Mechanical and Industrial Engineering Department. His research interests include vehicle dynamics and control, human factor, and driver behavior.

George Liu (geo.liu@utoronto.ca) received an honors bachelor of arts degree from the University of Toronto and is currently a master's student in industrial engineering at the University of Toronto and holds a second master's degree in urban transportation planning. His research focuses on driver feedback design and human factors analysis.

Konstantinos N. Plataniotis (kostas@ece.utoronto.ca) is a professor and the Bell Canada Chair in Multimedia with the Department of Electrical and Computer Engineering, University of Toronto, Ontario, Canada. He is a Registered Professional Engineer in Ontario and a fellow of the Engineering Institute of Canada. He has served as the edi-

tor-in-chief of *IEEE Signal Processing Letters* and the technical cochair of the 2013 IEEE International Conference in Acoustics, Speech, and Signal Processing. He is the IEEE Signal Processing Society vice president for membership (2014–2016) and the general chair of the 2018 IEEE International Conference on Image Processing. He received the IEEE Canada Engineering Educator Award for contributions to engineering education and inspirational guidance of graduate students. His research interests include statistical signal processing, knowledge and digital media design, multimedia systems, biometrics, image processing, biomedical signal processing, and pattern recognition. He is a Fellow of the IEEE.

Huei-Yen Winnie Chen (win.chen@mail.utoronto.ca) received her B.A.Sc. and M.A.Sc. degrees in systems design engineering from the University of Waterloo in 2005 and 2007, respectively, and her Ph.D. degree in industrial engineering from the University of Toronto in 2013. She is currently a postdoctoral researcher at the University of Toronto in the Mechanical and Industrial Engineering Department. Her research interests are understanding and improving human-technology interactions in dynamic, complex systems using advanced analytical methods. Her research areas have included applications in transportation, health care, defense, and aerospace.

Zohreh Sojoudi (zohreh.sojoudi@mail.utoronto.ca) received her bachelor's degree in clinical psychology from the University of Tehran, Iran, in 2012. She is currently in her fourth year of studying cognitive science at the University of Toronto, Faculty of Arts and Science. Her research focuses on behavioral driver monitoring and cognitive neuroscience.

References

- [1] M. Peden, R. Scurfield, D. Sleet, D. Mohan, A. A. Hyder, E. Jarawan, and C. Mathers "World report on road traffic injury prevention," World Health Organization, Geneva, 2004.
- [2] "Overview: 2013 data," National Highway Traffic Safety Administration, Washington, DC, Tech. Rep. DOT HS 812 169, 2015.
- [3] S. Singh, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," National Highway Traffic Safety Administration, Washington, DC, Tech. Rep. DOT HS 812 115, 2015.
- [4] C. Owsley, K. Ball, G. McGwin Jr., M. E. Sloane, D. L. Roenker, M. F. White, and E. T. Overley, "Visual processing impairment and risk of motor vehicle crash among older adults," *JAMA*, vol. 279, no. 14, pp. 1083–1088, 1998.
- [5] H. A. Deery, "Hazard and risk perception among young novice drivers," *J. Safety Res.*, vol. 30, no. 4, pp. 225–236, 2000.
- [6] B. A. Jonah, "Sensation seeking and risky driving: A review and synthesis of the literature," *Accid. Anal. Prev.*, vol. 29, no. 5, pp. 651–665, Sept. 1997.
- [7] M. Haglund and L. Åberg, "Speed choice in relation to speed limit and influences from other drivers," *Transp. Res. F Traffic Psychol. Behav.*, vol. 3, no. 1, pp. 39–51, Mar. 2000.
- [8] F. Barbé, J. Pericas, A. Munoz, L. Findley, J. M. Anto, A. G. Agusti, and M. de Luc Joan, "Automobile accidents in patients with sleep apnea syndrome: an epidemiological and mechanistic study," *Am J. Respir. Crit. Care Med.*, vol. 158, no. 1, pp. 18–22, 1998.
- [9] R. W. Proctor and T. Van Zandt, *Human Factors in Simple and Complex Systems*, 2nd ed. Boca Raton, FL: CRC Press, 2008.
- [10] J. F. Coughlin, B. Reimer, and B. Mehler, "Monitoring, managing, and motivating driver safety and well-being," *IEEE Pervasive Comput.*, vol. 10, no. 3, pp. 14–21, 2011.

- [11] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, 2011.
- [12] T. Ersal, H. J. Fuller, O. Tsimhoni, J. L. Stein, and H. K. Fathy, "Model-based analysis and classification of driver distraction under secondary tasks," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 692–701, 2010.
- [13] M. Bayly, K. L. Young, and M. A. Regan, "Sources of distraction inside the vehicle and their effects on driving performance," in *Driver Distraction: Theory, Effects, and Mitigation*, M. A. Regan, J. D. Lee, and K. L. Young, Eds. Boca Raton, FL: CRC Press, 2008, ch. 12, pp. 191–214.
- [14] K. Yabuta, H. Iizuka, T. Yanagishima, Y. Kataoka, and T. Seno, "The development of drowsiness warning devices," SAE Technical Paper, Tech. Rep., paper number 856043, 1985.
- [15] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting driver drowsiness based on sensors: A review," *Sensors*, vol. 12, no. 12, pp. 16 937–16 953, 2012.
- [16] F. Tango, C. Calefato, L. Minin, and L. Canovi, "Moving attention from the road: A new methodology for the driver distraction evaluation using machine learning approaches," in *Proc. IEEE Conf. Human Syst. Interactions*, 2009, pp. 596–599.
- [17] D. Sandberg, T. Åkerstedt, A. Anund, G. Kecklund, and M. Wahde, "Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 1, pp. 97–108, 2011.
- [18] T. Åkerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual," *Int. J. Neurosci.*, vol. 52, no. 1–2, pp. 29–37, 1990.
- [19] C. C. Liu, S. G. Hosking, and M. G. Lenné, "Predicting driver drowsiness using vehicle measures: Recent insights and future challenges," *J. Safety Res.*, vol. 40, no. 4, pp. 239–245, 2009.
- [20] D. Dinges and R. Grace, "Perclous: A valid psychophysiological measure of alertness assessed by psychomotor vigilance," National Highway Traffic Safety Administration, Washington, DC, Tech. Rep. FHWA-MCRT-98-006, 1998.
- [21] M.-H. Sigari, M.-R. Pourshahabi, M. Soryani, and M. Fathy, "A review on driver face monitoring systems for fatigue and distraction detection," *Int. J. Sci. Adv. Technol.*, vol. 64, pp. 73–100, Mar. 2014.
- [22] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-time system for monitoring driver vigilance," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 63–77, 2006.
- [23] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 1–511.
- [24] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness," *IEEE Trans. Intell. Transport Syst.*, vol. 11, no. 2, pp. 300–311, 2010.
- [25] Q. Ji, Z. Zhu, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1052–1068, 2004.
- [26] R. Oyini Mbouna, S. G. Kong, and M.-G. Chun, "Visual analysis of eye state and head pose for driver alertness monitoring," *IEEE Trans. Intell. Transport Syst.*, vol. 14, no. 3, pp. 1462–1469, 2013.
- [27] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 8, pp. 1418–1426, 2007.
- [28] U. Agrawal, S. Giripunjé, and P. Bajaj, "Emotion and gesture recognition with soft computing tool for drivers assistance system in human centered transportation," in *Proc. IEEE Int. Conf. Systems, Man, and Cybernetics*, 2013, pp. 4612–4616.
- [29] C. Yan, B. Zhang, and F. Coenen, "Driving posture recognition by convolutional neural networks," in *Proc. Int. Conf. Natural Computing*, 2015, pp. 680–685.
- [30] E. Ohn-Bar, A. Tawari, S. Martin, and M. M. Trivedi, "On surveillance for safety critical events: In-vehicle video networks for predictive driver assistance systems," *Comput. Vis. Image Understand.*, vol. 134, pp. 130–140, May 2015.
- [31] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neurosci. Biobehav. Rev.*, vol. 44, pp. 58–75, July 2014.
- [32] B. Mehler, B. Reimer, J. Coughlin, and J. Dusek, "Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers," *Transp. Res. Rec. J.*, vol. 2138, pp. 6–12, 2009.
- [33] B. Mehler, B. Reimer, and J. F. Coughlin, "Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task an on-road study across three age groups," *Hum. Factors*, vol. 54, no. 3, pp. 396–412, 2012.
- [34] D. Taylor, "Drivers' galvanic skin response and the risk of accident," *Ergonomics*, vol. 7, no. 4, pp. 439–451, 1964.
- [35] W. Boucsein, *Electrodermal Activity*. New York: Springer Science, 2012.
- [36] S. H. Fairclough and L. Mulder, "Psychophysiological processes of mental effort investment," in *How Motivation Affects Cardiovascular Response: Mechanisms and Applications*, R. Wright and G. Gendolla, Eds. Washington, DC: American Psychological Assoc., 2011, pp. 61–76.
- [37] C. Zhao, M. Zhao, J. Liu, and C. Zheng, "Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator," *Accid. Anal. Prev.*, vol. 45, pp. 83–90, Mar. 2012.
- [38] N. Hjortskov, D. Rissén, A. K. Blangsted, N. Fallentin, U. Lundberg, and K. Sjøgaard, "The effect of mental stress on heart rate variability and blood pressure during computer work," *Eur. J. Appl. Physiol.*, vol. 92, no. 1-2, pp. 84–89, 2004.
- [39] K. Ryu and R. Myung, "Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic," *Int. J. Ind. Ergon.*, vol. 35, no. 11, pp. 991–1009, 2005.
- [40] E. M. Peck, D. Afergan, B. F. Yuksel, F. Faloutsos, and R. J. Jacob, "Using fNIRS to measure mental workload in the real world," in *Advances in Physiological Computing*. London, U.K.: Springer, 2014, pp. 117–139.
- [41] S. Tak and J. C. Ye, "Statistical analysis of fNIRS data: A comprehensive review," *NeuroImage*, vol. 85, pp. 72–91, 2014.
- [42] A. S. Aghaei, M. S. Mahanta, and K. N. Plataniotis, "Separable common spatio-spectral patterns for motor imagery bci systems," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 15–29, 2016.
- [43] H. Schulz, "Rethinking sleep analysis," *J. Clin. Sleep Med.*, vol. 4, no. 2, pp. 99–103, 2008.
- [44] N.-E. El Faouzi, H. Leung, and A. Kurian, "Data fusion in intelligent transportation systems: Progress and challenges—a survey," *Inf. Fusion*, vol. 12, no. 1, pp. 4–10, 2011.
- [45] B.-G. Lee and W.-Y. Chung, "A smartphone-based driver safety monitoring system using data fusion," *Sensors*, vol. 12, no. 12, pp. 17 536–17 552, 2012.
- [46] J. J. Scott and R. Gray, "A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving," *Hum. Factors*, vol. 50, no. 2, pp. 264–275, 2008.
- [47] B. Donmez, L. N. Boyle, and J. D. Lee, "Safety implications of providing real-time feedback to distracted drivers," *Accid. Anal. Prev.*, vol. 39, no. 3, pp. 581–590, 2007.
- [48] J. Engström, T. Victor, and G. Markkula, "Attention selection and multitasking in everyday driving: A conceptual model," in *Driver Distraction and Inattention: Advances in Research and Countermeasures*, M. Regan, J. Lee, and T. Victor, Eds. Surrey, U.K.: Ashgate, 2013, pp. 27–54.
- [49] J. Feng and B. Donmez, "Design of effective feedback: Understanding driver, feedback, and their interaction," presented at the 7th Int. Driving Symp. Human Factor in Driver Assessment, Training, and Vehicle Design, New York, 2013.
- [50] P. Chapman and G. Underwood, "Forgetting near-accidents: the roles of severity, culpability and experience in the poor recall of dangerous driving situations," *Appl. Cogn. Psychol.*, vol. 14, no. 1, pp. 31–44, 2000.
- [51] L. Gugerty, "Situation awareness in driving," in *Handbook for Driving Simulation in Engineering, Medicine and Psychology*, D. L. Fisher, M. Rizzo, J. K. Carid, and J. D. Lee, Eds. 2011, vol. 1, pp. 1–25.
- [52] M. R. Endsley, "Measurement of situation awareness in dynamic systems," *Hum. Factors*, vol. 37, no. 1, pp. 65–84, 1995.
- [53] P. Stahl, B. Donmez, and G. Jamieson, "Anticipation in driving: The role of experience in the efficacy of pre-event conflict cues," *IEEE Trans. Human-Mach. Syst.*, vol. 44, no. 5, pp. 603–613, 2014.
- [54] G. Underwood, P. Chapman, N. Brocklehurst, J. Underwood, and D. Crundall, "Visual attention while driving: sequences of eye fixations made by experienced and novice drivers," *Ergonomics*, vol. 46, no. 6, pp. 629–646, 2003.
- [55] L. Jackson, P. Chapman, and D. Crundall, "What happens next? Predicting other road users' behaviour as a function of driving experience and processing time," *Ergonomics*, vol. 52, no. 2, pp. 154–164, Feb. 2009.
- [56] P. Stahl, B. Donmez, and G. A. Jamieson, "Supporting anticipation in driving through attentional and interpretational in-vehicle displays," *Accid. Anal. Prev.*, vol. 91, pp. 103–113, Jun. 2016.
- [57] M. Tonnis, C. Lange, and G. Klinker, "Visual longitudinal and lateral driving assistance in the head-up display of cars," in *Proc. IEEE and ACM Int. Symp. Mixed and Augmented Reality*, 2007, pp. 91–94.
- [58] F. Laquai, F. Chowanetz, and G. Rigoll, "A large-scale LED array to support anticipatory driving," in *Proc. IEEE Int. Conf. Systems, Man, and Cybernetics*, 2011, pp. 2087–2092.
- [59] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Trans. Syst., Man Cybern., A*, vol. 30, no. 3, pp. 286–297, 2000.

Fuliang Weng, Pongtep Angkititrakul,
Elizabeth E. Shriberg, Larry Heck,
Stanley Peters, and John H.L. Hansen

Conversational In-Vehicle Dialog Systems

The past, present, and future

Automotive technology rapidly advances with increasing connectivity and automation. These advancements aim to assist safe driving and improve user travel experience. Before the realization of a full automation, in-vehicle dialog systems may reduce the driver distraction from many services available through connectivity.

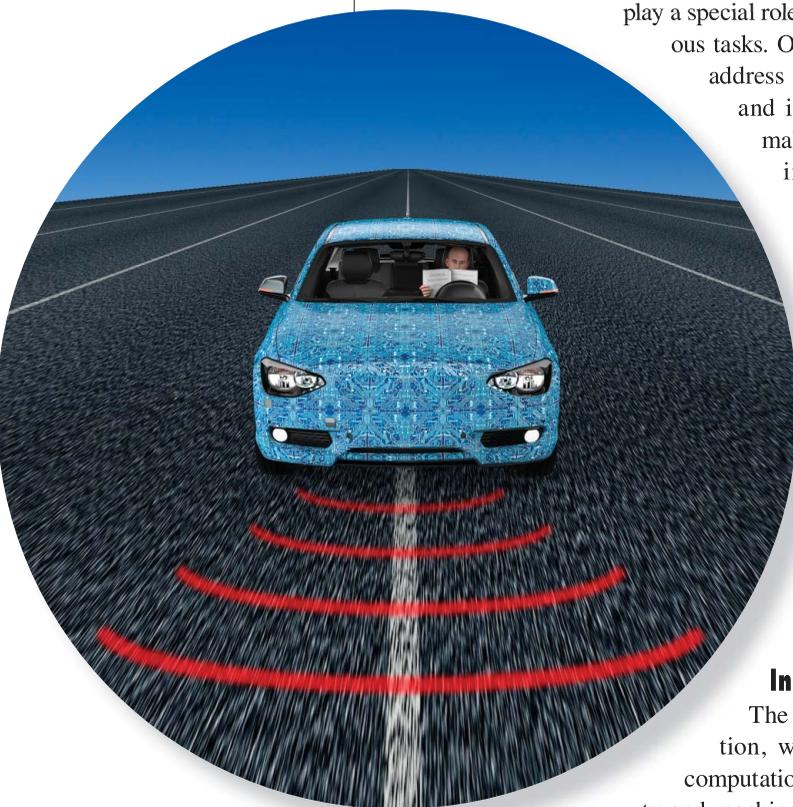
Even when a full automation is realized, in-vehicle dialog systems still play a special role in assisting vehicle occupants to perform various tasks. On the other hand, in-vehicle use cases need to address very different user conditions, environments, and industry requirements than other uses. This makes the development of effective and efficient in-vehicle dialog systems challenging; it requires multidisciplinary expertise in automatic speech recognition, spoken language understanding, dialog management (DM), natural language generation, and application management, as well as field system and safety testing. In this article, we review research and development (R&D) activities for in-vehicle dialog systems from both academic and industrial perspectives, examine findings, discuss key challenges, and share our visions for voice-enabled interaction and intelligent assistance for smart vehicles over the next decade.

Introduction

The automotive industry is undergoing a revolution, with progress in technologies ranging from computational power and sensors to Internet connectivity and machine intelligence. Expectations for smart vehicles are also changing, due to an increasing need for frequent engagement with work and family, and also due to technological progress of products in other areas, such as mobile and wearable devices.

This article will address human-machine interaction (HMI) and advanced intelligent assistance during the use of smart vehicles with a focus on voice-based technology. Speech is a primary means for human-to-human communication, capable of conveying rich content, colorful emotion, and human intelligence. Therefore, it is the most suitable

SKY IMAGE LICENSED BY GRAPHIC STOCK
©ISTOCKPHOTO.COM/POSTERIORI



Digital Object Identifier 10.1109/MSP.2016.2599201
Date of publication: 4 November 2016

driver-vehicle modality, allowing drivers to keep their eyes on the road and their hands on the wheel. A strong user need for a true voice-enabled intelligent assistance will soon make it possible to perform many daily tasks while driving, which previously could not be achieved safely. During the transition from highly manual vehicle operation to highly automated operation, moving from simple dialog systems to sophisticated driving assistance is a major trend, and both challenges and opportunities arise.

Review of past major activities

The results from well-designed driver data collections reveal useful insights into HMI. Innovations from R&D projects as well as related assistance systems provide a good foundation for future developments. An overview for past major activities is given next.

In-vehicle dialog data collection

In the past two decades, vehicles equipped with both human and vehicle sensors have been deployed to collect realistic data on drivers, vehicles, and driving environments for the development of human-centered technologies. Among the in-vehicle data, speech corpora were collected for developing in-vehicle automatic speech recognition (ASR) and spoken dialog systems. Speech and audio data collections in cars are necessary to capture specific driver speech patterns and in-vehicle noise characteristics. For ASR, collected data can also be used to expand natural language coverage in recognition grammars. Video and data from a variety of sensors (e.g., haptic input, brake and gas pedal pressure, steering wheel angle, vehicle position, etc.) were synchronously collected to capture multimodal interactions between the driver and vehicle, as well as driver's states and cognitive loads. Started in 1998, the SPEECHDAT-CAR effort was the first international program to form a multilanguage speech corpus for automotive applications with 23 research groups from nine European countries and one U.S. site. A number of additional large-scale automotive speech data collection were completed [1]–[9]. Some of them are fully or partially open to the public for research purposes; others are privately owned by companies for product development. To study naturalistic conversational speech during driving, most data collection experiments deployed a Wizard of Oz (WoZ) technique [10], typically involving a human navigator (wizard) sitting in a separate room to simulate a dialog system. Another common methodology is to collect actual driving data with event data recorders. In general, these systems only record specific vehicle information over a short period of time during a crash and usually do not include audio and video data. However, these data are very useful for investigating accident causes and designing interfaces.

A common finding among most data collection efforts validated very significant speaker variability across different driving conditions and maneuver operations. CU-Move focused on

in-car noise conditions across different cabin structures. It found that having the windows open has more effect on the recognition accuracy than increasing the vehicle speed. AVICAR showed that combining both audio and video information significantly improves the speech recognition over the audio-only information under noisy conditions but is less beneficial in quiet conditions. SPEECHDAT-CAR reported that about 50% of collected speech data in vehicle is from speaker noise such as breathing, and coughing, while mispronunciation and incomprehensible speech could contribute up to 5% of data. From the corpus of the Center for Integrated Acoustic Information Research, it found that the number of fillers, hesitations, and slips per utterance unit was 0.31, 0.06, and 0.03, respectively. The Conversational Helper for Automated Tasks (CHAT) data collection shows that nearly 30% of proper names were partial names [9], and disfluent and distracted speech was prevalent [10].

Key findings from past R&D projects for in-vehicle use cases

Over the past decade, a number of publicly funded projects specifically address in-vehicle use cases, e.g., Virtual Intelligent Codriver (VICO), Tools for Ambient Linguistic Knowledge (TALK), and CHAT [9]. In the automotive industry, it is well known that driver distraction is a major source of fatal traffic accidents; thus, minimizing driver distraction in automotive HMI has been a key goal. Driving-related activities have been classified as 1) the primary tasks of vehicle maneuvering, which require one's eyes on the road for navigation and obstacle avoidance, one's hands on the wheel for steering, and one's feet on the pedals for moving or stopping the vehicle and 2) secondary tasks, such as maintaining occupant comfort and accessing infotainment. While these primary

Speech is a primary means for human-to-human communication, capable of conveying rich content, colorful emotion, and human intelligence.

tasks have been stable, increasingly, the secondary tasks are becoming richer and more diverse, especially as Internet connectivity and availability of information and services have become common. Manual controls for such sophisticated secondary tasks such as buttons and knobs are difficult for drivers to safely operate and are inadequate for complex tasks. As speech technology has matured, it has become an increasingly natural choice due to its convenience and robust capabilities. The CHAT-Auto addresses conversational dialog systems-related challenges, such as voice-operated multitasking with imperfect speech input and communicating a large amount of information content to drivers with limited cognitive capacity. It further demonstrated the feasibility of speech technology for representative vehicle use cases such as navigation by destination entry, point-of-interest (POI) finding, and music search. The European Union (EU) VICO project initiated voice-enabled assistance system prototyping for navigation-related tasks such as POIs and address input. The EU TALK project covered on adaptive multimodal and multilingual dialog systems with a main focus on reusability by allowing for the core dialog system to be separated from specific

applications, languages, or modalities. It also intends to show the capability of learning dialog strategies to improve communication with drivers when an ASR engine make errors. A statistical framework of a partially observable Markov decision process was used to address uncertainty in speech recognition results [11]. Despite different foci of these research projects, they share many important building blocks in an advanced speech dialog system, as shown in Figure 1.

Influential automotive infotainment HMI products on the market include BMW iDrive, Mercedes COMAND, Audi MMI, Ford Sync, GM CUE, Lexus Enform, and Acura AcuraLink. In general, these HMI products are used to control multiple in-car functionalities including navigation, entertainment, multimedia, telephony, vehicle dynamics, and so on. Existing HMI input technologies fall into two categories: haptic based and voice based. Of haptic-based input methods, the two most common are control knobs and touch screens. German cars, e.g., BMW iDrive and Audi MMI, typically use control knobs. While these control knobs are simple to operate, it often takes multiple steps to access a function deep within a menu. Thus, these interfaces impose cognitive demands on drivers, requiring them to memorize where menu items are located and navigate through the menu while driving. In contrast, Japanese cars tend to use touch screens. Navigating touch-screen interfaces can be more intuitive and efficient, but they impose a spatial limitation on drivers, as they must reach out to different areas of touch screen to access predesignated buttons in a shallower but still hierarchical menu structure. As both knobs and touch screens require hand-eye coordination, haptic interfaces pose inherent

safety risks. Research shows that glancing away from the road for two seconds or longer may increase the risk of an accident from four to 24 times [12]. Such research evidence has led some original equipment manufacturers to prohibit the use of some of interfaces with hierarchical menu structures during driving. The voice-based systems have been shown to reduce driver look-away time and lessen spatial and cognitive demands on the driver. It is expected that recent advancement in speech recognition technology; other sensing technologies such as gesture on touch pad,

on steering wheel surface, and in air; as well as deeper modality fusion technologies would greatly liberate the designers so that many new devices and services can be incorporated without many physical constraints in the cockpit [13], [14].

In the past two decades, vehicles equipped with both human and vehicle sensors have been deployed to collect realistic data on drivers, vehicles, and driving environments for the development of human-centered technologies.

Related intelligent assistant technologies

Likewise, many efforts were devoted to the field of general-purpose voice-enabled intelligent personal assistants (IPAs). The U.S. Defense Advanced Research Projects Agency (DARPA) funded a number of key IPA-related programs. The

DARPA Communicator project was a major early effort in developing robust multimodal speech-enabled dialog systems with advanced conversational capabilities for engaging human users in mixed-initiative interactions. From 2003 to 2008, DARPA funded the Cognitive Agent that Learns and Organizes (CALO) project. CALO attempted to build an integrated system capable of true artificial intelligence's (AI's) key features such as the ability to learn and adapt in adverse situations, and comfortably interact with humans [15]. The CALO project had

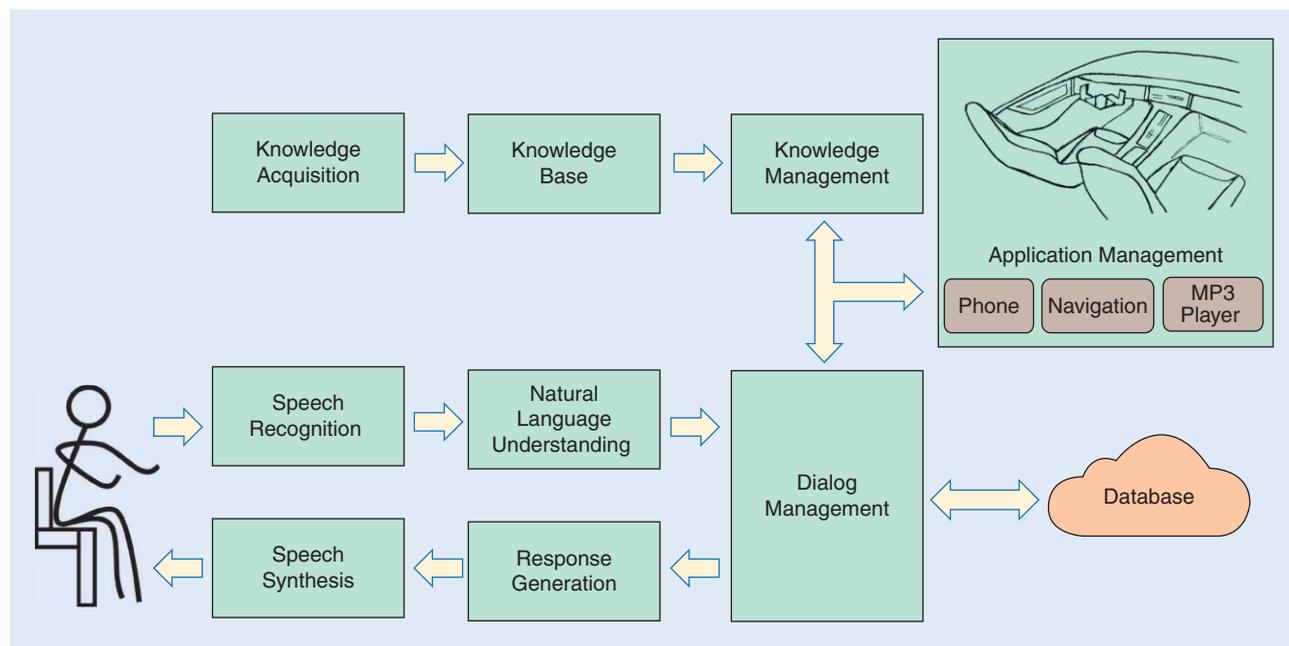


FIGURE 1. A generic block diagram for a typical in-vehicle spoken dialog system.

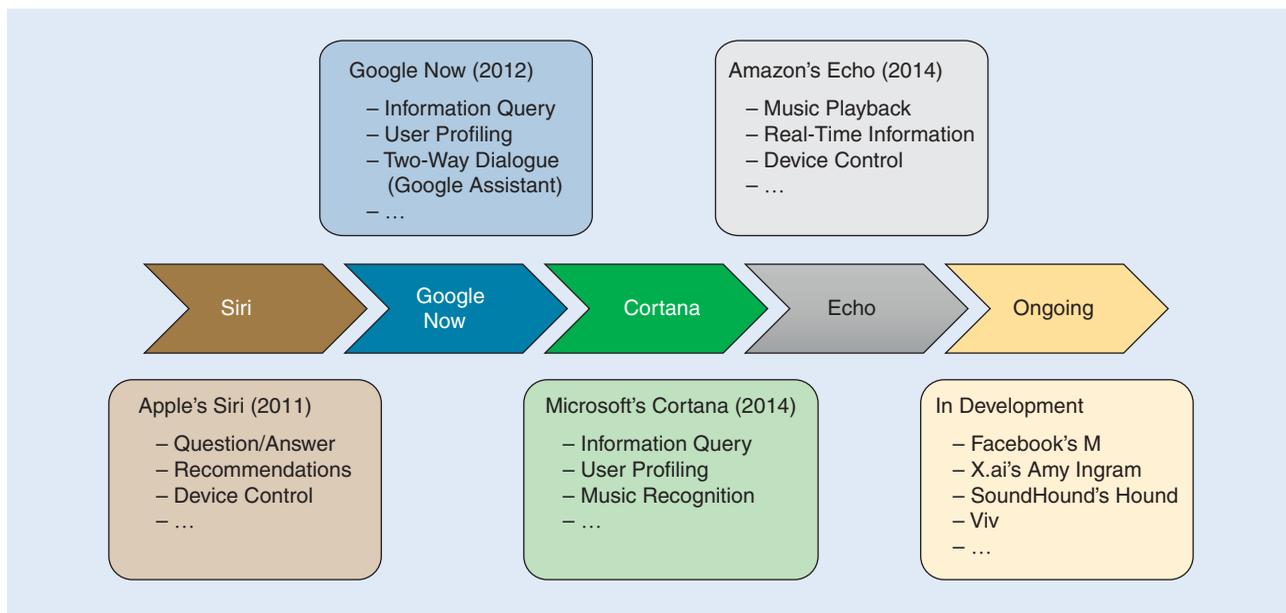


FIGURE 2. The development of voice-enabled IPAs.

a number of spin-offs, most notably SRI International's Siri intelligent software assistant.

Nuance speech technologies are behind many successful products in different markets. Due to limitations in speech understanding accuracy and coverage, early Nuance speech dialog systems in call-center applications strictly followed a visual menu design, rather than typical human interaction patterns, leading to low user satisfaction in many cases.

After the acquisition of Siri in 2011, Apple's launch of the IPA on the iPhone marked a turning point in the mass acceptance of speech technologies. With the massive computational power available through the cloud, more applications and AI technologies started to be integrated into dialog systems; see Figure 2. Consequently, IPAs were developed that allow users to operate devices, access information, and manage personal tasks in a much richer way. Notable IPAs include Apple Siri, Google Assistant, Microsoft Cortana, Amazon Echo, IBM Watson, and Baidu. In addition, the major players in the field are building up developer platforms around these IPAs to enable their own AI ecosystems. However, these systems typically do not have much dialog capability and, in most cases, focus on single-turn question-answers (Q&A) and simple actions. As a contrast, text-based chatbots from Facebook, Google, and others make use of dialog technologies in automating services via Messenger, bypassing the dependency on speech technology.

The rapid increase of high-quality cloud-based IPAs has been partly attributed to recent advances in deep learning technologies—especially deep neural networks (DNNs) [16]. With the exception of speaker recognition in the late 1990s [17], deep learning methods have only recently surpassed hidden

Markov with Gaussian mixture model (GMM-HMM)-based approaches in speech recognition performance. With powerful computing resources and advances in learning algorithms beyond early approaches [17], [18], hidden Markov models with DNN (DNN-HMM)-based approaches are able to reduce error rates in speech recognition by 50% [18]. Various studies have been conducted on noise robustness with DNN. Some focus on using environmental information via feature enhancement [19]–[22]. Others include different conditions, such as model noise, reverberation, or speaker variation, in an end-to-end speech recognition training with a recurrent neural network [23]. The latter contrasts with traditional approaches of separating ASR optimization into front-end signal processing and back-end model training. Effective front-end approaches have

been developed to address acoustic echo, background noises, and multiple sound sources with microphone array processing [24]–[26]. The improved recognition accuracy significantly benefits the usability and adoption of the general IPAs.

The automotive industry has taken two different approaches to address the arrival of cloud-based, voice-enabled assistance systems from major IT companies. The shallow integration approach leverages in-

vehicle microphones, loudspeakers, control buttons, or a head unit screen to enable mobile devices to synchronize the systems' look and feel. This approach is represented by the MirrorLink standard from Car Connectivity Consortium, Apple's CarPlay, and Google's Android Auto. The deep integration approach requires the integration of the embedded and cloud systems at a component level. Recent experiments have shown that combining ASR results from embedded and cloud-based engines can reduce word error rate by up to almost 30% for

The rapid increase of high-quality cloud-based IPAs has been partly attributed to recent advances in deep learning technologies—especially deep neural networks.

Table 1. Major technical challenges for the in-vehicle dialog system.

Dialog System/Components	Features	Challenges
ASR	Noise robustness	Robust ASR under different driving conditions with different types of noise environments, e.g., road, wind, or traffic noises.
	Cognitive state sensitivity	Learning the driver cognitive states from speech for better recognition: what his/her emotional state is, whether he/she is busy with maneuvering the vehicle, drowsy, listening to radio or music, or thinking about something else.
	Addressee detection	Discriminate between system-directed from human-directed (or ambient) speech without using push-to-talk button. Tracking user conversation with the system for multiple dialog turns (also DM).
IU	Spoken language understanding	Make sense from broken human speech with high level of hesitation, revision, or wrong word order or high speech recognition error rate.
	Multimodal understanding	Integration of input from other natural modes (e.g., gesture on surface or in air, eye gaze) and accurate understanding of user intent.
DM		Decision about what system states need to be communicated to the drivers and at what conditions in coordination of the requests from the user in the dynamic changing context. Careful management of driver emotion, and selected content to the driver for the condition.
Knowledge management		Domain knowledge to be selected based on dynamic context and personal preference via user modeling and used to facilitate the IU as well as recommendation and decision making.
Natural language generation		Context-dependent content generation and prioritization. Discourse-aware flexible expression generation under different driving conditions and requested item availability as the template-based approach will no longer be sufficient.
Application manager	Fast multidomain integration	Multiple nomadic devices may be brought into the vehicles. To specify devices and services so to voice-enable them, integrate devices dynamically (plug and play) or perform service composition.
	Multiagent integration	Coordination, integration, and enrichment of the output from multiple assistant agents brought in primarily through various smart mobile devices, such as Siri, Google Now, Cortana, or embedded systems.
System level	Testing	Systematic testing of whole-system behavior with respect to system latency, response appropriateness, and task-completion rate.
	Specification/validation	Develop a detailed specification with many different possible combinations of scenarios, and validate the resulting system according to system specification based on standards, e.g., ISO 26262.

limited in-vehicle domains [27]. These gains will likely carry over to other modules in the future.

While these general IPAs benefit from leveraging huge computing resources in the cloud, they are highly dependent on Internet connectivity. However, many remote locations, where drivers need navigation systems the most, do not have Internet connectivity. This crucial limitation provides a strong justification for the role of an embedded global positioning system and associated embedded speech interface.

In summary, menu-structured multimodal interaction is the current state-of-the-art technology for operating in-vehicle infotainment products. Embedded voice interface solutions are less popular, due to limited recognition accuracy and lack of natural language capability. Cloud-based IPAs provide a good alternative for in-vehicle uses but with only a shallow integration. While many forward-looking building blocks in dialog systems—such as disfluency detection and content management—have been investigated in research projects, the conversion rate into in-vehicle embedded products has been slow. Among the

most well known in-vehicle dialog system products include those from Ford Sync (2007–2012). The GM CUE system took a major step in 2012 by incorporating natural speech beyond restricted grammars into its embedded dialog system, resulting in a noticeably improved user experience [28]. In the

meantime, cloud-based intelligent assistance systems have popularized in-vehicle speech use, due to their relatively high language-understanding accuracy and rich applications not available in embedded in-vehicle systems. However, the impact of cloud-based speech technologies is mostly limited to the ASR accuracy and text-to-speech (TTS) naturalness for a dialog system development. Despite all the technology and product limitation, it is

clear that a general direction for the in-vehicle interaction is conversational and intelligent dialog (CID) systems.

Challenges

In many ways, the technical challenges for in-vehicle intelligent dialog systems bear similarity to those intended for general-purpose dialog systems, especially in the context of hands-busy and

While many forward-looking building blocks in dialog systems have been investigated in research projects, the conversion rate into in-vehicle embedded products has been slow.

eyes-busy scenarios. Because of the special nature of in-vehicle use, however, additional challenges are apparent from in-vehicle data collection efforts and in-vehicle dialog system development activities mentioned in the section “Review of Past Major Activities.” Some important challenges are summarized in Table 1.

One may look at these challenges as arising from the three interdependent factors: 1) the driver, 2) the environment, and 3) the automotive industry.

- *The driver.* Drivers typically have short attention spans during interaction with in-vehicle dialog systems, as driving is their primary task. From 2020 to 2030, drivers will begin to be exposed to autonomous driving technologies, freeing them from constant attention to vehicle control. However, according to projections by McKinsey [39], the adoption of self-driving vehicles will likely be less than 15% by 2030. HMI while driving will continue to be important challenge. As a result, CID systems have to handle challenges on many different levels [9]. Drivers’ speech may be fragmented, disfluent, and repetitive. Drivers may need to hear shorter or simpler responses; they may expect the system to provide straightforward recommendations instead of a potentially overwhelming number of choices. When systems make recommendations, the systems should anticipate driver’s contextual needs to the point that the driver has the feeling that she or he doesn’t need to state the obvious, so driver behaviors and personal preferences need to be taken into consideration to avoid excessive conversational turns. Additionally, in one vehicle, there may be multiple passengers, each requiring a separate preference profile. A sophisticated dialog system should be able to coordinate requests from different users, know when to not interrupt human–human conversations, and come up with an optimized response for the group.

- *The environment.* The in-car environment is generally more dynamic and has higher stakes than the contexts in which other dialog systems are deployed. Inside the vehicle, information from sensors reflects vehicle status changes, some of which need immediate attention (an engine breaking down), while others can be handled at a later time (an oil change warning). Differences in design, interior materials, and mechanics create different acoustic environments and background noises inside vehicles. Outside the vehicle, the physical environment is also diverse and dynamic. A vehicle may be on a highway or a city road, accelerating or decelerating, on gravel or asphalt—creating significantly different background noises. Traffic conditions can vary significantly: the driver may be stuck in stop-and-go traffic, or may be traveling unimpeded at high speed. Harsh weather, such as wind, rain, hail, or thunder, typically requires additional attention from the driver and alters the in-vehicle acoustics. In such cases, the dialog system may need to use different dialog strategies with respect to taking the initiative

to engage the driver. Available services (e.g., gas station, or parking) via TTS from users’ mobile devices add another dimension of complexity and are competing for drivers’ attention. Managing multiple assistance systems from different devices will become an important development consideration.

- *The automotive industry.* Two key issues are critical to automotive companies’ decisions about whether or not to adopt new technologies. One is safety. Improper implementation of certain technologies could lead to fatal accidents. The other is the reliability. When a component is installed in a vehicle, it must work properly for a long time. Vehicle parts and features installed are typically required to function properly for at least ten years. For the safety and reliability reasons, this industry is highly regulated by the government. Relevant regulations and guidelines include ISO 26262 [29], ISO 15504 (Automotive SPICE), and IEC

61508 [30]. Due to such requirements, common practice in the consumer electronics or Internet world of fast product development cycle may not be directly applied in this industry. In the automotive industry, a rigorous process covering design specification, development, testing, and validation is followed to ensure the resulting product quality meets requirements. For example, to develop a CID system, one needs to specify the system coverage and performance by listing many phrasal and interactive variations under different noise conditions by speakers from different dialect regions, leading to a huge number of combined testing cases. A complete testing of these cases against various requirements becomes more and more challenging, especially with increasing coverage needs from the users.

The combination of these three factors causes many technical challenges. We next highlight some critical ones that have long-term impact in the areas of speech recognition and understanding, multiple speaker conversation coordination, the effect of driver behaviors and states for safety, as well as the integration of general intelligent assistance systems.

Challenges in speech recognition and understanding

From a historical perspective, speech recognition in the car started with small vocabulary systems primarily for command and control, along with optimization of either microphone placement or multimicrophone array processing to suppress the diverse noise sources present for in-vehicle scenarios [31]. More recent efforts have focused on expanding speech recognition coverage to additional in-vehicle domains.

The in-vehicle acoustic environment is complex and dynamic. Factors in this acoustic environment include noise from air conditioning units, wiper blades, the engine, external traffic, the road surface, wind, open windows, and inclement weather. The level of background noise while windows were

The automotive industry has taken two different approaches to address the arrival of cloud-based, voice-enabled assistance systems from major IT companies.

open 1–2 inches traveling at 65 miles/hour can be very close to speech level in the frequency band 0–1 kHz. Noise is not only intrusive and reduces the drivers' concentration but also degrades the ASR performance and interrupts the dialog flow. While some noise conditions are quite similar across vehicle types, some in-vehicle noise conditions vary significantly across vehicles such as noise produced by engines, turn signals, or wiper blades. Within the same vehicle, noise levels can vary from very low when the engine is idle and windows are closed, to very high when traveling at high speeds with open windows. Similar to other dialog systems, environmental noise also impacts how speakers speak, affecting a wide range of speech characteristics (including task stress, emotion, and Lombard effects).

Another challenge for in-vehicle speech recognition and understanding comes from imperfect speech input. When drivers are under stress, their speech can be less fluent and predictable. Their speech tends to contain more word fragments, restarts and repairs, hesitations, and alternative phrasings. These deviations from standard speech result in degraded speech recognition performance as a result of variation in acoustics and language.

For acoustic modeling, the classical cross-word modeling approach becomes less effective due to word fragmentation and hesitation. Similarly, speech endpoint detection becomes very difficult as the engine does not know whether a silence is a long pause or the end of a request, especially in the presence of noise. For language modeling, word fragments pose a challenge since there are often many proper names to cover and names can be interrupted leaving only a fragment. The more word fragments are included in the system, the more confusability is added, and it is harder to build a low-perplexity language model to constrain search space.

Initial attempts have been made to improve the detection of disfluent speech [32]. As DNN-HMM and Connectionist Temporal Classification (CTC) are overtaking GMM-HMM [18], [33], DNN-based technologies are predicted to better handle disfluent speech although their potential benefits need to be validated through real-world in-vehicle uses.

Challenges in coordinating conversations

Beyond the challenges of word recognition itself, interacting with increasingly capable voice technologies in the car will require more sophisticated coordination of human-machine conversations. Drivers freed from lower-level driving tasks will have more opportunity for social interactions both inside and (via mobile) outside the car. And the systems in the car that they do communicate with will have advanced knowledge and complexity. As a result, conversation management presents challenges and new opportunities even if one assumes that high-quality word recognition is available.

Another challenge for in-vehicle speech recognition and understanding comes from imperfect speech input.

One basic challenge is determining the addressee of an utterance, since a driver (or passenger) may be speaking to the system, or to another person, or to a person outside the car (e.g., on a call), or even to an automatic assistant on, e.g., his or her mobile phone. A system needs to know when it is being addressed, and respond only then and not in other contexts. In addition, the interruption of conversation will generally need to put another “on hold”—something that people are used to doing with other people, but generally not with systems. Addressee detection will need to scale to be able to suspend as well as resume conversations. It is worth noting that the current practice of “hot” words (wake-up words that are used to engage with a system) serves to initiate an addressee, but not to effectively suspend, resume, or close these interactions.

Even given the correct addressee, another challenge is how to handle interruptions to fluency of incoming speech for the timely determination of system responses. As just noted, speech contains pausing and disfluencies [34], and the task of driving only enhances opportunities for distraction and coping with sudden changes in the car or surrounding environment. Such fluency breaks cause important ambiguities for turn-taking. For example, even a simple pause to a navigation system, as in “which road do I turn onto (long pause) after I cross the bridge” produces different results pre- (incorrect) versus post- (correct) pause that matter to timely interaction. Waiting induces system latency if the speaker was actually done. Work using acoustic-prosodic features of prepause speech [35], as well as incremental content [36], can be used to better determine whether a user is suspending versus finishing an utterance. Additional challenges exist for handling self-repairs [37] in the driver's speech in real time.

Future challenges in conversational management also include the system production of “conversational grounding,” which becomes more necessary as utterance length and complexity increases. In natural conversation, partners employ speech back channels such as “uh-huh,” as well as visual cues such as gaze and head nods, to convey to each other that they are “still listening.” Mobile interfaces currently display visual information to ground users, but as autonomy increases, audio rather than visual grounding offers the benefit of an eyes-free option. Research on system-produced back channeling [38] offers promise for the future of natural interactive systems, but scaling to grounding in the safety and multiconversation environment requires a better understanding of how users interact with system-produced grounding mechanisms in real time and under cognitive load.

For all of these conversational management tasks, the in-vehicle environment offers unusually rich opportunities for speaker-dependent modeling. With fewer lower-level driving tasks to attend to, drivers are expected in particular to vary dramatically in use of voice for reasons unrelated

to driving, including phone calls. Systems can learn driver behaviors over time to achieve better performance and safety outcomes.

Challenges in incorporating driver behavior and driver states within a vehicle assistance system for enhancing system and safe driving

A further set of challenges for in-vehicle dialog systems concerns driver behavior and cognitive state in the context of additionally available input from in-vehicle sensors, Internet content, and Internet services, and their impact on dialog system design and development.

Driving is a highly dynamic process in which the level of environmental demands can change rapidly, while human cognitive resources are limited. Therefore, the most important goal in designing any in-vehicle driver-vehicle interaction system is to optimize information processing of drivers while operating such systems. Driver behavior is a crucial factor in traffic safety and interaction with in-vehicle intelligent systems. Intelligent dialog systems that minimize driver distraction require a thorough analysis and a good understanding of driver behavior. In addition, dialog systems should operate together with vehicle Advanced Driver Assistant Systems such as Driver Inattention Monitoring Systems and Driver Alert Control Systems [36]. In future autonomous vehicles, drivers may engage in many tasks other than driving. Accurate identification of driver state and behavior plays an important role in deciding hand over of vehicle control.

Two basic challenges are how to accurately observe driver behavior using different kinds of sensors and then, based on the observations, how to objectively identify driver behavior and cognitive states, which are highly subjective. Previous studies have used nonintrusive techniques combining eye movements, gaze variables, head orientation, heart rate, CAN-bus signal, vehicle position, and road geometry to capture driver-behavior signals. These approaches could be enhanced by integrating personal characteristics such as gender, age, and medical conditions. Earlier attempts monitored the behavior of driver and vehicle separately, whereas recent attempts focused on monitoring driver, vehicle, and driving environment simultaneously to effectively associate driver's behavior and states corresponding to contextual information. To infer the driver behavior from sensor data, recent common approaches perform some kind of probabilistic techniques to capture both static and dynamic behavioral characteristics. The challenge is how to define to normal versus abnormal driving status. Nevertheless, drivers possess an ability to manage their workload capacities and adjust their behavior to the environment under hazardous situations—but not in every situation. In addition to safety, monitoring driver behavior is important for a dialog system to dynamically adapt itself to the driver state (e.g., emotion)

to avoid negative experience of users from system errors, as well as nurture the positive experience.

Challenges to integrate general intelligent assistance technologies

As the demand for general cloud-based IPAs continues to increase, drivers will bring them into vehicles while driving. The general cloud-based IPAs have been designed primarily for uses other than in-vehicle; for example, responses are mostly displayed on the screen. Significant adaptation will be necessary to accommodate in-vehicle uses. Another challenge for the in-vehicle use of general IPAs would be to address different types of noises effectively to achieve high understanding accuracy with drivers and passengers. Yet another one will be about how well the general IPAs will be integrated into embedded spoken dialog systems with overlapping functionalities and different personalities. For such cases, one needs to decide which IPA will take the task and provide a solution. If multiple IPAs are used, the integration of multiple solutions for a task is

rather challenging and has to be resolved to offer consistent user experience. This would be especially important if different IPAs keep different user profiles with different learned preferences.

Taking into the consideration of all these factors and their possible combinations in a dialog system realization, we summarize some major technical challenges in Table 1.

Future trends and CID system outlook

Looking toward the future, voice-enabled in-vehicle assistance technologies will be influenced by two major trends on the horizon: increased automation in driving, with independent sensing and artificial intelligence capabilities; and increased vehicle connectivity to online IPAs, with driving-related services enhanced by traffic infrastructure and sensor advancement.

CID systems in the context of autonomous driving

Because the average American drives alone nine times as often as he or she drives or rides in a car with someone else, the primary emphasis regarding in-car speech systems has been and today still is on such systems' interactions with drivers and mitigating driver distraction. As technological advances push automotive design toward increasing autonomy in coming years, perhaps eventually culminating in fully self-driving cars, drivers will become increasingly like passengers. The nature as well as the risks of distraction vary along this spectrum of automotive autonomy. Even as driving becomes increasingly automated, the modality of speech retains essential advantages over other modalities for several reasons: listening does not detract from the visual attention needed for driving, language has richest expressive capability, speech technology is very flexible

Driver behavior is a crucial factor in traffic safety and interaction with in-vehicle intelligent systems.

with respect to form factor, and the speech channel is little utilized in single-occupant vehicles.

At all points on the vehicle autonomy spectrum, drivers and passengers can avail themselves of Internet connectivity to rich content and services only recently accessible in-vehicle. Time in the car can be used to entertain oneself or complete useful tasks, to the extent that these activities do not interfere with driving responsibilities. Audio channel is preferable to visual for these activities because reading in cars causes motion sickness for many people and because vehicle vibration caused by road conditions interferes with reading but listening is much less affected. Therefore, CID systems will remain a desirable interface channel for content and service consumption.

When drivers have greater responsibility for controlling their vehicles, CID systems hold the potential to reduce distraction and increase safety by, for example, communicating information about vehicle health and road conditions in natural language instead of with cryptic warning lights or not at all. When autonomous cars take over greater responsibility for controlling vehicles, the drivers' trust or sense of a safe ride need to be built up over time. This trust-building process for the adoption of autonomous vehicles can be facilitated by the CID systems through communicating information about vehicle controlling capability together with vehicle health and road conditions.

Further along the autonomy spectrum, where smarter cars and intelligent traffic systems provide ever greater driver assistance, the temptation for drivers to pay less attention to driving and more to unrelated content and services will naturally grow as the primary driving tasks diminish. The potential for boredom increases drivers' risk of becoming inattentive and occupying themselves with nondriving activities. Smarter cars equipped with smarter speech systems can reduce this risk of distraction by keeping drivers engaged to an appropriate extent, and preparing them to take over greater vehicle control as necessary. For instance, such cars could explain aloud their automated responses to significant changes in driving circumstances. When approaching a complex intersection, encountering difficult vehicle or pedestrian traffic, or upon detecting an unexpected road closure, cars can announce pertinent information preparing the driver to handle the situations.

When fully autonomous driving becomes available, transitions from autonomous-driving mode to human-driver mode and vice versa will need to be very intuitive and natural without any additional training. When the vehicle requires a driver to take over the control, the most natural way for drivers will be for the alert to come via speech request. Likewise, allowing drivers to request the vehicle to take over control using speech commands will be natural and convenient for drivers.

The increasing automation of driving may allow cars to contain larger screens. Virtual reality or mixed reality could make use of this screen space to present content, possibly as a three-dimensional (3-D) virtual world. CID systems may play a special role in navigating such 3-D worlds while driving; for example, one may, with a speech request, switch the display to a place not shown on a screen and preview its surroundings.

Recognizing that CID technology can both increase driving safety and improve user experience allows one to see new useful opportunities for in-car speech systems. These, however, require more conversational intelligence than is currently available to ensure that drivers' and passengers' voice interactions with systems are natural and not cognitively demanding or distracting.

CID systems in the context of the Internet of Things

Today, a single car can contain more than 100 sensors, supporting various vehicle functionalities and detecting occupant states. As vehicle technology advances, increasing numbers and types of sensors will appear in cars. Some of these will improve vehicle automation and safety, some will sense the external environment, for example, measure air quality, and others will sense occupants' physical states and actions, such as posture or alertness. Such sensors can provide much of the information needed for improving the conversational intelligence of CID systems.

With the increase of connectivity in the Internet of Things (IoT), devices and sensors will increasingly become more diverse with much richer information-exchange functionalities, expressing additional device status, features, operating instructions, or maintenance needs, etc. With new form factors of devices and sensors, the success of touch screens on smart phones for information exchange may not be easily repeated in cars where physical spaces are limited or design flexibility is required. As the screen size decreases, speech will become a more preferred information exchange modality.

The IoTs will also bring in much more content and many more services for drivers to access. One may receive a scenic spot description or hotel vacancy advertisement along a highway not through traditional billboards but rather from the Internet in real time. The general IPAs may intend to support people for these needs. However, given the typically fast-changing environment involved in driving, the in-vehicle use cases add much more dynamic and context-sensitive requirements for such assistance. In-vehicle CID systems have the potential in using in-vehicle sensors, such as gas-tank level, to find better solutions for the drivers. One will expect a tight integration of the general IPAs and embedded CID systems to offer the drivers synergized benefits from both systems.

At all points on the vehicle autonomy spectrum, drivers and passengers can avail themselves of Internet connectivity to rich content and services only recently accessible in-vehicle.

Expectation for better in-vehicle CID systems

In smarter cars with high autonomy and connectivity, speech systems need more conversational intelligence, as we have seen. Fortunately, the popular demand for better IPAs and mobile voice technology is driving improvements in conversational agents by increasing conversational intelligence in many of the ways previously mentioned in this article to be necessary for in-vehicle CIDs. Integrating developments in safety-optimized autonomy and conversational intelligence offers much promise for the CIDs that can meet automotive requirements for naturalness, ease of use, low cognitive demand on users, and minimal distraction from interacting with the CID itself. Of course, automobiles constitute a unique environment for speech in a multimodal setting, and additional research is needed specific to the automotive milieu, as discussed previously.

Moving forward, we foresee that CID systems would increasingly offer explanations about the vehicle itself, including functional operations, vehicle status, maintenance requirements, or even recommended driving styles for extended uses or environmental impact. CID systems will further act as a mediator to synchronize the content and services from different IPAs and integrate them with in-vehicle information. They will collaboratively support drivers on various activities with expert

advisories, and communicate properly based on their cognitive and emotional state (Figure 3).

A successful deployment of sophisticated in-vehicle CID systems in the future would require a breakthrough in the system development process from specification, development, testing, and validation in the automotive industry to ensure high-quality but low-cost software. The complexity in introducing many additional sensors into the vehicle combined with much more content and many more services from the cyberworld should not be underestimated. It is quite possible that additional new layers will be introduced in the

infotainment architecture to simplify the development and testing process. New standards may need to be introduced to facilitate industry-wide collaboration and incentivize the adoption of new technologies with affordable cost. Additional features may be also introduced to support the privacy and security in both the physical or cyberworlds by using CID systems to recognize and track drivers' identity and set up proper access or operation restrictions while driving.

A key aspect with in-vehicle dialog systems is that the interaction is often carried out while a driver is operating a vehicle as a secondary task.

Conclusions

An in-vehicle dialog system is a complex system that involves broad interdisciplinary knowledge and technologies from automatic speech recognition, spoken language understanding,

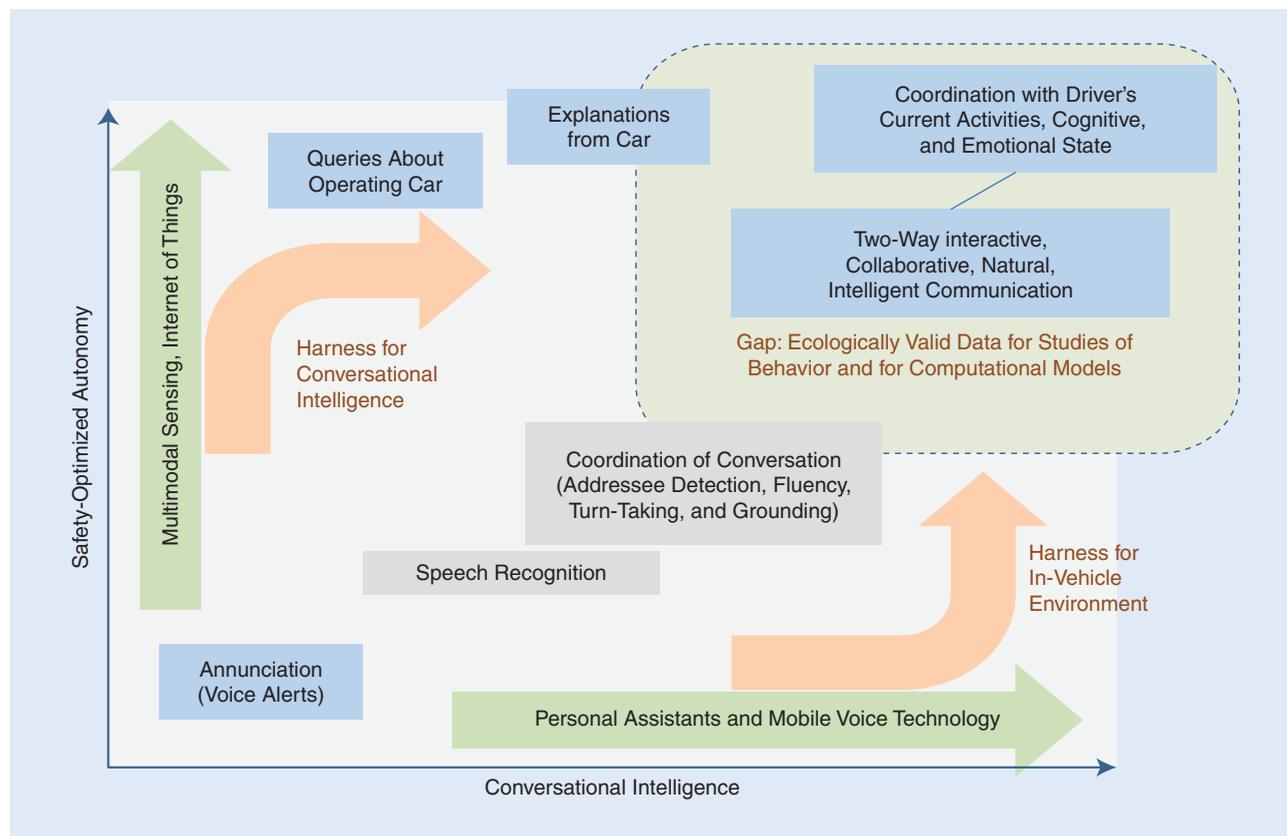


FIGURE 3. Future directions of CID systems.

DM, natural language generation, and TTS synthesis. In addition to challenges in development of other spoken dialog systems, a key aspect with in-vehicle dialog systems is that the interaction is often carried out while a driver is operating a vehicle as a secondary task. To ensure driving safety and intelligent interaction, it is necessary to provide an effective and user-friendly interaction between the driver and vehicle. The development of such a system requires the support of both the automotive and high-tech industries.

As sensors have become more reliable and accurate, the incorporation of multimodal interaction beyond audio itself allows the dialog system to detect additional nonverbal communications such as the intention and emotion of drivers. This information is advantageous for the dialog system to manage its interaction toward accomplishing its task based on the state or status of the driver, vehicle, and environment. Design of the intelligent in-vehicle dialog system also exploits such information for modeling behavior and usage patterns of a driver to adapt itself toward more effective interaction with an individual driver.

In the context of autonomous driving and the IoT, we expect to see more integration of speech-enabled technology with general IPAs fully connected with in-vehicle systems. We believe that in-vehicle dialog system technology will remain on demand with much more enriched features in the future for both the current human-centric driving paradigm and autonomous driving paradigm.

Acknowledgment

We would like to thank Dr. Liberty Lidz for her support on the preparation of the manuscript.

Authors

Fuliang Weng (Fuliang.weng@us.bosch.com) received his B.S. degree from Fudan University, Shanghai, China, in 1984. He was in a joint M.S. and Ph.D. program of Fudan University from 1984 to 1989, and completed all requirements but his Ph.D. thesis. In addition, Fuliang received an M.S. degree from New Mexico State University, Las Cruces, in 1993. He is the general manager of Bosch Language Services. Previously, he served as chief expert and director of user technologies at Bosch Corporate Research. He has more than 70 research publications and 60 issued or pending patents and has received multiple national and industrial awards.

Pongtep Angkittrakul (pongtep.angkittrakul@us.bosch.com) received his B.Eng. degree from Chulalongkorn University, Thailand, and his M.S. and Ph.D. degrees in electrical engineering from the University of Colorado at Boulder. Currently, he is a lead engineer at Robert Bosch LLC. Previously, he was a visiting researcher at Toyota Central R&D, Japan, where he contributed to algorithm development for advanced driver assistance system and driver monitoring technology, as well as the Toyota

To ensure driving safety and intelligent interaction, it is necessary to provide an effective and user-friendly interaction between the driver and vehicle.

award-winning project for drowsiness detector product. He organized and served as the general cochair of the Biennial Workshop for Digital Signal Processing for In-Vehicle Systems and Safety in 2015.

Elizabeth E. Shriberg (elizabeth.shriberg@sri.com) is a principal scientist at SRI International. She is also affiliated with Johns Hopkins University, the University of Maryland, and the Inter-

national Computer Science Institute, and she was recently at Microsoft Research. Her government and commercial work models spoken communication with focus areas in emotion, cognition, health, prosody, fluency, dialog, and speaker modeling, in which she holds multiple patents. She has authored more than 250 research publications and serves on the International Speech Communication Association (ISCA) Advisory Council and on journal, conference, and review boards for the speech science and speech technology communities. She is a fellow of SRI and ISCA.

Larry Heck (larry.heck@ieee.org) received his B.S.E.E. degree from Texas Tech in 1986 and his M.S.E.E. and Ph.D. degrees from the Georgia Institute of Technology in 1989 and 1991, respectively. He worked at SRI (1992–1998) in the Speech Technology and Research Laboratory. He was with Nuance (1998–2005) as vice president of research and development, where he led natural language processing, automatic speech recognition, voice authentication, and text-to-speech synthesis. At Yahoo! (2005–2009), he was the vice president of search and advertising sciences. While at Microsoft (2009–2014), he cofounded the Cortana personal assistant as chief scientist of speech. He is currently a research director for dialog at Google. He is a Fellow of the IEEE and holds more than 50 U.S. patents.

Stanley Peters (peters@csl.stanford.edu) is emeritus professor of Linguistics and Symbolic Systems at Stanford University, where he chaired the Department of Linguistics (1996–2002) and directed the Center for the Study of Language and Information (1987–1990 and 2008–2013). He previously taught at the University of Texas at Austin (1966–1983). His research interests are mathematical properties of grammar, logical semantics of natural languages, artificially intelligent tutoring systems, human–robot interaction, computational modeling of conversation, and spoken dialog systems. He has authored three books and more than 120 papers and holds five patents. He is a fellow of the Linguistic Society of America.

John H.L. Hansen (john.hansen@utdallas.edu) received his B.S.E.E. degree from Rutgers University, New Jersey, in 1982 and his M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology in 1983 and 1988, respectively. He serves as associate dean for research at the University of Texas–Dallas, where he founded the Center for Robust Speech Systems. He has been an associate editor for numerous IEEE publications

and organized the International Speech Communication Association (ISCA) INTERSPEECH 2002 and coorganized and was the technical program chair for the IEEE International Conference on Acoustics, Speech, and Signal Processing 2010 and the IEEE Spoken Language Technology Workshop 2014. He is a Fellow of the IEEE and ISCA.

References

- [1] H. van den Heuvel, J. Boudy, R. Comeyne, S. Euler, A. Moreno, and G. Richard, "The SPEECHDAT-CAR multilingual speech databases for in-car applications: Some first validation results," in *Proc. European Conf. Speech Communication and Technology*, Budapest, Hungary, 1999, pp. 2279–2282.
- [2] J. H. L. Hansen, P. Angkittrakul, J. Plucienkowski, S. Gallant, U. Yapanel, B. Pellom, W. Ward, and R. Cole, "CU-move: Analysis and corpus development for interactive in-vehicle speech systems," in *Proc. ISCA Interspeech Conf.*, Aalborg, Denmark, Sep. 2001, pp. 2023–2026.
- [3] N. Kawaguchi, S. Matsubara, I. Kishida, Y. Irie, H. Murao, Y. Yamaguchi, K. Takeda, and F. Itakura, "Construction and analysis of a multi-layered in-car spoken dialogue corpus," in *Proc. ISCA Interspeech Conf.*, Aalborg, Denmark, Sep. 2001, pp. 2027–2030.
- [4] E. Levin, S. Narayanan, R. Pieraccini, K. Biatov, E. Bocchieri, G. Di Fabrizio, W. Eckert, S. Lee, A. Pokrovsky, M. Rahim, P. Ruscitti, and M. Walker, "The AT&T-DARPA communicator mixed-initiative spoken dialogue system," in *Proc. Int. Conf. Spoken Language Processing*, 2000, pp. 122–125.
- [5] B. Lee, M. Hasegawa-Johnson, C. Goudeseune, S. Kamdar, S. Borys, and T. Huang, "AVICAR: Audio-visual speech corpus in a car environment," in *Proc. ISCA Interspeech Conf.*, Korea, 2004, pp. 2489–2492.
- [6] K. Takeda, J. H. L. Hansen, P. Boyraz, L. Malta, C. Miyajima, and H. Abut, "International large-scale vehicle corpora for research on driver behavior on the road," *IEEE Trans. Intell. Transport. Syst.*, vol. 12, no. 4, pp. 1609–1623, Dec. 2011.
- [7] P. Geutner, F. Steffens, and D. Manstetten, "Design of the VICO spoken dialogue system: Evaluation of user expectations by Wizard-of-Oz experiments," in *Proc. 3rd Int. Conf. Language Resources and Evaluation*, Canary Islands, 2002, pp. 1588–1593.
- [8] O. Lemon, K. Georgila, J. Henderson, M. Gabsdil, I. Meza-Ruiz, and S. Young, "D4.1: Integration of learning and adaptivity with the ISU approach," TALK Project No. IST-507802, Information Society Technologies, Germany, 2005.
- [9] F. Weng, S. Varges, B. Raghunathan, F. Ratiu, H. Pon-Barry, B. Lathrop, Q. Zhang, H. Bratt, T. Scheideck, K. Xu, M. Purver, R. Mishra, A. Lien, M. Raya, S. Peters, Y. Meng, J. Russell, L. Cavedon, E. Shriberg, and H. Schmidt, "CHAT: A conversational helper for automotive tasks," in *Proc. ISCA Interspeech Conf.*, 2006, pp. 1061–1064.
- [10] B. Lathrop, H. Cheng, F. Weng, R. Mishra, J. Chen, H. Bratt, L. Cavedon, C. Bergmann, T. H. Bender, H. P. Barry, B. Bei, M. Raya, and E. Shriberg, "A Wizard of Oz framework for collecting spoken human-computer dialogs: An experiment procedure for the design and testing of natural language in-vehicle technology systems," in *Proc. 12th World Congress Intelligent Transport Systems*, San Francisco, CA, 2005.
- [11] B. Thomson and S. Young, "Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems," *Comp. Speech Lang.* vol. 24, no. 4, pp. 562–588, Oct. 2010.
- [12] D. Hurwitz, E. Miller, M. Jannat, L. Boyle, S. Brown, A. Abdel-Rahim, and H. Wang, "Improving teenage driver perceptions regarding the impact of distracted driving in the Pacific Northwest," *J. Transport. Safety Security*, vol. 8, no. 2, pp. 148–163, 2016.
- [13] C. Muller, G. Weinberg, and A. Vetro, "Multimodal input in the car, today and tomorrow," *IEEE Multimedia*, vol. 18, no. 1, pp. 98–103, 2011.
- [14] Y. C. Cheng, K. Li, Z. Feng, F. Weng, and C.-H. Lee, "Online whole-word and stroke-based modeling for hand-written letter recognition in in-car environments," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2013, pp. 847–852.
- [15] G. Tur, A. Stolcke, L. Voss, D. Hakkani-Tur, J. Dowding, B. Favre, R. Fernandez, M. Frampton, M. Frandsen, C. Frederickson, M. Graciarana, D. Kintzing, K. Leveque, S. Mason, J. Niekraz, M. Purver, K. Riedhammer, J. Tien, D. Vergyri, and F. Yang, "The CALO meeting assistant system," *IEEE Trans. Audio Speech Lang. Processing*, vol. 18, no. 6, pp. 1601–1611, Aug. 2010.
- [16] Deep learning (2016). [Online]. Available: https://en.m.wikipedia.org/wiki/Deep_learning
- [17] N. Morgan and H. Bourlard, "Continuous speech recognition: An introduction to the hybrid HMM/connectionist approach," *IEEE Signal Process. Mag.*, vol. 12, no. 3, pp. 25–42, May 1995.
- [18] G. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Process. Mag.*, vol. 29, pp. 82–97, Nov. 2012.
- [19] L. Heck, Y. Konig, M. Sonmez, and M. Weintraub, "Robustness to telephone handset distortion in speaker recognition by discriminative feature design," *Speech Commun.*, vol. 31, no. 2, 2000, pp. 181–192.
- [20] Y. Konig, L. Heck, M. Weintraub, and M. Sonmez, "Nonlinear discriminant feature extraction for robust text-independent speaker recognition," presented at the RLA2C, ESCA Workshop, Avignon, France, 1998.
- [21] M. L. Seltzer, D. Yu, and Y. Wang, "An investigation of deep neural networks for noise robust speech recognition," in *Int. Conf. Acoustics, Speech and Signal Processing*, Vancouver, Canada, 2013, pp. 7398–7402.
- [22] X. Feng, B. Richardson, S. Amman, and J. Glass, "On using heterogeneous data for vehicle-based speech recognition: A DNN-based approach," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing*, Brisbane, Australia, 2015a, pp. 4385–4389.
- [23] A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, A. Coates, and A. Y. Ng. (2014). Deep speech: Scaling up end-to-end speech recognition [Online]. Available: <https://arxiv.org/abs/1412.5567>
- [24] S. Nordholm, I. Claesson, and M. Dahl, "Adaptive microphone array employing calibration signals: Analytical evaluation," *IEEE Trans. Speech Audio Process.*, vol. 7, pp. 241–252, May 1999.
- [25] J. H. L. Hansen, J. Plucienkowski, S. Gallant, B. Pellom, and W. Ward "CU-Move: Robust speech processing for in-vehicle speech systems," in *Proc. Int. Conf. Spoken Language Processing*, 2000, vol. 1, pp. 524–527.
- [26] X. Zhang and J. H. L. Hansen, "CSA-BF: A constrained switched adaptive beamformer for speech enhancement and recognition in real car environments," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 733–745, Nov. 2003.
- [27] Z. Zhou, R. Botros, H. Lin, Y. Hao, Z. Feng, and F. Weng, "Hybrid speech recognition using pairwise classification and deep neural network based frameworks with long-distance constraints," Bosch Internal Report No. CR/RTC-NA-847, Robert Bosch LLC, Palo Alto, CA, 2016.
- [28] UKIP Media & Events Ltd [Online]. Available: http://www.automotive-interiors-expo.com/english/awards_14_winners.php
- [29] ISO [Online]. Available: <https://www.iso.org/obp/ui/#iso:std:iso:26262:-6:ed-1:vi:en>
- [30] IEC [Online]. Available: <http://www.iec.ch/functionalsafety/>
- [31] F. Soong and E. Woudenberg, "Hands-free human-machine dialogue-corpora, technology and evaluation," in *Proc. Int. Conf. Spoken Language Processing*, 2000, vol. 4, pp. 41–44.
- [32] Q. Zhang, F. Weng, and Z. Feng, "A progressive feature selection algorithm for ultra large feature spaces," in *Proc. 21st Int. Conf. Computational Linguistics*, 2006, pp. 561–568.
- [33] X. Maas, D. Jurafsky, Z. Xie, and A. Y. Ng, "Lexicon-free conversational speech recognition with neural networks," presented at the North American Association for Computational Linguistics, Denver, CO, 2015.
- [34] H. H. Clark, "Managing problems in speaking," *Speech Commun.*, vol. 15, no. 3–4, pp. 243–250, Dec. 1994.
- [35] H. Arsikere, E. Shriberg, and U. Ozertem, "Computationally-efficient end-pointing features for natural spoken interaction with personal-assistant systems," in *Proc. Int. Conf. Acoustics Speech Signal Processing*, 2014, pp. 3241–3245.
- [36] L. Ferrer, E. Shriberg, and A. Stolcke, "Is the speaker done yet? Faster and more accurate end-of-utterance detection using prosody in human-computer dialog," in *Proc. Int. Conf. Spoken Language Processing*, 2002.
- [37] E. E. Shriberg, "To 'errrr' is human: Ecology and acoustics of speech disfluencies," *J. Int. Phonetic Assoc.*, vol. 31, no. 1, pp. 153–169, 2001.
- [38] N. Ward and W. Tsukahara, "Prosodic features which cue back-channel responses in English and Japanese," *J. Pragmatics*, vol. 32, no. 8, pp. 1177–1207, 2000.
- [39] P. Gao, H.-W. Kass, D. Mohr, and D. Wee, "Automotive revolution: Perspective towards 2030—Advanced industries," McKinsey & Company, New York, Jan. 2016.

Prasanga N. Samarasinghe,
Wen Zhang, and Thushara D. Abhayapala

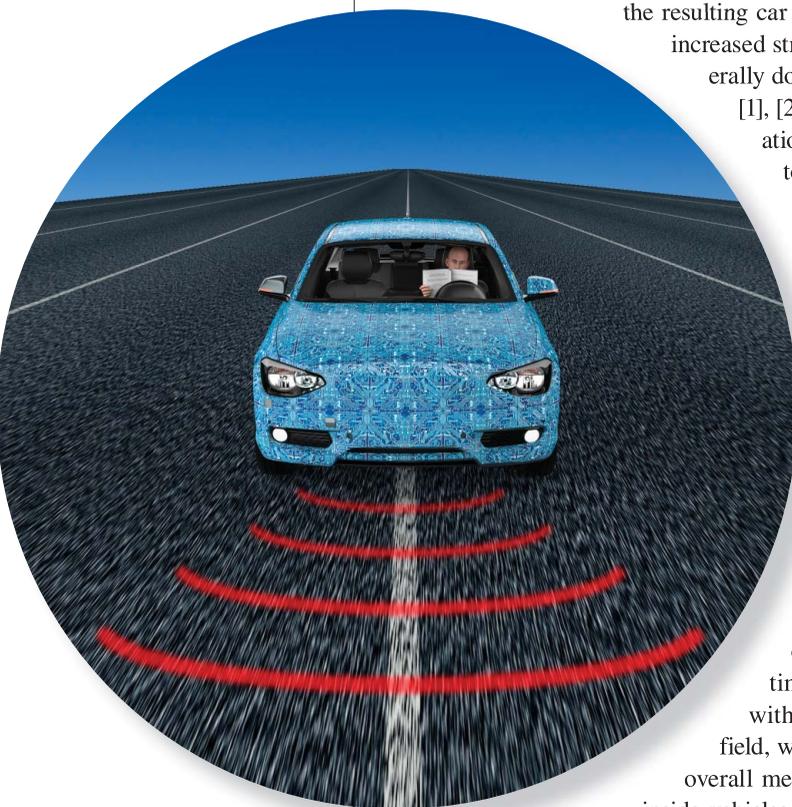
Recent Advances in Active Noise Control Inside Automobile Cabins

Toward quieter cars

Minimization of interior cabin noise has been a key topic of research in the automobile industry for the last two decades. This problem was initially approached via passive noise cancellation methods, where physical treatments such as structural damping and acoustic absorption were used. However, with vehicle manufacturers striving for more economical and lightweight designs, the resulting car interiors invariably became noisier due to the increased structural vibrations. These noise fields are generally dominated by low frequencies (i.e., 0–500 Hz) [1], [2], hence, the conventional passive noise cancellation approaches are less effective. In an attempt to resolve the aforementioned problem, active noise control (ANC) methods were developed where secondary sources were proposed to attenuate the noise inside the cabin. With modern in-car entertainment systems providing four to six built-in loudspeakers, the addition of an ANC system comes at a relatively low additional cost.

In practice, in-car ANC is achieved by producing a secondary signal(s) that cancels the noise generated by the noise source(s). The residual difference between these two components is measured using a microphone(s) placed inside the cabin and is minimized using a feedforward/feedback control system [3]. Feedforward systems use a time-advanced “reference signal(s)” correlated with the noise signal to attenuate the primary noise field, whereas feedback systems tend to attenuate the overall measured noise. The basic concept behind ANC inside vehicles is described in Figure 1, which shows how a secondary sound field cancels out the undesired noise field utilizing an adaptive controller. Since the noise observed inside vehicle cabins is often random and time varying, the aforementioned control systems are required to be adaptive. While the theory and concept behind ANC systems are quite straightforward, their practical implementation and performance are often hindered by factors such as the noise field complexity inside the car geometry, cost, adaptive system convergence time, system stability, noncausality and poor spatial coverage. Over the

SKY IMAGE LICENSED BY GRAPHIC STOCK
©ISTOCKPHOTO.COM/POSTERIOR



Digital Object Identifier 10.1109/MSP.2016.2601942
Date of publication: 4 November 2016

last 30 years researchers and car manufacturers have done extensive amounts of experimental research to overcome these limitations [1], [4]–[6].

In this article, we review ANC techniques for noise cancellation inside automobiles. We explain the different types of noise fields present inside vehicles, the theoretical basics of ANC techniques, and their applicability in canceling the aforementioned noise fields. We focus on recent advances made in vehicle ANC during the past 10–15 years including commercial developments available in mass production vehicles. We aim to show that in-car ANC is an exciting field of research with the potential to substantially improve the passenger experience in acoustically challenging environments.

Noise source and noise field inside vehicles

Different kinds of noise sources exist in automobiles, such as engine, road-tire, and wind noise, each with their own distinct acoustic properties. The vehicle compartment can be considered a very small room and depending on the frequency, the acoustics inside the car exhibit completely different physical behaviors. This section reviews the attributes of typical noise sources and noise models, which effectively describe the noise fields inside automobiles.

Attributes of noise sources

Engine noise

Most road vehicles with four or more wheels employ a reciprocating, four stroke, internal combustion engine [7]. The noise produced by the internal combustion engine is dominated by two processes, the piston crank mechanism and the combustion process. The piston crank mechanism, such as the movement of pistons and their lines, generate an impulsive noise with a flat spectrum; the combustion process, on the other hand, produces a tonal noise, which is directly related to the

Different kinds of noise sources exist in automobiles, such as engine, road-tire, and wind noise, each with their own distinct acoustic properties.

rotational speed of the engine. ANC strategies are generally more effective in controlling combustion noise processes due to its predictive nature. Based on in-car noise measurements, there exist two simple relationships between the engine type and the resulting combustion noise. Given the engine size expressed in terms of the number of cylinders and their individual capacity

in liters, the average noise level can be empirically estimated using the following equation [6]

$$\text{Noise Power Level} = 10 \log_{10}(\text{no. of cylinders}) + 23 \log_{10}(\text{cylinder capacity}).$$

With the knowledge of the engine rotation speed, the fundamental noise frequency/firing frequency/dominant engine order is

$$f_0 = \frac{\text{rotation speed}}{2 \times 60} \times \text{no. of cylinders}.$$

Derivation of the dominant engine order at any given rotation speed [revolutions/minute (r/minute)] is straightforward. First, the r/minute is converted to Hertz by a multiplication of 1/60 (e.g., an engine spinning at 1,800 r/minute can be said to be running at 30 Hz). Second, since a four-stroke engine fires each cylinder only once every two crank revolutions, the rotations per second is multiplied by half the number of cylinders (e.g., for a six-cylinder engine spinning at 1,800 r/minute, the dominant engine order is at 90 Hz), which gives the fundamental frequency. In a six-cylinder engine, it's also called the *third engine order* because the frequency is three times that of the engine's rotation in Hz. While the dominant engine order defines the engine's distinctive sound character, its overall timbre is decided by multiple variables such as its structure, plumbing, and materials, which cause additional engine orders to become active. Therefore, typical engine induced noise carries multiple engine orders.

In automatic transmission powertrains, the torque converter and torque converter clutch are critical devices governing the overall power transfer efficiency. They create a one-to-one connection between the output of the engine and the input of the transmission. With increasing demand for fuel economy, the recent trend is to apply the torque converter clutch over a wider range of driving conditions. This increases powertrain high torque fluctuation and causes noise and vibration. While there exist many passive control solutions for this issue, active control of noise and vibration is one of the most efficient solutions because active control avoids the addition of extra weight [8].

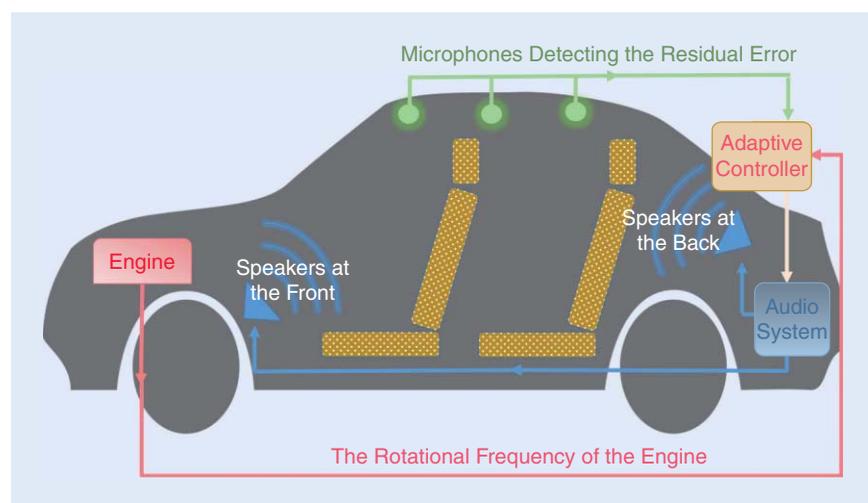


FIGURE 1. ANC inside an automobile cabin.

Road-tire noise

Road-tire noise is produced due to the interaction between the road surface and the tire, which can be classified mainly into two kinds, air pumping noise and vibration induced noise (or road booming noise) [7]. Air pumping noise is caused by the tire/tread pumping when on a rough road surface. Normally, this kind of noise is dominated by high frequencies and its level is dependent on factors such as the size of the road, the size of tire cavities, the load on the tire, and the pressure inside the tire. In contrast, vibration induced noise (or road booming noise), is mainly due to the nonuniformity of road surfaces, change of vehicle speed and irregularities in the tire tread pattern. There are several characteristics of vibration induced noise. First, it is generated by a combination of independent vibrations of the four wheels, therefore it is hardly reduced by ANC techniques with one or two reference sensors. Second, the spectrum and properties of vibration induced noise vary continuously with varying road profiles and vehicle speed. Third, transfer function between wheel vibrations and road booming noise is nonlinear. The above characteristics are often unique to the vehicle of interest and therefore, ANC of road noise is rather difficult to achieve in realistic conditions as opposed to laboratory setups.

Wind noise

The wind noise is the predominant component of interior noise at speeds above 100 km/hour [2]. It can be classified by the noise production mechanisms, such as 1) a low-frequency broadband noise due to the vehicle moving through air at mid- to high speeds or turbulent air flow through holes (e.g., vehicle windows and doors), 2) an impulsive noise above 300 Hz due to the external varying wind conditions, and 3) a narrowband beating noise due to air flow over open windows or sunroofs. The acoustic energy inside vehicle cabins due to wind noise is generally concentrated at the frequency band 50–500 Hz.

Noise fields inside vehicles

The vehicle compartment is generally considered as a small room and has distinct acoustic properties, similar to room modes. Room modes or acoustic modes are a collection of resonances that exist in a room when the room is excited by an acoustic source. Most rooms have their fundamental resonances in the 20–200 Hz region, each frequency being related to one or more of the room's dimension's or a divisor thereof. At low frequencies, the sound field is dominated by a limited amount of acoustic modes. At high frequencies, as the number of acoustic modes increases, the sound field becomes increasingly diffuse, making it more accurate to be modeled using statistical methods. In this region, a typical phenomenon is the spectral coloration, where the spectral peaks and dips do not correspond to eigen frequencies of acoustic modes but

are the result of numerous overlapping resonances. Often the Schroeder cut-off frequency [9] is used to separate the low- and high-frequency regions. This is about 300 Hz in a typical car compartment [6].

In current automotive ANC systems that are commercially available, the main objective is to globally cancel the dominant engine order(s) inside the vehicle cabin such that all passenger seats are covered. Global cancelation of noise requires

In current automotive ANC systems that are commercially available, the main objective is to globally cancel the dominant engine order(s) inside the vehicle cabin such that all passenger seats are covered.

the entire noise field inside the car to be considered. This is best done by describing the noise field in terms of acoustic modes, which depend on the noise field's frequency content and the structural-acoustic coupling. (When a vibrating structure is in contact with air, some of the energy from the structure escape to the air as sound.) This interaction is referred to as *structural-acoustic coupling* of the enclosure. ANC control systems thus require the speakers positioned to control the aforementioned acoustic modes and the error microphones

positioned to observe the relevant acoustic modes. The effective frequency range of a given ANC system is determined by the modal density (the average number of modes in a unit frequency interval) of the automobile enclosure and the available number and placement of microphones and loudspeakers. This limitation is purely physical and does not depend on the control algorithm or software used for achieving ANC.

In the automotive industry, low-frequency noise fields inside vehicles are often simulated using computer-aided engineering (CAE) technology to assist production. CAE often uses the finite element method [10] and boundary element method [11] to model the vibrations and structural-acoustic coupling inside vehicle cabins. Since CAE is largely useful for ANC, and the current direction of commercial vehicle development leans toward shortening the development time while decreasing the prototype vehicle quantity, there is an urgent need to improve the computer-based prediction technologies at the planning stage [12].

ANC techniques

ANC involves a system that cancels the primary (unwanted) noise based on the principle of superposition. In the time domain, for a single-input, single-output (SISO) system,

$$e(n) = d(n) + y'(n), \quad (1)$$

where $e(n)$ represents the error signal and $d(n)$ and $y'(n)$ represent the noise and secondary sound fields present at the error sensor respectively. Typical ANC systems require one or more loudspeakers to produce the secondary sound field, one or more microphones to measure the residual error signal(s) present at the observation point(s) of interest, and an adaptive control system to drive the loudspeaker(s) while minimizing the residual error. In-car ANC systems can be broadly classified as, feedforward systems and

feedback systems depending on whether reference sensors are present, or broadband and narrowband depending on the operating bandwidth. The next two sections will review each system in detail, while providing their recent applications, mainly related to engine noise cancellation and road noise cancellation.

Feedforward control systems

Feedforward systems use an additional sensor(s) (acoustic/mechanical/optical/electric) to measure the primary noise field or to generate a signal related to the noise generation mechanism. This reference signal(s) is processed by the ANC system to drive the loudspeaker(s), to minimize the residual error. The performance of feedforward ANC systems is dependent on the coherence between the reference signal and the primary noise, thus there exists an inherent requirement for the reference sensors to be placed close to the noise source. The performance of a feedforward ANC system also relies on the frequency spectrum of the primary noise. If the primary noise field is a random broadband sound field, it is important that the reference signal and the adaptive system continuously track it. Furthermore, if the adaptive system's electrical delay (due to the processing time) is larger

than the acoustic delay from the reference microphone to the canceling loudspeaker, the controller response becomes noncausal and the system performance will be substantially degraded. However, in narrowband feedforward systems, the continuous tracking requirement and the causality condition are largely preserved because the reference signal is often predictable.

Feedforward control algorithms

Broadband feedforward ANC

Broadband feedforward control systems are utilized when the primary noise field has a broadband frequency response (e.g., road noise). The system identification framework of a basic SISO broadband feedforward control system is illustrated in Figure 2. The primary path system function $P(z)$ consists of the acoustic response from the reference sensor location to the error sensor, and the secondary path system function $S(z)$ consists of a combination of 1) the electronic response between the adaptive filter $W(z)$ and the loudspeaker, and 2) the acoustic response from the loudspeaker to the error microphone. If the secondary path transfer function is not taken in to account (i.e., $W(z) = P(z)$), the system will become unstable, because the error signal will not be correctly time-aligned with the reference signal. Since the primary path of a vehicle cabin is often dynamic, the role of the adaptive filter $W(z)$ is to continuously track the time variations in the primary noise source [through the reference signal $x(n)$] and minimize the residual error signal $e(n)$. The most common form of adaptive filters is the transversal filter using the least-mean-square (LMS) algorithm. From Figure 2, the Z-transform of the error signal is

$$E(z) = (P(z) - S(z)W(z))X(z). \quad (2)$$

In the ideal case, $E(z) = 0$ after the adaptive filter converges, which implies that the optimal filter response is $W(z) = P(z)/S(z)$. To achieve this result, the adaptive filter has to simultaneously model $P(z)$ and inversely model $S(z)$. Since an inverse does not always exist for $S(z)$, Morgan [13] suggested a more effective approach, where an identical filter [to $S(z)$] was proposed to be placed along the reference signal path to the weight update of the LMS algorithm. This modification compensates for the secondary path effects. It is also the origin of the well-known filtered-XLMS (FXLMS) algorithm, for which the corresponding block diagram is given in Figure 3. The term $\hat{S}(z)$ in Figure 3 refers to an estimated value of the secondary path. For a more detailed derivation of the FXLMS algorithm, see [3]. Another practical limitation that arises with feedforward systems is the effect of feedback. This involves the upstream of the antinoise output from the secondary loudspeakers to the reference sensor, which corrupts the reference signal. The simplest approach to solving this problem is to model the feedback path transfer function $F(z)$ and neutralize it with a separate feedback cancelation filter [3]. However, this leads to additional stability issues in

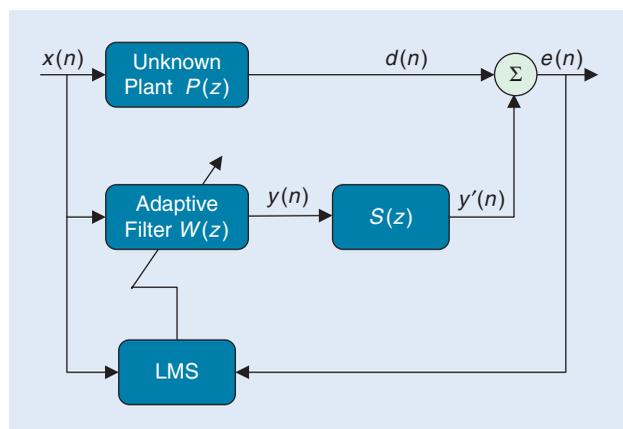


FIGURE 2. A block diagram of a feedforward ANC system.

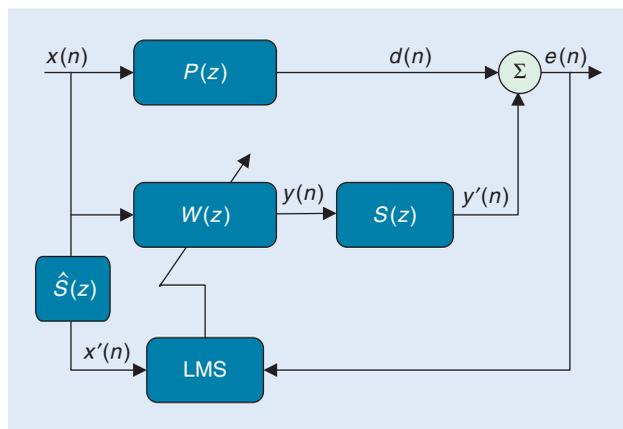


FIGURE 3. A block diagram of a feedforward ANC system using the FXLMS algorithm.

practice, and incorrect modeling of the feedback path can lead to instability.

Narrowband feedforward ANC

Noises generated by mechanical components such as engines, compressors, motors, and fans are generally narrowband and periodic. To monitor such noise, it is sufficient to use a non-acoustic sensor (e.g., tachometer), which provides an electrical reference signal that contains the fundamental frequency and all of the harmonics of the primary noise source. This technique has several advantages compared to a broadband ANC system using acoustic reference sensors; 1) the effect of feedback is eliminated, 2) aging and nonlinearities associated with acoustic reference sensors are avoided, and 3) the causality condition is preserved due to periodicity. In narrowband ANC systems, once an electric reference signal is available, a corresponding acoustic reference signal is internally generated to assist the noise cancellation process.

MIMO feedforward ANC

When the noise field of interest increases in size and bandwidth, the number of active acoustic modes increases. To control multiple acoustic modes, it is necessary to use multiple-channel ANC systems with multiple secondary sources, error signals, and reference signals. A MIMO feedforward ANC system employs J reference sensors to observe the primary noise, M error sensors to measure the residual noise, and K secondary sources to produce the antinoise. Figure 4 illustrates the block diagram of a broadband adaptive MIMO feedforward ANC system with feedback and secondary path transfer functions. The wide arrows represent a flow of vectors (multichannel acoustic or electrical signals). The matrix P represents $M \times J$ primary path transfer functions, matrix S represents $K \times M$ secondary path functions, matrix F represents $K \times J$ feedback path functions, and W represents a matrix of $K \times J$ adaptive filters each serving an individual feedforward channel.

Application of feedforward control to car noise cancellation

In the application of car-noise cancellation, feedforward ANC systems are mostly applied for engine noise cancellation. This is because it's easier to obtain a reference signal directly from the engine resulting in high coherence between the reference and error signals. Furthermore, as mentioned previously, engine noise is often periodic. Therefore, ANC for engine noise is often approached via a cost-effective narrowband feedforward system using an engine speed reference sensor, low-cost microphone error sensor(s), and the vehicle's built-in loudspeaker system as control sources. Such ANC systems have been commercially implemented by a number of manufacturers as discussed later in the section "Commercial Systems."

The control bandwidth of feedback control systems are inversely proportional to the spacing in between the error sensor(s) and secondary source(s).

Ideally, noise cancellation inside a vehicle would require the total acoustic energy distributed over the entire global region to be minimized. Since this is an impractical task, the control region is often sampled using one or more error sensors distributed over the control region. The total acoustic energy to be minimized is then approximated by the sum of squares of the sensor output as

$$J_p = \sum_{m=1}^M p_m^2 \tag{3}$$

where p_m is the sound pressure at the m th error sensor position ($m = 1 \dots M$). The effects due to the above approximation is often comparable for low frequency control but gets increasingly noticeable at high frequencies. The accuracy of the above approximation largely depends on the location of the error sensors because, as mentioned in the section "Noise Fields Inside Vehicles," noise field characteristics inside an enclosure are largely related to the enclosure's active number of acoustic modes and structural-acoustic coupling. The effects of

structural-acoustic coupling on ANC inside vehicles have been thoroughly studied over the last 20 years [1], [14], [15]. In [1], Elliot et al. showed that at low frequencies, a single sensor is capable of achieving significant control, however with increasing complexity of the sound field (frequency and geometry), multiple acoustic modes become active and, therefore, multiple sensors are required to successfully couple into all of them. The frequency limit of global control was shown to be directly related to the modal overlap or the number of acoustic modes that are significantly excited at a given frequency f , which increases with the cube of f . Therefore, to achieve control over the entire global region with a size of a car, the number of sensors required are often impractical.

To improve the control bandwidth, an alternative control strategy needed to be developed. Such a strategy was recently investigated in [6] and [16], which attempts to control the sound field within smaller spatial regions (regional

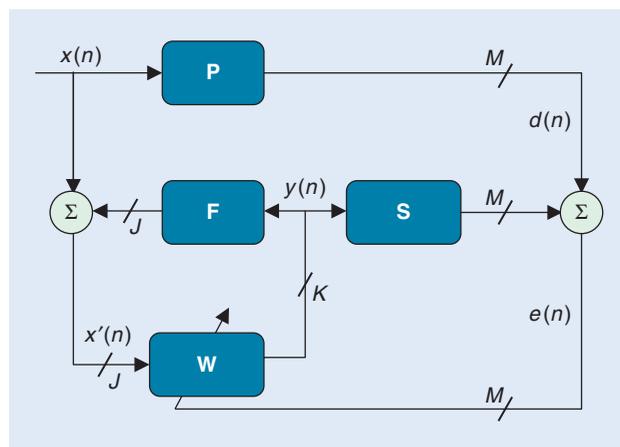


FIGURE 4. A broadband adaptive MIMO feedforward ANC system with feedback and secondary paths.

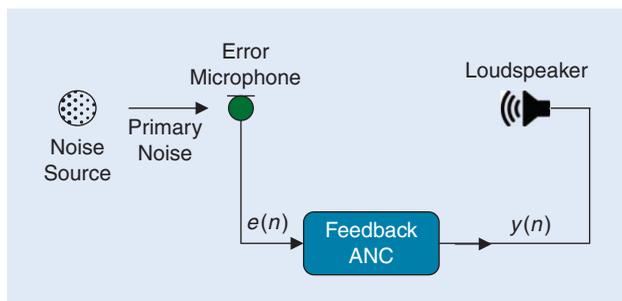


FIGURE 5. A SISO adaptive feedback ANC system.

control), particularly surrounding the driver’s/passenger’s head. Regional or local control of noise reduces the volume over which the noise energy has to be minimized and therefore reduces the constraints on the control system and increases the control bandwidth. In [6], the authors investigated a regional narrowband feedforward system over two regions, one rectangular region across the front seats and a second rectangular region across the rear seats. The performance of the system was investigated through simulations and synthesis based on transfer functions measured in a rectangular car cabin mock-up. The control system comprised four secondary sources positioned at the standard car audio loudspeaker positions, and eight error sensors positioned at the four head rest positions (two sensors on each headrest). The acoustic potential energy within the control regions were shown to be significantly reduced at frequencies up to 370 Hz. This is around twice the control bandwidth of global feedforward control using a similar system. The authors mention a potential issue with the regional feedforward control strategy, i.e., the system is said to be susceptible to unobservable modes that result in enhancements in the regional acoustic potential energy. However, it has also been shown that these effects can be limited by using control effort weighting parameters.

In addition to engine noise cancellation, feedforward systems are also applied in road noise cancellation. In [17], Oh et al. presented a leaky constraint MIMO feedforward ANC system for road booming noise control in a midsize passenger vehicle using two accelerometers, two control loudspeakers,

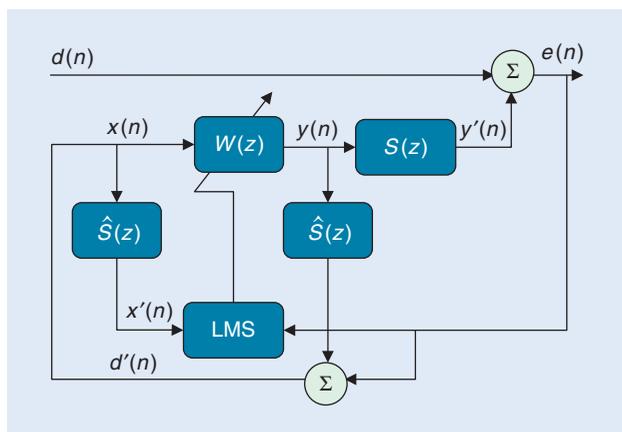


FIGURE 6. A SISO adaptive feedback ANC system using the IMC architecture.

and a single error microphone. Based on experimental results, the optimum positioning for the aforementioned devices were chosen, and during driving tests on rough asphalt and curvy roads at 60 km/hour, a reduction of 6-dB A-weighted sound pressure level (A-weighting accounts for the relative loudness perceived by the human ear) in the road booming noise was achieved. Due to the high cost of accelerometers, feedforward ANC is generally regarded as unsuitable for mass production, however, with the recent introduction of low-cost MEMS accelerometers, it is expected to change.

Feedback control systems

This section discusses the adaptive feedback control systems used for broadband ANC. Unlike feedforward systems, feedback systems directly employ the signal(s) from error sensor(s) to drive the secondary source(s) via a controller. Since the error sensor signal is fed back to the secondary source, the system cannot be optimized on a frequency-by-frequency basis as in feedforward control, and, therefore, the whole frequency response (broadband) must be considered at all times. The performance of feedback control systems are limited by their stability, which is largely dependent on the system delay. Therefore, the control bandwidth of feedback control systems are inversely proportional to the spacing in between the error sensor(s) and secondary source(s).

Feedback control algorithms

SISO feedback control

A single-channel adaptive feedback ANC system, as shown in Figure 5, was first proposed in [18]. Based on the internal model control (IMC) architecture, an adaptive feedback system can be viewed as an adaptive feedforward system, which synthesizes or regenerates its own reference signal using the error signal and the adaptive filter output. The basic concept of this model is to estimate the primary noise $d(n)$ present at the error sensor and use it as a reference signal $x(n)$ for the adaptive filter. If the secondary path transfer function $S(z)$ is known, the primary noise signal $d(n)$ can be synthesized using

$$X(z) \equiv \hat{D}(z) = E(z) + \hat{S}(z)Y(z), \quad (4)$$

where the notation “ $\hat{}$ ” represents an estimated value. Therefore, the reference signal synthesis technique filters the secondary source signal $y(n)$ using the secondary path estimate $\hat{S}(z)$ and combines it with $e(n)$ to regenerate the primary noise. Figure 6 shows a complete SISO feedback ANC system using the FXLMS algorithm with secondary path cancellation as discussed in the section “Broadband Feedforward ANC.” When applying a feedback control system, the overall control system stability is very important next to the noise reduction level. This is generally analyzed by checking the Nyquist stability criterion, which states that the polar plot of the open-loop response must not enclose the Nyquist point $(-1, 0)$ as ω increases from $-\infty$ to $+\infty$ [19]. In a practical system, since the open-loop response often varies with time, it is typical to

set the feedback gain to stabilize the system despite its variations. A system that stabilizes with such a gain is said to have robust stability.

MIMO feedback control

The SISO feedback ANC system is extendable to a multiple-channel system with K secondary sources and M error sensors. Such a system will have $M \times K$ secondary paths. Each path $S_{mk}(z)$ is from the k th secondary source to the m th error sensor and needs to be estimated by a filter $\hat{S}_{mk}(z)$. These estimated filters along with the K secondary/control signals $y_k(n)$ and the M error signals $e_m(k)$ will synthesize M reference signals $x_m(n)$ for the corresponding $K \times M$ adaptive filters $W_{km}(z)$. A multiple-channel FXLMS algorithm will be required to calculate the coefficients of the adaptive filters. Figure 7 illustrates a block diagram of the entire process described previously. In practice, the extension of a SISO feedback system to a MIMO feedback system is somewhat complex due to the need to calculate the eigenvalues of the open-loop response to assess the controller stability.

Application of feedback control to car noise cancelation

Feedback control is predominantly used to minimize the effects of road noise. Sano et al. [5] designed and implemented a SISO feedback control system for boom noise control of a Honda station wagon. The system mainly attempts to control the boom noise at 40 Hz present in the front seats of the car, which is due to the first acoustic longitudinal mode of the vehicle's enclosure. The feedback system employs a single-error microphone, positioned under the front seat and the two front door loudspeakers of the car's built-in speaker system (the two speakers are driven in-phase, hence the control system is SISO) to achieve noise reduction up to 10 dB. The authors observed an undesired side effect in the rear seats, where the boom noise was increased by 3 dB. To avoid this, they proposed a simultaneous fixed feedforward control system, which uses the previous system's error microphone as a reference signal to minimize the boom noise present at the rear seats. The secondary system utilized the two rear door loudspeakers from the vehicle's built-in speaker system driven in-phase. This approach managed to achieve a 10 dB reduction of the boom noise at the front seats while avoiding the increase of sound level at the rear seats, where the boom noise is not significant.

Similar to feedforward control, the performance of SISO feedback control is largely limited by increasing frequency and the enclosure size. The theory involved with employing MIMO feedback control for road noise control inside automobiles was first presented by Elliot and Sutton in [4]. Recent work on this approach including a practical investigation was carried out by Cheer et al. in [6] and [20], where the authors utilized a nonrigid car cabin mock-up. With eight error sensors and four control sources, the MIMO system

required a total of 32 FIR filters. The authors simulated road noise using uncorrelated structural vibrations for which the system managed to cancel an increased number of acoustic modes compared to the modal feedback controller. When the road noise was simulated using uncorrelated point sources, the system was capable of canceling all the active modes up to a control bandwidth of 100 Hz. The performance of the above system in a practical automobile environment was also

investigated by Cheer et al. [6], [19] inside a small city car. In this work, the authors utilized 16 error microphones (eight on the floor with a pair near each tire, and eight on the seats with a pair on each headrest) and the four built-in door loudspeakers as control sources. The system performance was synthesized offline based on the transfer function measurements inside the car.

The authors managed to achieve significant noise reduction up to 8 dB in the low-frequency range and an average reduction of 3 dB between 80–200 Hz where the road noise is prominent.

As discussed in the section “Application of Feedforward Control to Car Noise Cancelation,” a recent approach [21] to improve control bandwidth and reduce system complexity is to simplify the global control requirement by regional control where the control region is shrunk to a small area around the head position(s). In [22], the authors implemented a regional feedback control system in a Ford S-Max employing a horizontal grid array of 25 error microphones positioned in front of the headrest on the front passenger seat and two control loudspeakers mounted on the headrest. For smooth driving conditions at 80 km/hour, the regional control system achieved noise reduction up to 300 Hz. Also, when the error was only averaged over four microphones close to typical ear positions, the control bandwidth was

Similar to feedforward control, the performance of SISO feedback control is largely limited by increasing frequency and the enclosure size.

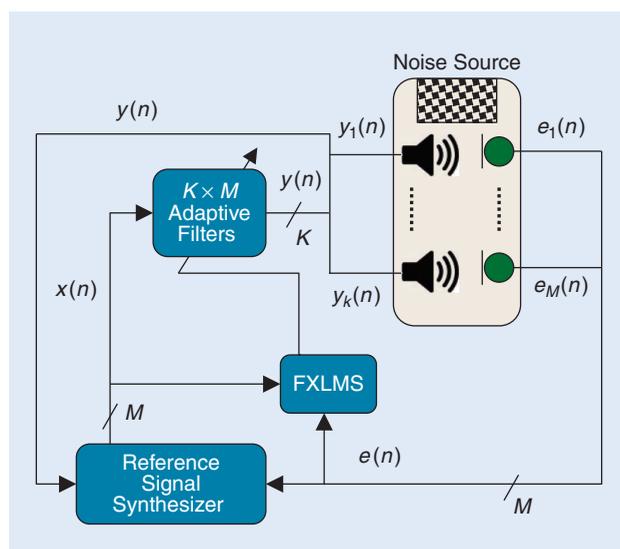


FIGURE 7. A MIMO adaptive feedback ANC system using M error sensors and K secondary sources.

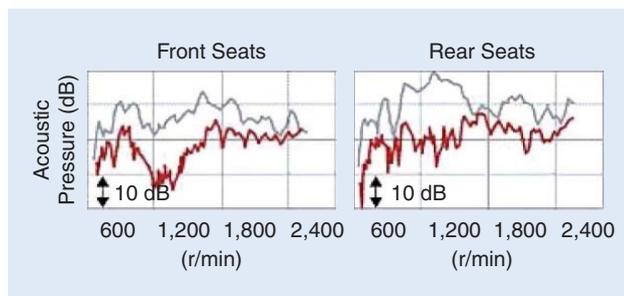


FIGURE 8. Noise reduction levels of the HONDA ANC system at (a) front seats and (b) rear seats. (Figure adapted from [36] and [35] and used courtesy of Honda.)

extended up to 500 Hz, revealing the potential advantages of regional control systems.

Commercial systems

In the commercial automotive space, noise control is still predominantly achieved via passive control. However, to overcome limitations related to passive control, more companies are increasingly applying active control to mass production. Commercial active control systems are typically applied for both noise and vibration control. While ANC utilizes an acoustic system, AVC typically comprises an active engine mount (ACM), which not only reduces vibrations but also reduces noise inside the cabin. More recently, the concept of active sound control (ASC) was introduced to commercial automotive solutions, which improves the driving experience by synthesizing certain sounds that are essential for the perceptual sound quality inside/outside the car. This process, which is similar to ANC, typically uses adaptive algorithms to change the coefficients of a set of digital filters such that not only are some selected frequencies canceled by secondary loudspeaker(s) generating an inverse disturbance signal(s), but others are controlled to a predetermined level, or even enhanced [12], [23]. In this section, we present the chronological advancement of active control in commercial automotive applications with the main focus on ANC.

Research and development of ANC became popular in the latter half of 1980 [1]. The earliest ANC systems were feedforward arrangements based on MIMO FXLMS for tonal engine noise (or booming noise) control inside cars [24]. Implementation of such a system was initially carried out in collaboration with the University of Southampton and Lotus Engineering, where four loudspeakers were adjusted at the engine's firing frequency and its harmonics to minimize the mean-square pressure at eight error microphones located on the headrests [25], [26]. ANC in mass-produced vehicles was first introduced by Nissan in 1991 [27] for booming noise, where a limited grade Nissan midsize car was optionally installed with a separate ANC system that consisted of additional loudspeakers, microphones, and a MIMO FXLMS

control system. The cost of implementation was quite high and the resulting level of noise reduction was not deemed to be significant. Therefore, at the time, it was not generally accepted as a useful technology.

In addition to engine booming noise reduction, research has also been actively carried out on road noise reduction since the early 1990s [28], [29]. In 2000, Honda introduced ANC for low-frequency narrowband road noise control. It was applied as standard equipment in a station wagon, where a fixed feedback controller based on control engineering theory was utilized [5].

Nearly a decade after Nissan's attempt, commercial interest in engine booming noise reduction started regaining attention due to the integration of the ANC system to the vehicle's built-in audio system. In 2003, Honda introduced ANC for booming noise caused by the Honda V6 engine model, which employed the variable cylinder management (VCM) technology to improve fuel economy by providing three-cylinder operation [30], [31]. This ANC system employed an adaptive notch filter-based MIMO feedforward controller and was combined with an active control engine mount (ACM) to reduce vibrations. Currently all VCM engine models from Honda are equipped with ANC and ACM. In 2006, Honda combined an ASC

system with their existing ANC solution for engine booming noise control. The ASC system was introduced to improve the internal cabin sound by synthesizing engine acceleration sounds for speeds above 2,500 r/minute such that it delivers a sporty feel to the driver [32]. In 2008, both Toyota [33] and GM introduced ANC for booming noise in a midsize car and a midsize SUV, respectively. Both

solutions were based on adaptive MIMO feedforward controllers. Soon after in 2009, Nissan reintroduced a MIMO ANC system based on adaptive feedforward control for engine booming noise in a midsize car [34]. In 2011, Honda introduced commercial solutions for low-frequency road noise by integrating an extra feedback controller (adaptive notch filter) to their existing booming noise controller (MIMO feedforward) [35]. In the frequency range below 100 Hz, this system is claimed to achieve 10 dB reduction of noise level inside a midsize car [36] (see Figure 8). The aforementioned road noise controller by Honda was updated in 2015 with an expanded low-frequency range [37].

In addition to the ANC solutions provided by automobile manufacturers, leading audio system developers such as Bose and Harman have also developed noise management solutions for automobiles. The Bose Active Sound Management System (ASM) is an example for such a solution [38]. The main technologies used in ASM are Bose Engine Harmonic Cancellation (EHC), Bose Engine Harmonic Enhancement (EHE), and Bose Rapid Mode Transition (RMT). The EHC technology is an ANC solution that minimizes booming noise utilizing a feedforward control system [39]. The

We present the chronological advancement of active control in commercial automotive applications with the main focus on ANC.

EHE and RMT systems are both ASC solutions that synthesize artificial sound to improve the driving experience. The EHE technology provides desirable linear (or sporty) sounds by masking sound anomalies that occur during acceleration [40] while the RMT technology provides additional control parameters to synthesize a seamless sound experience during variable powertrain and cabin modes (e.g., cylinder deactivation/reactivation, hybrid operation) [38]. Bose ASM was previously available only for vehicles with Bose sound system hardware however, since 2013, it was released as a software solution integrated in a chip for global auto manufacturers. Currently, Bose ASM is integrated in vehicles manufactured by GM, Nissan, Audi [41], Porsche etc., particularly in their luxury divisions.

Introduced recently in 2015, HALOsonic is another commercially available noise management solution provided by Harman International and Lotus Cars, which comprises a suite of four technologies to enhance the in-car audible environment and improve pedestrian safety [42], [43]. The two technologies directly related to ANC are the Road Noise Cancellation (RNC) system and the Engine Order Cancellation (EOC) system. The RNC system is a broadband feedforward control system with accelerometers as reference sensors. It is based on road noise cancellation solution originally presented in [28]. The HALOsonic EOC system, is a combined system with feedforward control to reduce noise due to engine rotations and a feedback controller to reduce noise due to internal combustion engine and exhaust components. Note that the design and specifications of the above systems may vary based on the size, shape, and cost of the vehicle model of interest. The remaining two technologies of HALOsonic focus on ASC or electric sound synthesis in quiet cars (e.g., electric cars, hybrid cars). This is done in two areas; 1) internally, to improve the passenger experience and 2) externally, to improve the safety of pedestrians. The internal (iESS) system helps synthesize an exhilarating engine sound that adds a very emotional element to the overall driving experience. It helps reinstate original engine sound in case of sound loss due to original equipment manufacturer (OEM) features such as downsized engines and the use of turbochargers. iESS also offers multiple engine sound modes (e.g., normal, moderate, and sporty engine sound) for the same car, thereby enhancing car occupants' emotional experience. The external (eESS) system mainly provides safety to quieter cars, where the active system is optimized to operate in urban environments with the greatest risk of a collision with pedestrians, especially high-risk groups such as the elderly, children, cyclists, and particularly the blind and their guide dogs. eESS helps automakers comply with governmental safety regulations. The sound of a car engine is an integral part of the experience behind the wheel and plays a crucial role in defining the DNA of the car. eESS is

capable of creating custom-designed engine sounds, thereby helping to retain an OEM-specific sound DNA for the car [42].

Practical limitations of current ANC systems

In this section, we discuss the main limitations related to mass production of ANC systems inside automobiles. As reviewed in [44], these include

- effects due to constrained number of microphones/speakers
- stability issues in feedback control systems
- system latency
- uncertainties
- ANC system integration issues
- system production tuning issues.

A brief discussion on each of the aforementioned constraints is as follows. Due to cost restrictions, the numbers of speakers and microphones employed in ANC systems are limited. Currently, a typical production set-up consists of four microphones and four–five speakers, which can only achieve

global noise control over 30–250 Hz (the lower limit is determined by the speaker characteristics, whereas the higher limit is determined by vehicle interior and component placement). Control up to at least 200 Hz is often desired because 200 Hz is equivalent to the firing frequency of an I4 engine at 6,000 r/minute, which is the dominant cause of booming noise. The active number of engine orders of a typical automobile is however much higher, and with new technologies like cylinder deactivation,

there exist extra noise components that need to be addressed. Therefore, improved ANC performance requires increased numbers of microphones/speakers with increased computational requirements, memory, and higher tuning efforts due to the increased complexity.

Another practical constraint that affects the performance of ANC systems, particularly employing feedback control loops, is stability. To minimize side effects due to stability, it is important to carefully calculate the sensitivity function of the closed loop. At present, there exist advance modeling techniques to understand the system behavior through computer-aided technology [41], which is expected to make significant progress in the coming years. Within the working frequency range of an ANC system, another practical limitation that arises is system latency. This is the signal latency added by processes like analog-to-digital and digital-to-analog conversion, and digital signal processing (DSP) latency. It is a difficult task to pinpoint the exact instance when system latency starts to deteriorate the ANC performance. In practice, an ANC system latency (latency in the microphone input-ANC processing-speaker output loop) of 2 ms is considered to be acceptable while 3 ms is the upper limit to avoid significant degradation [44]. ANC performance in realistic automobile cabins is often affected by uncertainties such as the number and placement

In practice, the extension of a SISO feedback system to a MIMO feedback system is somewhat complex due to the need to calculate the eigenvalues of the open-loop response to assess the controller stability.

of seated passengers, the opening and closing of windows/doors, and vehicle interior production tolerances. To ensure consistent ANC performance, it is important to improve system robustness while minimizing uncertainties. The typical approach to achieve this is via measuring and modeling different components of overall uncertainty as accurately as possible under realistic conditions, and setting the ANC system parameters to guarantee robustness.

In addition to the aforementioned concerns, another key aspect that affects the implementation of ANC in production vehicles is the integration of the ANC system to the existing audio system for parallel usage. Initially, ANC solutions were added as an extra control unit (DSP audio amplifier) causing no impact on the existing audio system, however it was soon deemed to be ineffective with regard to cost, weight, and space. A subsequent ANC solution was to add some dedicated processing resources for ANC into the existing audio system (i.e., plug-in module for head unit with dedicated DSP), which omits the need for an additional control unit while minimizing added weight and space. A more recent and improved ANC solution is to fully integrate ANC in the form of software into the existing audio system. This is done by 1) adding ANC as a software on the amplifier without extra processing unit (e.g., Analog devices' SHARC processor) or 2) integrating functional software using system-on-chip (SOC) solutions (e.g., NXP chip for the Bose Active Sound Management System). While the commercial application of the aforementioned solutions are still limited, they are expected to be utilized more broadly in the near future. One more issue that affects ANC implementation in mass-produced vehicles is system tuning during production. This involves the measurement of secondary path transfer functions and the determination of algorithm parameters such as the number of engine orders to cancel. With the uncertainty issues mentioned above, and multiple available powertrains, it is important to tune the ANC system for each of the vehicle variants. This task requires a lot of time and manpower and, with increasing demand to shorten the vehicle development period, there is an urgent need to opt for advanced CAE technologies that enable faster tuning.

Spatial sound field control in ANC

Up to this point, we have only discussed ANC techniques that model the noise field in terms of acoustic modes and structural-acoustic coupling. By now, it is common knowledge that the aforementioned ANC is effective at low frequencies, but have limitations at high frequencies due to increased requirement of microphones/speakers and related cost. Recently, research has been carried out to model noise fields in an alternative domain such that characteristics like sparsity can be exploited to bring down the minimum requirement of microphones/speakers. This concept is based on spatial sound field control, and initial research on this topic is described next.

Spatial sound field control involves acoustic control over a continuous spatial region utilizing a finite set of transducers distributed over the region of interest. Two well-developed techniques to achieve spatial sound field control are wavefield synthesis [45] and higher-order ambisonics (HOA) [46]. Currently, HOA is the only technique utilized for spatial noise cancelation inside automobiles, and, therefore, the overview given in this section will be limited to HOA. HOA is conceptually based on the cylindrical/spherical harmonics-based solution to the wave equation. This solution represents the incident pressure at any arbitrary point x within a control region of radius R , with respect to its origin by [47]

$$p(x, k) = \begin{cases} \sum_{n=-N}^N \alpha_n(k) J_n(kr) e^{in\phi} & \text{2-D region with } x = (r, \phi) \\ \sum_{n=0}^N \sum_{m=-n}^n \alpha_{nm}(k) j_n(kr) Y_{nm}(\theta, \phi) & \text{3-D region with } x = (r, \theta, \phi) \end{cases}, \quad (5)$$

where $k = 2\pi f/c$ represents the wave number with f and c representing frequency and speed of sound respectively, α denotes the HOA harmonic coefficients, $J_n(\cdot)$ and $j_n(\cdot)$ represent the cylindrical and spherical Bessel functions of order n , respectively, $Y_{nm}(\cdot)$ denotes the spherical harmonic function, and $N = \lceil kR \rceil$ is the summation's truncation limit (commonly referred to as *sound field order*) derived based on inherent properties of Bessel functions. The main advantage of the aforementioned decomposition is that it gives the ability to record or produce an entire continuous spatial sound field by considering only a finite set of coefficients. When recording a spatial sound field, these coefficients have a direct relationship with the microphone outputs in the

Another key aspect that affects the implementation of ANC in production vehicles is the integration of the ANC system to the existing audio system for parallel usage.

form $\mathbf{P} = \mathbf{T}\alpha$, where α is a vector of recorded sound field coefficients, \mathbf{T} is a transformation matrix, and \mathbf{P} is a vector of microphone recordings. Similarly, when producing a sound field, the above coefficients have a direct relationship with the loudspeaker driving signals in the form $\alpha = \mathbf{T}_1 \mathbf{W}$, where α is now a vector of desired sound field coefficients, \mathbf{T}_1 is a transformation matrix, and \mathbf{W} is a vector of loudspeaker driving signals. Generally, when recording/producing an N th-order sound field, there exists a minimum requirement of $(2N + 1)$ or $(N + 1)^2$ sensors/loudspeakers for two-dimensional (2-D) and three-dimensional (3-D) sound fields, respectively. This is to avoid the undesired effects of spatial aliasing.

When HOA-based spatial sound field control is occupied in ANC, the residual field, the noise field, and the secondary sound field are first decomposed in to cylindrical/spherical harmonic coefficients. For example, in 3-D ANC, the frequency transform of (6) is decomposed into

$$e_{nm}(k) = d(k)_{nm} + y'(k)_{nm}. \quad (6)$$

The input and output of the adaptive controller are therefore spherical harmonic coefficients rather than the direct error sensor outputs or control speaker weights. As a result, the standard block diagram for feedforward and feedback systems needs to be updated by additional modal transformation blocks. In [48], Spors et al. designed and simulated a 2-D HOA-based massive feedforward ANC system with 80 reference sensors, 80 error sensors, and 80 loudspeakers, mainly for use in room noise cancelation.

A 2-D feedback control system following the HOA technique was recently proposed by Zhang et al. using 11 error sensors and 11 control sources [49]. This feedback system was later extended with a sparse FXLMS controller, particularly for the use in spatially sparse noise fields [50]. The main advantage of HOA-based ANC as observed in both feedback and feedforward systems mentioned above is the significant improvement of convergence time and the significant decrease of spatially averaged residual signal energy. However, due to the relationship $N = [kR]$, the minimum requirement of sensors/speakers to control a sizable enclosure is impractically high, especially if the system is to be utilized for noise reduction inside automobiles.

Application of spatial sound field control to car noise cancelation

In [16], Chen et al. investigated the applicability of spatial regional control in ANC inside automobiles. The main purpose of the study was to derive the performance bounds of a feedback system with the automobile's built-in speakers utilized as control sources. Results were synthesized using a fixed offline system based on transfer functions and noise measurements done at the front-left headrest of a Ford Falcon XR6 (see Figure 9). The sensor array used was a commercially available 32-microphone spherical array (Eigenmike). While the previously discussed spatial ANC systems preferred a spherical array of control sources, the proposed system simplified this constraint to the vehicle's own audio system, based on a novel model for the primary noise field. This model was derived utilizing the spherical harmonic decomposition of the recorded noise field such that it represents the primary noise field in terms of an alternative set of basis functions. Based on a diverse set of noise measurements obtained inside the car (e.g., engine only, AC only, road noise at specific speeds) the authors found out that the noise field

inside a vehicle is generally sparse in terms of the proposed noise model. In fact, for the head-sized region of interest, it was observed that only a single-noise mode was active at all times. Therefore, it was predicted that the vehicle's built-in two-channel audio system (the Ford Falcon XR6 has four loudspeakers with stereo control) will be sufficient to attenuate the active noise mode. Figure 10 shows the noise reduction observed over frequency for four different driving conditions, where it's observed that noise reduction is relatively consistent with the attenuation levels varying between 35–15 dB across 50–500 Hz. These results are quite promising in terms of the potential use of a

vehicle's own audio system for effective ANC. A robustness analysis of this system however is still to be carried out. The theory of the aforementioned design was later extended to support multiple-region ANC control and tested in the same vehicle [51]. From this investigation, the authors concluded that a vehicle's integrated loudspeakers, when used as a stereo system, are only capable of canceling the noise field up to 200 Hz at the head positions of two seats simultaneously. To achieve similar reductions over four headrest positions

Right now, the main drivers for cost are hardware components, particularly the extra requirement for error/reference microphones and control loudspeakers.



FIGURE 9. The hardware setup for regional and spatial feedback control.

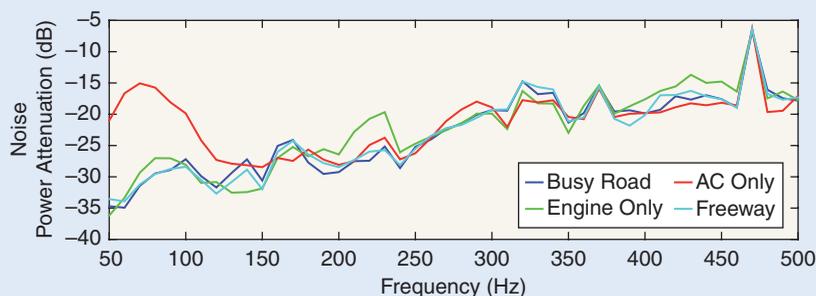


FIGURE 10. Performance bounds of spatial ANC: the noise power spectrum attenuation for four different driving conditions.

simultaneously, a minimum of four individually driven loudspeakers were needed.

Summary and future directions

In this article, a compact tutorial of ANC techniques was presented with a review of their application in reducing undesired noise inside automobiles. Some of the recent advances have demonstrated significant improvements in the noise reduction levels as well as the cost and implementation complexity. While the aforementioned techniques discussed may individually focus on a particular noise field (e.g., road noise only, engine noise only), it is proven through research and commercial products that a combination of these strategies can deliver significant benefits in realistic conditions.

Future opportunities for improving in-car ANC exist in 1) cost reduction, 2) practical implementation and commercialization of regional sound field control, 3) integration of regional ANC with future in-car infotainment systems, and 4) researching on alternative noise modeling techniques to bring down the system components. Right now, the main drivers for cost are hardware components, particularly the extra requirement for error/reference microphones and control loudspeakers. With the introduction of microelectromechanical systems (MEMS) microphones, and MEMS loudspeakers (e.g., Audio Pixels [52]), technically feasible low-cost ANC systems could be introduced, possibly with better performance. As mentioned previously, regional ANC is a well-researched topic that could reduce the overall system requirements. Regional ANC can be also extended for multiple regions serving individual passengers. Practical implementation and commercialization of these solutions are still minimal and has potential to reduce costs and improve efficiency. With the current global trend of instant connectivity, vehicles are evolving to provide infotainment systems rather than just radio. These include the availability of different wireless interfaces including Wi-Fi and Bluetooth, which forces all of the systems to move to digital. Future in-vehicle infotainment systems are predicted to be comprised of center stack computers [44] to process all types of media content and a network hub to serve multiple media/data streams, possibly on a per/seat, per/display basis. With the introduction of such systems, its essential for ANC to be reintroduced with appropriate low-latency audio processing that handles digital signals. Finally, another important future direction that could improve ANC efficiency is by looking for alternative modeling methods for the noise field inside the car. Even though the current ANC systems are largely restricted to low frequencies, an alternate model that describes noise fields in terms of a lower number of active modes may significantly enhance the system performance for the same number of microphones/speakers, specially when the noise field is directionally sparse.

Regional ANC is a well-researched topic that could reduce the overall system requirements and can also be also extended for multiple regions serving individual passengers.

Authors

Prasanga N. Samarasinghe (prasanga.samarasinghe@anu.edu.au) received her B.E. degree (with first-class honors) in electronic and electrical engineering from the University of Peradeniya, Sri Lanka, in 2009. She completed her Ph.D. degree at the Australian National University (ANU), Canberra in 2014. She is currently working as a research fellow at the Research School of Engineering at ANU. Her research inter-

ests include spatial audio, active noise control, and multichannel signal processing.

Wen Zhang (wen.zhang@anu.edu.au) received her B.E. degree in telecommunication engineering from Xidian University, Xi'an, China, in 2003, and her M.E. (with first-class honors) and Ph.D. degrees from the Australian National University, Canberra, in 2005 and 2010, respectively. From 2010 to 2012, she was a postdoctoral fellow at the Commonwealth Scientific and Industrial Research Organization Process Science and Engineering in Sydney, Australia. She is currently working as a research fellow in the College of Engineering and Computer Science at the Australian National University. Her research interests include spatial audio, binaural source localization, and active noise control. She was awarded the Discovery Early Career Researcher Award fellowship by the Australian Research Council in 2015.

Thushara D. Abhayapala (thushara.abhayapala@anu.edu.au) received his B.E. degree (with honors) in engineering in 1994 and his Ph.D. degree in telecommunications engineering in 1999, both from the Australian National University (ANU), Canberra. He is currently the deputy dean of the College of Engineering and Computer Science, ANU. He was the the director of the Research School of Engineering at ANU from January 2010 to October 2014 and the leader of the Wireless Signal Processing Program at the National ICT Australia from November 2005 to June 2007. His research interests are in the areas of spatial audio and acoustic signal processing, and multichannel signal processing. He has supervised more than 30 Ph.D. students and coauthored more than 200 peer-reviewed papers. He is an associate editor of *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. He is a member of the Audio and Acoustic Signal Processing Technical Committee (2011–2016) of the IEEE Signal Processing Society. He is a fellow of Engineers Australia.

References

- [1] P. A. Nelson and S. J. Elliott, *Active Control of Sound*. New York: Academic Press, 1991.
- [2] G. Cerrato, "Automotive sound quality: Powertrain, road and wind noise," *Sound Vibration*, vol. 43, pp. 16–24, Apr. 2009.
- [3] S. M. Kuo and D. R. Morgan, "Active noise control: A tutorial review," *Proc. IEEE*, vol. 87, no. 6, pp. 943–973, 1999.
- [4] S. J. Elliott and T. J. Sutton, "Performance of feedforward and feedback systems for active control," *IEEE Trans. Speech Audio Process.*, vol. 4, no. 3, pp. 214–223, 1996.

- [5] H. Sano, T. Inoue, A. Takahashi, K. Terai, and Y. Nakamura, "Active control system for low-frequency road noise combined with an audio system," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 7, pp. 755–763, 2001.
- [6] J. Cheer, "Active control of the acoustic environment in an automobile cabin," Ph.D. dissertation, Faculty of Engineering and the Environment, Univ. Southampton, United Kingdom, 2012.
- [7] D. Thompson and J. Dixon, "Vehicle noise," in *Advanced Applications in Acoustics, Noise and Vibration*. London, U.K.: Spon Press, 2004, ch. 6.
- [8] D. Robinette, M. Grimmer, J. Horgan, J. Kennell, and R. Vykydal, "Torque converter clutch optimization: improving fuel economy and reducing noise and vibration," *SAE Int. J. Engines*, vol. 4, no. 1, pp. 94–105, Apr. 2011.
- [9] M. R. Schroeder, "Frequency-correlation functions of frequency responses in rooms," *J. Acoust. Soc. Am.*, vol. 34, no. 12, pp. 1819–1823, 1962.
- [10] S. J. E. and M. E. Johnson, "Radiation modes and the active control of sound power," *J. Acoust. Soc. Am.*, vol. 94, no. 1453, pp. 2194–2204, 1993.
- [11] S. M. Kim, "Active control of sound in structural-acoustic coupled systems," Ph.D. dissertation, Univ. Southampton, Faculty of Engineering and the Environment, United Kingdom, Apr. 1998.
- [12] H. Sano, "Modern advancements in passive and active noise and vibration control technology in automobiles," in *Proc. 40th Int. Congr. InterNoise*, 2011, vol. 1.
- [13] D. R. Morgan, "An analysis of multiple correlation cancellation loops with a filter in the auxiliary path," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 28, no. 4, pp. 454–467, 1980.
- [14] J. Pan, C. Hansen, and D. Bies, "Active control of noise transmission through a panel into a cavity: I. Analytical study," *J. Acoust. Soc. Am.*, vol. 87, no. 5, pp. 2098–2108, 1990.
- [15] J. I. Mohammad, "The active control of random sound inside cars," Ph.D. dissertation, Univ. Southampton, Faculty of Engineering and the Environment, United Kingdom, 2006.
- [16] H. Chen, P. Samarasinghe, T. D. Abhayapala, and W. Zhang, "Spatial noise cancellation inside cars: Performance analysis and experimental results," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2015, pp. 1–5.
- [17] S.-H. Oh, H.-S. Kim, and Y. Park, "Active control of road booming noise in automotive interiors," *J. Acoust. Soc. Am.*, vol. 111, no. 1, pp. 180–188, 2002.
- [18] L. Eriksson, "Recursive algorithms for active noise control," in *Proc. Int. Symp. Active Control of Sound Vibration*, 1991, pp. 137–146.
- [19] J. Cheer and S. J. Elliott, "Multichannel control systems for the attenuation of interior road noise in vehicles," *Mech. Syst. Signal Process.*, vol. 60, pp. 753–769, Aug. 2015.
- [20] J. Cheer and S. Elliott, "Multichannel feedback control of interior road noise," in *Proc. Meetings on Acoustics Acoustical Society of America*, 2013, vol. 19, no. 1, p. 030118.
- [21] S. J. Elliott, W. Jung, and J. Cheer, "The spatial properties and local active control of road noise," in *Proc. Euro-Noise*, 2015, 2189–2194.
- [22] J. Cheer, S. Elliott, and W. Jung, "Sound field control in the automotive environment," in *Proc. 3rd Int. ATZ Automotive Acoustics Conf.*, Zurich, Switzerland, 2015.
- [23] L. E. Rees and S. J. Elliott, "Adaptive algorithms for active sound-profiling," *IEEE Trans. Audio Speech Lang. Process.*, vol. 14, no. 2, pp. 711–719, 2006.
- [24] S. Elliott, I. Stothers, P. Nelson, A. McDonald, D. Quinn, and T. Saunders, "The active control of engine noise inside cars," in *Proc. Inter-Noise and Noise-Control Congr. Conf.*, 1988, vol. 1988, no. 3, pp. 987–990.
- [25] S. J. Elliott, "A review of active noise and vibration control in road vehicles," *ISVR Tech. Memorandum*, vol. 981, 2008. [Online]. Available: <http://eprints.soton.ac.uk/65371/>
- [26] A. McDonald, S. Elliott, and M. Stokes, "Active noise and vibration control within the automobile," in *Proc. Int. Symp. Active Control of Sound Vibration*, 1991, pp. 147–156.
- [27] S. Hasegawa, T. Tabata, A. Kinoshita, and H. Hyodo, "The development of an active noise control system for automobiles," SAE Tech. Rep., 1992.
- [28] T. J. Sutton, S. J. Elliott, A. M. McDonald, and T. J. Saunders, "Active control of road noise inside vehicles," *Noise Contr. Eng. J.*, vol. 42, no. 4, pp. 137–147, 1994.
- [29] R. J. Bernhard, "Active control of road noise inside automobiles," in *Proc. Inter-Noise and Noise-Control Congr. Conf.*, 1995, vol. 1995, no. 5, pp. 21–32.
- [30] T. Inoue, A. Takahashi, H. Sano, M. Onishi, and Y. Nakamura, "NV countermeasure technology for a cylinder-on-demand engine-development of active booming noise control system applying adaptive notch filter," SAE Tech. Rep., 2004.
- [31] H. Matsuoka, T. Mikasa, and H. Nemoto, "NV countermeasure technology for a cylinder-on-demand engine-development of active control engine mount," SAE Tech. Rep., 2004.
- [32] Y. Kobayashi, T. Inoue, A. Takahashi, and K. Sakamoto, "Active sound control in automobiles," in *Proc. Inter-Noise and Noise-Control Congr. Conf.*, 2008, vol. 2008, no. 4, pp. 5001–5009.
- [33] H. Morita, T. Mochizuki, and H. Yoshida, "Development of active noise control-mechanism analysis and countermeasure of the waterbed phenomena," in *Proc. Soc. Automotive Engineers of Japan Annu. Congr.*, 2009, pp. 15–18.
- [34] N. Masahiro, O. Daisuke, H. N. Yoshiharu Nakaji, M. Masahiro, and H. Mineura, "Development of anc technology for reconciling acceleration sound with fuel economy and quietness," in *Proc. Soc. Automotive Engineers of Japan Annu. Congr.*, 2010, pp. 7–10.
- [35] A. Takahashi, T. Inoue, K. Sakamoto, and Y. Kobayashi, "Integrated active noise control system for low-frequency noise in automobiles," in *Proc. Inter-Noise and Noise-Control Congr. Conf.*, 2011, vol. 2011, no. 6, pp. 2105–2113.
- [36] Honda, New Zealand. (2014) Honda active noise cancellation. [Online]. Available: <http://www.honda.co.nz/technology/driving/anc/>
- [37] K. Sakamoto and T. Inoue, "Development of feedback-based active road noise control technology for noise in multiple narrow-frequency bands and integration with booming noise active noise control system," *SAE Int. J. Passenger Cars-Mech. Syst.*, vol. 8, no. 1, pp. 1–7, Apr. 2015.
- [38] Bose Corporation. (2016) Bose active sound management technology [Online]. Available: <http://www.bose.com/prc.jsp?url=/automotive/innovations/asm.jsp>
- [39] A. Ganeshkumar. (2011 Sept.) Engine harmonic cancelling system and operating method thereof [Online]. Available: <http://www.sumobrain.com/patents/WO201111247A1.html>
- [40] C. M. Hera, "Engine harmonic enhancement control," U.S. Patent 9 177 544, Nov. 3, 2015.
- [41] R. Schirmacher, R. Kunkel, and M. Burghardt, "Active noise control for the 4.0 TFSI with cylinder on demand technology in Audi's S-series," SAE Tech. Rep., 2012.
- [42] Harman International. (2016) Halosonic: Noise management solutions [Online]. Available: <http://www.halosonic.co.uk/>
- [43] Harman International. (2015) Halosonic: Future of sound [Online]. Available: <http://harmaninnovation.com/blog/halosonic-future-sound/>
- [44] R. Schirmacher, "Current status and future developments of ANC systems," *SAE Int. J. Passenger Cars-Mech. Syst.*, vol. 8, no. 3, pp. 892–896, June 2015.
- [45] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. 124th Audio Engineering Society Convention*, 2008, pp. 17–20.
- [46] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 6, pp. 697–707, 2001.
- [47] E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*. London, U.K.: Academic Press, 1999, pp. 115–125.
- [48] S. Spors and H. Buchner, "Efficient massive multichannel active noise control using wave-domain adaptive filtering," in *Proc. IEEE 3rd Int. Symp. Communications, Control and Signal Processing*, 2008, pp. 1480–1485.
- [49] J. Zhang, W. Zhang, and T. D. Abhayapala, "Noise cancellation over spatial regions using adaptive wave domain processing," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2015, pp. 1–5.
- [50] J. Zhang, T. D. Abhayapala, P. Samarasinghe, W. Zhang, and S. Jiang, "Sparse complex fxlms for active noise cancellation over spatial regions," in *Proc. Int. Conf. Acoustics, Speech and Language Processing*, Shanghai, China, 2016.
- [51] H. Chen, P. Samarasinghe, and T. D. Abhayapala, "In-car noise field analysis and multi-zone noise cancellation quality estimation," in *Proc. Asia-Pacific Signal and Information Processing Association Annu. Summit Conf.*, 2015, pp. 773–778.
- [52] Audio Pixels Limited. (2016) Audiopixels technology [Online]. Available: <http://www.audiopixels.com.au/index.cfm/technology/>

Robert Hult, Gabriel R. Campos, Erik Steinmetz,
Lars Hammarstrand, Paolo Falcone, and Henk Wymeersch

Coordination of Cooperative Autonomous Vehicles

Toward safer and more efficient road transportation

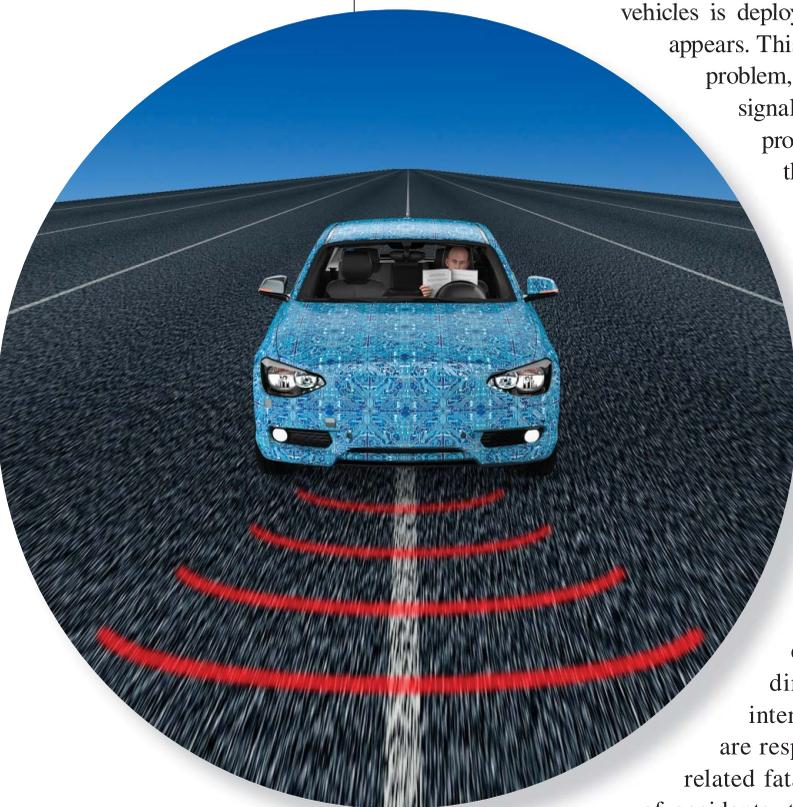
While intelligent transportation systems come in many shapes and sizes, arguably the most transformational realization will be the autonomous vehicle. As such vehicles become commercially available in the coming years, first on dedicated roads and under specific conditions, and later on all public roads at all times, a phase transition will occur. Once a sufficient number of autonomous vehicles is deployed, the opportunity for explicit coordination appears. This article treats this challenging network control problem, which lies at the intersection of control theory, signal processing, and wireless communication. We provide an overview of the state of the art, while at the same time highlighting key research directions for the coming decades.

Introduction

The purpose of intelligent transportation systems (ITS) is to leverage advances in information technology to alleviate major problems in the current road traffic system. Focus areas include the prevention and mitigation of accidents, reduction of greenhouse gas emissions, and efficiency in terms of energy and infrastructure utilization.

A particularly problematic subset of traffic scenarios in terms of both safety and efficiency is those where vehicles must coordinate the use of a common resource, such as intersections, roundabouts and on ramps. These are responsible for a significant fraction of traffic-related fatalities and injuries [1]. Due to the high risk of accidents, these traffic scenarios are among the most regulated, with vehicles guided simultaneously by traffic lights, signs, road markings, and right-of-way rules. The problems of traffic fatalities and inefficiency are expected to become even more pressing in the future, as the global number of light vehicles (e.g., passenger cars and light trucks) is forecast to rapidly increase. Proportional expansion of road infrastructure is undesirable in most countries, and might not even be possible given continued urbanization and the associated increase in population density. Hence, there is great interest to

SKY IMAGE LICENSED BY GRAPHIC STOCK
@STOCKPHOTO.COM/POSTERIORI



Digital Object Identifier 10.1109/MSP.2016.2602005
Date of publication: 4 November 2016

improve safety, energy, and traffic efficiency on the existing and planned road infrastructure.

Many of the aforementioned problems are caused by the human involvement in the coordination of traffic. Studies have shown that over 90% of traffic accidents are completely, or in part, due to human error [2]. This has led to a progressive shift in responsibilities from the human driver to dedicated control systems, most recently in the form of autonomous vehicles, which aim to provide more efficient, comfortable, and virtually accident-free road traffic. Autonomous vehicles are still limited in terms of their sensing and coordination capabilities, as their actions depend on the on-board sensory data and models of other vehicles' behavior. As an example, a summary of the Urban Grand Challenge [3] mentioned that a number of incidents could have been avoided if vehicles could anticipate the behavior of other vehicles, and that vehicles should cooperate for autonomous driving to reach its full potential. The benefit of cooperation was already recognized in a parallel track in vehicle automation, and platooning, which instead of complete autonomy promotes information sharing between vehicles and joint decision making. In a platoon, the vehicles rely on vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication to share information regarding the environment and internal states and choose safe and efficient control policies jointly [4].

The two tracks for automating vehicles have thus followed different approaches: one (the platooning track) explicitly relies on communication among vehicles, while the second (the autonomous vehicle track) does not. With the adoption of the IEEE 802.11p Standard, as well as the possibilities of V2V and V2I communication and other services under the future 5G wireless standard [5], the two tracks are expected to merge, leading to a new type of large-scale wireless networked control system. This merging will likely take place in a piecemeal fashion, with first a ubiquitous availability of wireless communications, and only later the gradual introduction of cooperative autonomous vehicles [6]. These vehicles will drive autonomously, but at the same time be able to leverage their communication capabilities for cooperative planning and control, as well as cooperative perception and sensing, thus eliminating many of the traffic-safety and efficiency problems. The design and operation of such networks of cooperating vehicles place enormous demands on the control, communications, and sensing subsystems, as they must operate in harmony across different brands and types of vehicle, with limited margin for error.

In this article, we give an overview of the coupling between control, communication, and sensing, as visualized in Figure 1. We provide a survey of the different control approaches and their associated signal processing challenges. We hope that this article can provide an introduction for signal processing professionals to the control-theoretic aspects of vehicle coordination and pave the way for a tighter collaboration.

Problem formulation

The problem of coordinating a set of vehicles can be phrased as calculating the best control trajectories for the individual vehicles that allow them to safely reach their destination in finite time (e.g., within a few tens of seconds). In general, any solution should meet the basic requirements of safety (i.e., no collisions occur) and liveness (i.e., destinations are reached eventually), while optimizing some performance metric. The most important requirement is safety. Hence, vehicles may never be steered to states from which future collisions are unavoidable. Second, the coordination algorithm must guarantee that all vehicles are allowed to both enter and exit the coordination area in finite time so that permanent stops and traffic deadlocks are avoided. Finally, a performance criterion is necessary to favor one among multiple solutions. In summary, a coordination problem can be stated as a constrained optimal control problem, where a performance criterion is optimized with respect to the vehicles' control input trajectories, subject to safety and liveness requirements:

$$\text{minimize} \quad \text{performance criterion} \quad (1a)$$

$$\text{subject to} \quad \text{safety constraints,} \quad (1b)$$

$$\text{liveness constraints.} \quad (1c)$$

The constrained optimal control framework clearly allows one to conveniently accommodate performance, safety, and

We give an overview of the coupling between control, communication, and sensing for control of automated vehicles.

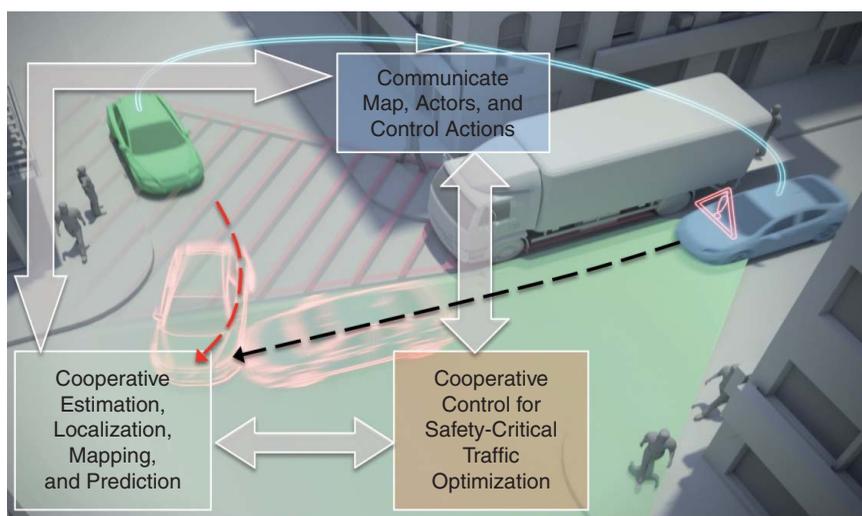


FIGURE 1. Vehicle coordination relies on a tight interaction between control, communication, and sensing.

liveness. However, as will be discussed later in this article, the partial lack of formal analysis tools limits their practical applicability. In particular, the impact of imperfect sensing data and communication impairments on stability and feasibility (i.e., the capability of finding a solution that meets safety and liveness requirements) of (1) is not completely understood under realistic communication protocols and sensing scenarios. In the absence of sensing and communication impairments, a simplified version of the generic problem (1) can be specified mathematically as follows.

Safety constraints

Consider a set of N vehicles (agents), whose motion is described by

$$\dot{x}_i(t) = f_i(x_i(t), u_i(t), t), \quad (2)$$

where $x_i \in \mathcal{X}_i \subseteq \mathbb{R}^n$ and $u_i \in \mathcal{U}_i \subseteq \mathbb{R}^m$ are the state and input/control vectors, respectively, $\dot{x}_i(t)$ denotes the time derivative of $x_i(t)$, and the sets $\mathcal{X}_i, \mathcal{U}_i$ reflect physical and design constraints. Examples of such constraints are acceleration limitations and the vehicles' minimum and maximum speeds. Examples of the state x_i are vectors comprising the vehicle's position and velocity in one, two, or three dimensions. Let $\mathcal{G}_i(x_i)$ describe the vehicle geometry, being the closed and compact set of spatial coordinates that vehicle i occupies when its state is x_i . Hence, a collision between two vehicles i and j occurs at time t if

$$\mathcal{G}_i(x_i(t)) \cap \mathcal{G}_j(x_j(t)) \neq \emptyset. \quad (3)$$

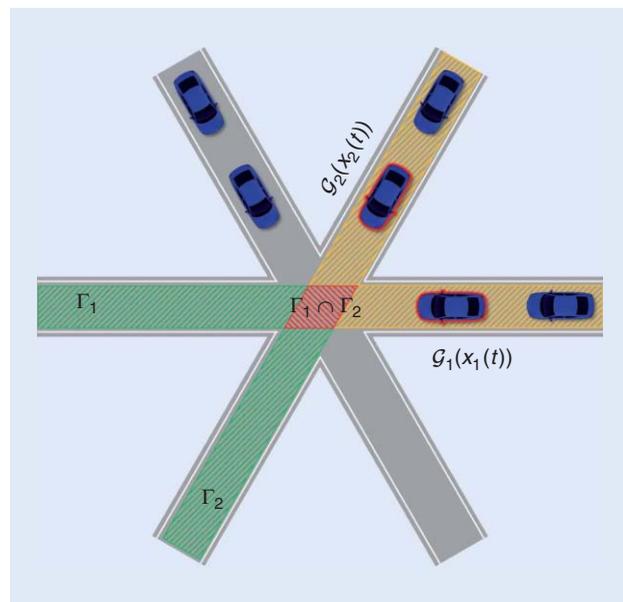


FIGURE 2. An example of a traffic coordination scenario at a three-way intersection. The geometries of vehicles 1 and 2 are highlighted in red, and their paths Γ_i are dashed and colored. The critical region, $\Gamma_1 \cap \Gamma_2$, is shown in dashed red, the target sets $\mathcal{T}_1, \mathcal{T}_2$ in green, and the paths before the critical region in yellow. The vehicle geometry $\mathcal{G}_i(x_i(t))$, depending on the vehicle state $x_i(t)$, is also depicted.

Furthermore, we denote by Γ_i the closed and connected set of spatial coordinates that comprises the paths of vehicle i so that a vehicle is on its path if $\mathcal{G}_i(x_i(t)) \subseteq \Gamma_i$. Provided that each vehicle stays on its path, a collision between vehicles can consequently only take place within a critical region where $\Gamma_i \cap \Gamma_j$, i.e., where paths fully or partly overlap (i.e., where paths are the same, cross, or merge).

Liveness constraints

Assume that all paths Γ_i are fixed and constant, and let the target set $\mathcal{T}_i \subset \Gamma_i$ be the set of spatial coordinates that vehicle i strives to reach (e.g., the road after an intersection, roundabout, or onramp). If $\mathcal{G}_i(x_i(t)) \subset \mathcal{T}_i$ is satisfied in finite time for all vehicles, the coordination is said to be deadlock-free, and all vehicles are eventually coordinated through the critical regions. For an illustration of the introduced notations, see Figure 2.

Performance criterion

In general, the cost for vehicle i , denoted by $J_i(x_i(t), u_i(t))$, can be expressed as

$$J_i(x_i(t), u_i(t)) = \int_0^{+\infty} \Lambda_i(x_i(t), u_i(t), t) dt,$$

where the stage cost $\Lambda_i(x_i(t), u_i(t), t)$ could be, e.g., instantaneous power consumption so that $J_i(x_i(t), u_i(t))$ is the total consumed energy. Other examples of $\Lambda_i(x_i(t), u_i(t), t)$ include a deviation from a target speed, or a measure of discomfort for the driver.

Overall problem and its receding horizon formulation

With the introduced notations and concepts, the N -vehicle optimal coordination problem (OCP) is now naturally formulated as the following infinite time, constrained optimal control problem.

Problem 1: OCP

$$\underset{\mathbf{x}(t), \mathbf{u}(t)}{\text{minimize}} \quad \sum_{i=1}^N J_i(x_i(t), u_i(t)) \quad (4a)$$

$$\text{subject to} \quad \dot{x}_i(t) = f_i(x_i(t), u_i(t), t), x_i(0) = x_{i,0} \quad (4b)$$

$$x_i(t) \in \mathcal{X}_i, u_i(t) \in \mathcal{U}_i \quad (4c)$$

$$\mathcal{G}_i(x_i(t)) \subseteq \Gamma_i \quad (4d)$$

$$\mathcal{G}_i(x_i(t)) \cap \mathcal{G}_j(x_j(t)) = \emptyset, \forall t \geq 0, i, j \neq i \quad (4e)$$

$$\exists T < \infty : \mathcal{G}_i(x_i(T)) \subseteq \mathcal{T}_i, \quad (4f)$$

where $\mathbf{x}(t) = [x_1^T(t), \dots, x_N^T(t)]^T$, $\mathbf{u}(t) = [u_1^T(t), \dots, u_N^T(t)]^T$ represent the states and control signal for each vehicle over the entire operating horizon (i.e., all $t \geq 0$). The OCP is thus the problem of finding the best admissible control inputs $u_i(t)$ for the dynamical systems $f_i(x_i(t), u_i(t), t)$ (4b), starting from the initial conditions $x_{i,0}$, that respect the state constraints (4c), while keeping all vehicles i within their paths Γ_i (4d), avoiding collisions between vehicles (4e), and eventually clearing

the coordination region (4f). The problem captures the basic requirements: safety through (4e), liveness through (4f), and performance through the objective function. Problem 1 can be conveniently reformulated in a discrete time domain by discretizing the systems dynamics (4b). Furthermore, to solve a finite dimensional problem, receding horizon control (RHC) or model predictive control (MPC) schemes can be used [7], where a finite time optimal control problem is solved every sampling time instant. In particular, as explained in Figure 3, RHC seeks future input and state trajectories at every sampling time instant over a finite time horizon, so as to minimize the cost function, subject to the constraints. The first element of the computed control input sequence is applied to the system, and, at the next time step, the problem is formulated and solved over a shifted time horizon. This RHC approach also allows us to account for the future, but by only committing the control action for the current time, we are able cope with limited disturbances (e.g., due to imperfect sensing or communication). This is important, as we will see in the next section.

Challenges in solving the coordination problem

Although finite dimensional, solving problem 1 in a receding horizon framework is extremely challenging, not only from the control perspective, but also due to imperfect communications as well as uncertainties induced by the sensors. While all these challenges are interrelated, we break them down as follows.

Control challenges

The main control-related challenges involve, first, the ability to compute good, feasible control actions, and, second, the ability to guarantee that the closed-loop system has certain desired properties. Regarding the former, note that the mathematical problem of finding the actions of N vehicles that allow them to pass the coordination zone without colliding is inherently a combinatorial problem. For a given initial configuration, a multitude of feasible temporal crossing orders (i.e., different orders in which one vehicle passes a coordination zone before another) might exist, and the optimal ordering can only be found by a structured exploration of the different alternatives. It is therefore no surprise that even the problem of finding a feasible solution to (4) is NP-hard in general [8]. Exact solutions to the OCP are therefore intractable for relevant problem sizes, and either heuristics or approximations must be employed. Regarding the properties of closed-loop control, there are several challenges. For instance, given the severity of constraint violations in the OCP, any controller needs to ensure persistent feasibility. If satisfied, this property ensures that any action taken does not put the system in a state from which no feasible actions exists, i.e., that no vehicle is ever put in a state from which a collision is inevitable. The closed-loop controller must also ensure stability, e.g., to make certain that the crossing order does not change every time the solution is recomputed. Additionally, the aforementioned issues are linked, as the computational challenges of the mathematical

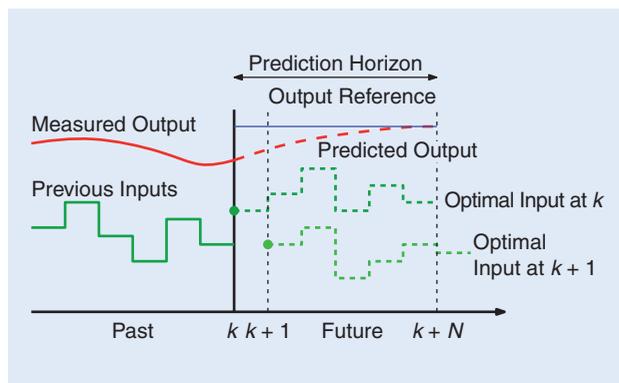


FIGURE 3. An illustration of an RHC scheme. The sketch depicts how, in an RHC algorithm, a present decision is made based on the current state of the system and its predicted future behavior.

coordination problem, for instance, might promote distributed solutions. In that case, the closed-loop controller needs to guarantee the aforementioned properties while the solution is obtained iteratively and possibly asynchronously over the wireless vehicular network.

Communications challenges

Irrespective of how the OCP is solved, information exchange is required between the involved entities (e.g., vehicles and possibly dedicated infrastructure). First and foremost, this includes the information necessary to formulate the OCP, e.g., the models of vehicle dynamics, road geometry, state measurements, and static and dynamic map information. In addition, it also includes information required to solve it, e.g., internal messaging in a distributed, iterative algorithm. Communication between entities will be greatly affected by the impairments associated with wireless channels, including the inherent randomness and correlation of the channel, interference due to simultaneous transmissions, and a limited communication range. In combination with limited communication resources (bandwidth, power), this results in packet drops and random latencies in packet arrivals. For automobile applications, it was pointed out that the current standards for V2V and V2I communication cannot ensure time-critical message dissemination in dense scenarios [9]. In general, it is desired to keep the communication load low, as wireless channel congestion is envisioned to be one of the major challenges related to vehicular networks [10], [11]. Overall, the communication subsystem forms a bottleneck for the OCP, related both to its formulation and the means by which it is solved.

Sensing challenges

The vehicles' own perception of their current locations and the positions of surrounding vehicles is fundamentally uncertain. Both are based on observations from sensors such as cameras, radar, lidar, the global navigation satellite system (GNSS), and inertial navigation sensors, which deliver observations that are corrupted by noise and clutter

(spurious nonobject detection). In addition, the sensors often fail to detect objects, e.g., vehicles and pedestrians, leading to uncertainty about whether all relevant objects are known to the sensing system. Moreover, as each autonomous vehicle is equipped with a different sensor setup with different types of observations, the accuracy of each vehicle's perception of the current traffic situation will typically vary over time and will not be coherent among the vehicles [12]. There are methods to handle time-varying and noncoherent uncertainties, but these require that an uncertainty description is communicated among the involved vehicles, further increasing the demand on the communication system. Even with perfect communication, it is still nontrivial to associate the information from one vehicle with other vehicles' local understanding of the same situation. In the literature, this is called the *data association problem*, which is known to be an NP-hard problem [13]. Processing and sharing a large amount of data requires suitable compression algorithms in combination with application-specific semantic representations, amenable for inclusion in the OCP. These different types of uncertainties are generally ignored in RHC solutions to the OCP, as long as the immediate future is relatively certain. However, in general we may not be able to guarantee performance, liveness, or safety of the resulting solutions. Moreover, the solutions may no longer be stable, which is undesirable from the point of view of the passengers.

Solving the coordination problem

Despite (or perhaps because of) the inherent difficulty of solving the OCP, many solutions have been presented in the context of automated vehicles. These works have been carried out by several communities, resulting in differences of focus and techniques. The resulting techniques can be classified into two groups: rule based and optimization based.

Rule-based solutions

In a large number of existing approaches, e.g., [14]–[16], the vehicle coordination problem is solved using a set of fixed rules, implemented through an interaction protocol. This protocol specifies the content and timing of communications, as well as the possible responses to actions of other participants. To simplify the set of rules, protocols generally assume that individual agents take on partial local responsibility (e.g., resolving rear-end collisions and lane keeping), while a coordination manager resolves any multipath conflicts at the intersection. A canonical protocol operates as follows: 1) a vehicle requests permission to enter the coordination zone at a given time with a given velocity; 2) the intersection manager takes the request and decides whether it can lead to a collision-free crossing (if so, the request is accepted; otherwise, it is denied) 3) when a vehicle's request is denied, it decelerates and sends a new request. Once a request is accepted, the vehicle applies a suitable control

action to meet specifications on when it is allowed to use the coordination zone. From these simple rules, it follows that only requests of vehicles with a safe option are accepted, while all vehicles for which no reservation can be found will slow down and eventually stop.

The benefits of rule-based schemes are the distribution of computation and the economic use of the communication resources (since the rule-set and interaction protocol is known to all participants). In terms of performance, rule-based solutions are generally only possible to evaluate a posteriori as the actions taken by the vehicles are generated implicitly by application of the rule set. The rules are usually claimed to be chosen to optimize some objective (commonly, throughput), but formal results are missing in most cases.

The general lack of formal guarantees in terms of the objective and constraints of the OCP forms the main weakness of rule-based solutions. Extensions of these works include liveness guarantees [17], and refinements on the individual control policies [18]. The approaches presented in [15], [16], [19], and [20] share similar concepts but differ in terms of the set of rules determining the priority of each vehicle. In summary, while rule-based methods may outperform current regulatory mechanisms, they most likely underutilize the potential of automated vehicles in coordination scenarios.

Optimization-based solutions

In this second group of solution approaches, the coordination problem is treated as a mathematical program from the outset and solved using standard tools and algorithms from optimal control. By doing so, one can potentially separate the feasibility and optimality aspects, and use general, multiobjective performance measures so derive formal guarantees for both performance and safety. However, as a consequence, the computational complexity issues are inherited from the original problem. The contributions of the surveyed papers [21]–[24] are therefore mainly reformulations, approximations, and heuristics that aim to remedy the computational tractability issues. One class of solutions [21], [22] casts an equivalent of the OCP as a safety verification problem. The verification problem entails determining the largest set of (infinite horizon) control actions that avoid any conflict at all future times, and is used in [21] to synthesize a least-restrictive supervisor for human drivers: if the verification fails, the supervisor overrides the human's command, e.g., desired acceleration. Determining the overriding control signals can be posed as a type of OCP wherein the objective corresponds to minimizing the total time needed to clear the intersection or the deviation from the desired control signal, given by the human driver [22]. A more general approach to the OCP was considered in [23], presenting a hierarchical decomposition of (4), where the problem is split up into one centralized time-slot allocation problem and several local vehicle-level optimal control problems. In the latter,

The benefits of rule-based schemes are the distribution of computation and the economic use of the communication resources.

each vehicle computes approximations of its local solutions, parameterized by its occupancy time interval T_i . Loosely speaking, the local optimal cost is expressed as a function of locally feasible T_i 's and transmitted to the central controller. The controller can then find the optimal values of T_i and broadcast this information to the vehicles. Using this approach, the OCP is posed as the search for approximately optimal, nonoverlapping occupancy time slots, which is a rather small mixed-integer optimization problem. Finally, yet another approach was taken in [24], where the combinatorial aspect is resolved through a cooperatively predetermined decision order (or priority), enabling sequential decision making. Once a decision order is agreed upon, the highest priority vehicle solves a local optimal control problem, ignoring all of the remaining vehicles. The solution is communicated to the second vehicle in the ordering, which uses it to solve its own problem: finding the best solution that crosses the coordination zone either before or after the first vehicle. In general, vehicle i in the order will have access to the occupancy intervals of all higher priority vehicles and solves a small local problem. An RHC extension of this approach was developed and demonstrated in [24], where at each time instance both the priority assignment and sequential decision making is repeated.

The primary benefits of the optimization-based approaches are the inherent flexibility and ease with which different and tunable objective functions, dynamics, and physical constraints are included in the design phase, but also modified in the operating phase of the coordination system. This gives to the designer, operator, or passenger the control over what kind of solution the system outputs, and the ability to change this during operation. Furthermore, extensive results regarding the issues of persistent feasibility, stability, and robustness of model-based and optimization-based control schemes (i.e., as in MPC) are available in the literature, as are approximate schemes with quantifiable suboptimality. The optimization-based coordination schemes could potentially leverage such results and formally provide the required safety guarantees. The major weakness of the optimization-based schemes is the complexity, directly inherited from the original formulation (4), which grows exponentially with the number of possible conflict relationships among the vehicles.

The role of signal processing in the OCP

From the aforementioned discussion, it is clear that the OCP explicitly relies on sensing and perception algorithms as well as on wireless communication for its formulation and solution, even though sensing and communication aspects have largely been ignored in the development of control algorithms. Conversely, specific control applications and their demands are usually not considered in the design of sensing

and perception algorithms, nor in the design of wireless communication systems. In this section, we describe recent progress in sensing and wireless communication, and how it relates to solving the OCP. Furthermore, we discuss the need for a tighter integration between the different subsystems, and present ideas on how smart signal processing can be utilized to achieve this.

Wireless communication

Current vehicular communication standards (IEEE Wireless Access in Vehicular Environments and ETSI ITS G5) rely on Wi-Fi-like communication over 10-MHz channels in

the 5.9-GHz band and have defined both periodic awareness messages and event-triggered safety messages. These standards can support low-rate (up to 10 Hz) broadcast messages between vehicles within a communication range of about 500 meters [10], but will fail under the high load of the ultrafast communication that is needed to solve the OCP. In contrast, OCP-like problems have been considered explicitly in 5G research [25], with assumed reliability of 99.9% and status updates of 100 ms, considering a steering frequency of 10 Hz. However, these numbers only relate to the dissemination of the final control signal, not the collection of information needed to pose the OCP, nor the iterative message

exchange needed to solve it nor do they consider scalability with a large number of vehicles. To get a rough indication for how many vehicles can be supported in a centralized implementation of the OCP, consider a communication system operating in time division multiple access (TDMA) mode, between N vehicles and a controller. This means that each vehicle is assigned a time slot where it during the uplink (UL) phase can transmit its state information to the controller. Assume these time slots last around $100 \mu\text{s}$, accounting for the actual payload (see Figure 4), as well as overhead (OH) in terms of guard intervals (GIs), training sequences (TSs) and cyclic redundancy check (CRC) bits (for

Signal processing can relieve the burden on communications by censoring less critical information, by tailored compression and semantic algorithms, and by assigning communication resources to those vehicles that are critical to the optimal control problem.

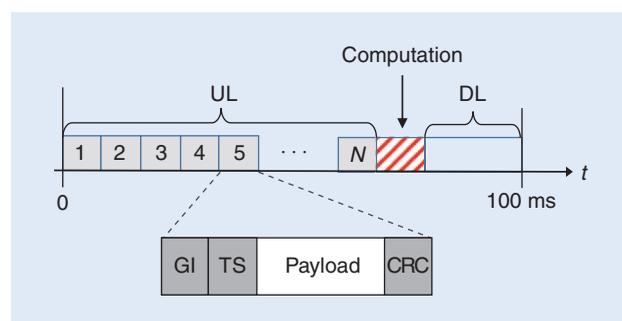


FIGURE 4. An illustration of a TDMA protocol and required communication overhead. In each time step of the RHC, vehicles send state information during the UL phase to the controller, which computes a new control signal and sends this back to the vehicles during the DL phase.

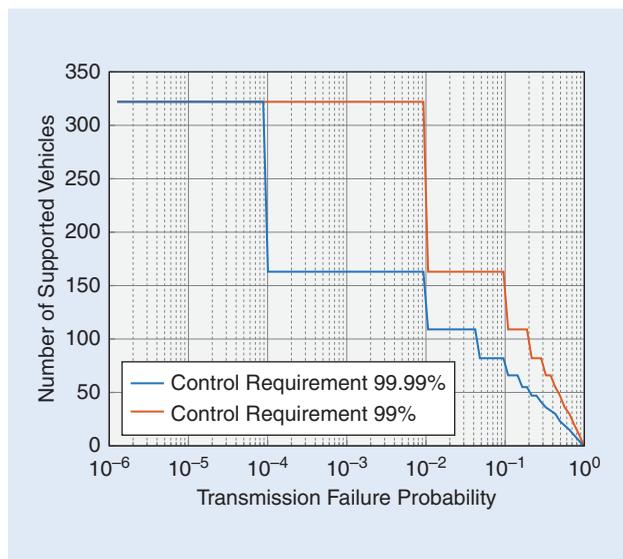


FIGURE 5. The number of vehicles that can be supported in the OCP as a function of the communication failure probability.

comparison, preamble and tail bits in 802.11p adds an overhead of approximately $45 \mu\text{s}$). The computation, assumed to scale linearly in N , is set to $10 \mu\text{s}$ per vehicle, for some value of the prediction horizon. In the downlink (DL), $200 \mu\text{s}$ data

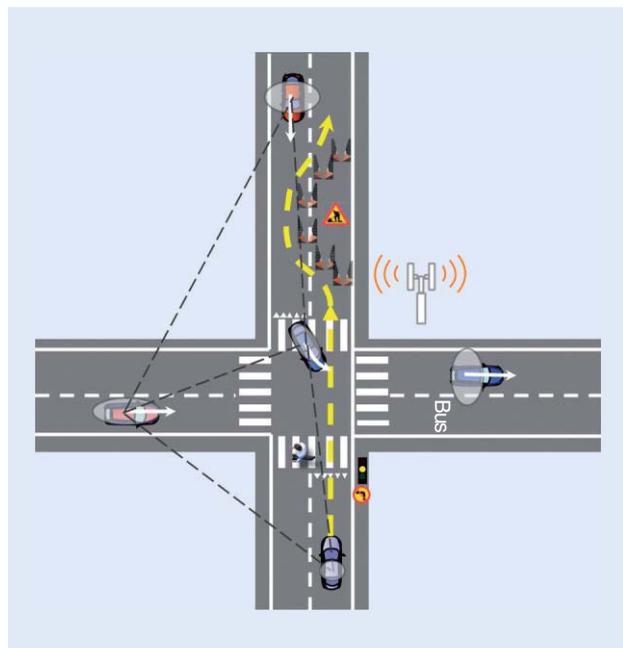


FIGURE 6. An illustration of the sensing problem to support the OCP. The host vehicle (bottom) is approaching an intersection. Information about the intersection is stored in a detailed map containing position of landmarks, geometry of lanes, traffic rules, etc. The aim of the perception subsystem is to position the host vehicle and other relevant road users as well as other obstacles (construction site) in the map such that the OCP can calculate an optimal path (shown as dashed lines). Both the host position and the state of other vehicles (pose and velocities) are described, including uncertainty measures (depicted as gray ellipses). In this example, four of the vehicles are cooperative and exchange information about their positions and current perception to each other as well as the construction site.

packets with the control signals are sent back to each of the vehicles. Since communication can never be guaranteed to succeed, we will consider that the OCP operates under a minimal requirement in terms of the fraction of information that is needed from the vehicles (say, 99% or 99.9%). The communication system is then designed to retransmit data until the requirement is met. Given this information, Figure 5 depicts the number of vehicles that can be supported for different control requirements and different communication failure probabilities. We observe that, when either the channel is very reliable or when the control requirement is loose, then more than 300 vehicles can be supported. However, this number drops quickly when the channel becomes more unreliable. Note that packet error rates in excess of 10% are not uncommon in practice. Based on this quick analysis, it is easy to see that the communication forms a bottleneck for the OCP, especially in urban scenarios where several hundred vehicles can be within communication range. To deal with the conflicting demands on the communication system in terms of latency, throughput, and node density, new V2X communication architectures were proposed in [26]. In addition, dedicated physical layer communication techniques, as well as highly optimized medium access control algorithms, will need to be developed, to complement and support these architectures. Signal processing can further relieve the burden on the communication subsystem by censoring less critical information, by tailored compression and semantic algorithms, and by assigning communication resources to those vehicles that are expected to be the most critical to the OCP. At the same time, signal processing is also expected to play an increasing role in the security of the OCP (through ultrafast authentication and verification), privacy preservation (rendering the OCP and its solutions anonymous and untraceable), and analytics (both within a vehicle and between vehicles, particularly after accidents).

Sensing and perception

To support the OCP, the sensing and perception subsystems have two main goals: 1) to estimate the host vehicle's current location (typically on a highly detailed map) and 2) to determine the position of other road users using noisy sensor observations from onboard sensors such as camera, radar, lidar, and GNSS. Both of these problems are challenging in themselves but can be alleviated by allowing information exchange from cooperating vehicles. Figure 6 depicts an illustration of the problem.

The self-localization problem, in this context, is typically solved by matching current sensor observations of the position of landmarks/features with position of sensor landmarks/features stored in a detailed map. This map is either preconstructed offline and streamed to the vehicle from the cloud, or constructed sequentially and jointly with the estimation of the vehicle's position (simultaneous localization and mapping, also known as SLAM [27]). In the case where the map is pre-constructed, the mapping and the localization problem can be separated, and only the localization

part needs to be solved online [28]. However, offline mapping is time-consuming and may need to be repeated periodically. In contrast, under SLAM, there is no need for offline mapping, rendering it less sensitive to changes in the environment. However, the SLAM problem is inherently more difficult than self-localization in a preconstructed map and thus tends to give inferior positioning accuracy. The problem of estimating the position of relevant road users, including uncertainty measures, is known as a multi-sensor and multiobject tracking problem, which is a well-studied problem within several applications. In contrast to the classical formulation, objects in an automotive setting typically give rise to multiple radar and lidar measurements, thus violating the classical point-source assumption (one measurement per object). Instead, objects such as vehicles, need to be treated as extended objects, which is less studied and typically leads to more complex algorithms. However, including multiple measurements per object also allows for a richer description of the object such as orientation and physical dimensions.

Both self-localization and estimating the position of other road users can be performed cooperatively [29]. For instance, for the latter problem, in addition to exchanging position estimates, information about the physical extension can be shared, thus greatly simplifying the inference and reducing the uncertainty in the position of the objects. However, as the sensor observations are typically not labeled, to use the measurements properly we need to be able to correctly associate them with the information coming from the other vehicles and accurately match them to the local view of the traffic situation, adjusting for delays due to data transmission and asynchronous sensor operation. For self-localization with offline mapping, the map resides in the cloud and can thus easily be shared among the cooperating vehicles. By sharing position estimates in the joint map, together with uncertainty measures, each vehicle can jointly estimate a more accurate ego-position as well as the position of all the other vehicles by fusing with the local perception from the on-board sensors [30]. This way, the self-location problem and positioning of other road users are solved simultaneously. This also leads to the possibility of quickly detecting and sharing changes in the map (e.g., the construction site in Figure 6). To increase the positioning accuracy, estimates of relative position to a selected set of high-quality landmarks can be exchanged between the vehicles and used in a similar manner. For SLAM, cooperation is also beneficial. There are two types of cooperative-SLAM (C-SLAM): centralized and distributed. In the former, the cooperating systems communicate their position estimates and current sensor observations to the cloud where a joint map is formed and shared among the systems [31]. In the distributed versions, however, this information is instead communicated to the individual vehicles, which build and keep their own map using all the information. Both of these C-SLAM methods require that the cooperating entities have a fairly homogeneous sensor setup such that landmarks seen

by one system are also detectable by the other systems. In addition, for the detailed sensors typically used for autonomous vehicles, communicating raw sensor observations is probably not feasible; thus, compression and semantic labeling is needed.

Performance of the OCP in the presence of communication and sensing impairments

To illustrate the role sensing and communication play in solving the coordination problem, we consider an intersection scenario of the type illustrated in Figure 2, where incoming vehicles periodically measure and send their state information (UL) to a centralized controller. The control is performed in a receding horizon fashion, where the controller solves a finite time OCP, and broadcasts the resulting control actions to the vehicles (DL). We simplify the OCP by modeling vehicles as points with positions $x_i(t)$, velocities $\dot{x}_i(t)$, and controls/accelerations $u_i(t) = \ddot{x}_i(t)$ along one-dimensional trajectories, aligned with the center of each road. The intersection is then modeled as an interval $[L_i, H_i]$ on each trajectory. The objective (4a) is chosen to be

$$J_i(x_i(t), u_i(t)) = Q_i \int_0^{t_f} (v_i^{\text{ref}} - \dot{x}_i(t))^2 dt + R_i \int_0^{t_f} u_i^2(t) dt, \quad (5)$$

where v_i^{ref} is a constant reference speed, t_f is a time horizon, and $Q_i > 0, R_i > 0$ are weights set by the user. The liveness constraint (4f) is stated as $x_i(t_f) \geq H_i$ for all vehicles. Finally, the problem is discretized and solved using standard optimization tools.

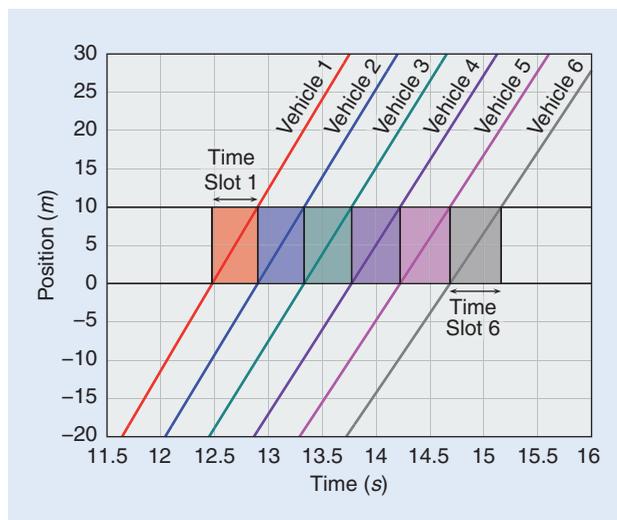


FIGURE 7. The visualization of a part of the position trajectories along each road for six coordinated vehicles under perfect communication and sensing. For each vehicle, the intersection starts at 0 [m] and ends at 10 [m]. The colored lines represent the trajectories of each vehicle. The correspondingly colored boxes visualizes the time slots during which the intersection is occupied by each vehicle. Note that collisions would occur if the time slots were to overlap. In this idealized case, the time slots are tightly packed. Hence, there is no safety margin, and the performance of the system in terms of the objective (5) is pushed to its limits.

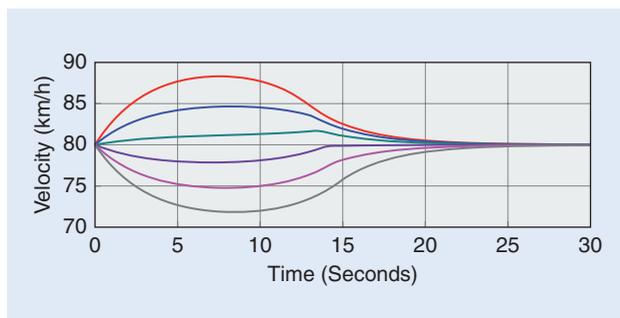


FIGURE 8. The velocity profile for the six coordinated vehicles. Note that the coordination is performed with relatively small adaptations to the velocity of the vehicles well before the intersection is reached by the first vehicle (this happens at around 12.5 seconds, as can be seen in Figure 7).

First, we study an idealized case with perfect communication, no measurement errors, and a perfect match between the dynamics used in the controller and the actual dynamics of the vehicle. We consider an instance of the problem where $N = 6$ vehicles start 300 m away from the intersection at a desired speed of 80 km/hour. Figure 7 shows the solution to the idealized coordination problem in terms of position along their trajectories. We see that the vehicles cross the intersection, one right after the other. Figure 8 shows the velocity profiles of each of the vehicles. The vehicles immediately adjust their speeds to avoid collisions in such a way that minimizes their total cost.

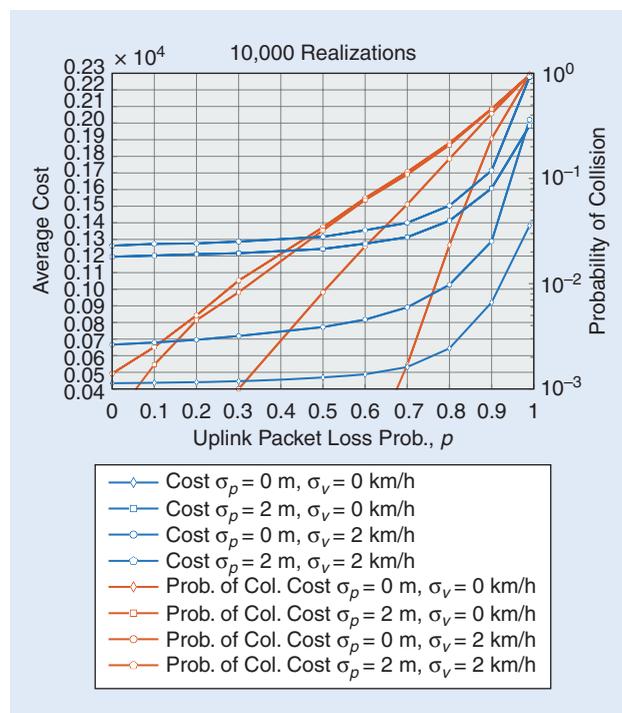


FIGURE 9. The average cost (blue) and probability of collision (red) as a function of UL packet loss probability p for different combinations of sensing uncertainty. From bottom to top: no measurement noise (diamonds), velocity measurement noise (circles), position measurement noise (squares), and both position and velocity measurement noise (pentagons).

To analyze the impact of communication and sensing errors, we reduce the problem size and focus on a two-vehicle case. Both vehicles start 80 m away from the intersection with a speed of 70 km/hour, which will lead to a collision if the central coordinator does not intervene. We introduce a slight mismatch between the true dynamics and the controller model of the dynamics, to avoid degenerate behavior in the presence of packet losses. Packet losses can occur with a probability $p \in [0, 1]$ in the UL communication, while the DL communication is assumed to be perfect. In case a packet is lost, the controller can use the latest received message from a vehicle to predict its current position and speed: the vehicle is simply assumed to obey the previously issued control command. Sensing errors are generated by adding Gaussian noise with a standard deviation σ_p and σ_v to vehicle i 's position and velocity, respectively. Performance is evaluated in terms of 1) the total cost realized by each vehicle and 2) the frequency with which collisions occur. Figure 9 shows the average cost as well as the collision probability, based on 10,000 Monte Carlo runs, as a function of the UL packet loss probability p for different combinations of sensing uncertainty. When there is no sensing uncertainty, we observe that both the probability of collision and the average cost are small, provided p is small (since the true dynamics and OCP model are well matched), but rapidly increase when $p > 0.5$. The reason for this increase is twofold. First of all, it is possible that the coordinator receives the first packet only when the cars are quite close to the intersection, thus requiring more aggressive control and possibly leading to collisions. Second, the controller may operate based on highly outdated information when successive UL transmissions fail, leading to an integration of the mismatch between true dynamics and OCP models, and thus to severe state estimation errors at the controller. When there is sensing uncertainty in either position or velocity, the controller must frequently revise its plan, leading to increases in cost, even under perfect communication. Interestingly, position uncertainty has a higher impact than velocity uncertainty, since velocity uncertainty must be accumulated over multiple failed transmissions to become significant. With increasing packet losses, there is relatively limited impact on the cost, but collisions become more and more frequent. We conclude that even when the OCP has an accurate model for each of the vehicles, sensing uncertainties quickly lead to severe problems, unless the communication system is very reliable. These problems can be avoided by formulating robust versions of the OCP, accounting for worst-case uncertainties, but at a significant cost in terms of the performance.

The road ahead

The coordination problem for cooperative, autonomous vehicles has unique properties compared to other networked control applications due to its safety-critical nature and the challenging communication environment. In this article, we cast such coordination problems as constrained optimal control problems. We have highlighted key challenges in

control, communication, and sensing, along with an overview of recent progress in each of these disciplines. In particular, the theory of distributed optimal control is deemed promising to further develop coordination algorithms that simultaneously accommodate safety and performance. Moreover, uncertainties due to pedestrians and legacy vehicles can be included in such formulations. Nevertheless, scalability and robustness are still challenging problems that deserve further study, in particular, with respect to inherently unreliable exchange of information and limited sensing of the surrounding environment.

Joint design paradigms, where control, sensing, and communications are simultaneously designed, are a promising path forward. While communication-aware control design paradigms exist (i.e., networked control design frameworks), control- and sensing-aware communications need to be further developed to facilitate the solution of the coordination problem. In particular, control-aware communication and sensing systems can establish relationships between the vehicles' mathematical models and the minimal communication and sensing resources necessary to guarantee convergence, stability, and feasibility of the coordination algorithm. Hence, algorithms could be designed so that the coordination plan and the communication or sensing resource allocation are simultaneously decided, thus inherently prioritizing the information exchange of critical vehicles (e.g., vehicles close to the intersection or vehicles that are most likely to be involved in a predicted collision). We believe that joint design of control, communication and sensing systems will pave the way for safer, more efficient, and sustainable road transportation, and that algorithmic aspects associated with signal processing implementations can help address the associated real-world challenges. Proposing and analyzing such joint design frameworks are formidable and long-term research challenges, which will require cooperation among signal processing, communication, and control communities.

Acknowledgments

This work has been supported by the Swedish Research Council under grant 2012-4038; Chalmers' Area of Advance in Transportation; SAFER; the National Metrology Institute hosted at the SP Technical Research Institute of Sweden, which in turn is partly funded by VINNOVA under the program for national metrology (project 2015-06478); COPPLAR (CampusShuttle Cooperative Perception and Planning Platform), funded under Strategic Vehicle Research and Innovation grant 2015-04849; the European Research Council under grant 258418 (COOPNET); the EU Horizon 2020 project HIGHTS (High-Precision Positioning for Cooperative ITS Applications) MG-3.5a-2014-636537; and the grant ADI4VARI02-Progetto ERC BETTER CARS-Sottomisura B.

Authors

Robert Hult (robert.hult@chalmers.se) received his B.S. degree in mechanical engineering in 2011 and M.S. degree in systems, control, and mechatronics in 2013 from Chalmers University, where he is currently pursuing his Ph.D. degree. His current research interests include distributed and cooperative control of automated vehicles and other control topics within intelligent transport systems.

Gabriel R. Campos (gabriel.rodrigues.decampos@polimi.it) received his B.S. degree in electrical engineering in 2007 from the Instituto Superior de Engenharia de Coimbra, Portugal, and from the Université de Reims-Champagne Ardenne, France. In 2009, he received the M.S. degree in electrical engineering from the same university. In 2012, he received his Ph.D. degree in automatic control from the Université de Grenoble, France. He is currently a postdoctoral research associate with the DEIB Department at Politecnico di Milano, Italy. Previously, he was with the Department of Signals

and Systems, Chalmers University of Technology, Sweden. His research interests include cooperative control and autonomous systems.

Erik Steinmetz (estein@chalmers.se) received his M.S. degree in electrical engineering from Chalmers University of Technology, Sweden, in 2009. He is currently a research and development engineer with SP Technical Research Institute of Sweden. He is also affiliated with the Department of Signals and Systems at Chalmers University of Technology, where he is working toward his Ph.D. degree. His research interests include positioning, sensor fusion, communication and controls applied within the fields of intelligent vehicles, cooperative automated driving, and active safety.

Lars Hammarstrand (lars.hammarstrand@chalmers.se) received his M.S. and Ph.D. degrees in electrical engineering from Chalmers University of Technology, Sweden, in 2004 and 2010, respectively. Between 2004 and 2011, he was with the Active Safety and Chassis Department at Volvo Car Corporation, Gothenburg, conducting research on tracking and sensor fusion methods for active safety systems. Currently, he is an assistant professor with the Signal Processing group at Chalmers University of Technology where his main research interests are in the fields of estimation, sensor fusion, and self-localization and mapping, especially with application to self-driving vehicles.

Paolo Falcone (paolo.falcone@chalmers.se) received the Laurea degree in 2003 from the University of Naples "Federico II." He received his Ph.D. degree in information technology in 2007 from the University of Sannio, Benevento, Italy. He is an associate professor with the Department of Signals and Systems of the Chalmers University of Technology, Sweden. He is an associate editor of *Control*

We believe that joint design of control, communication and sensing systems will pave the way for a safer, more efficient, and sustainable road transportation, and that algorithmic aspects associated with signal processing implementations can help address the associated real-world challenges.

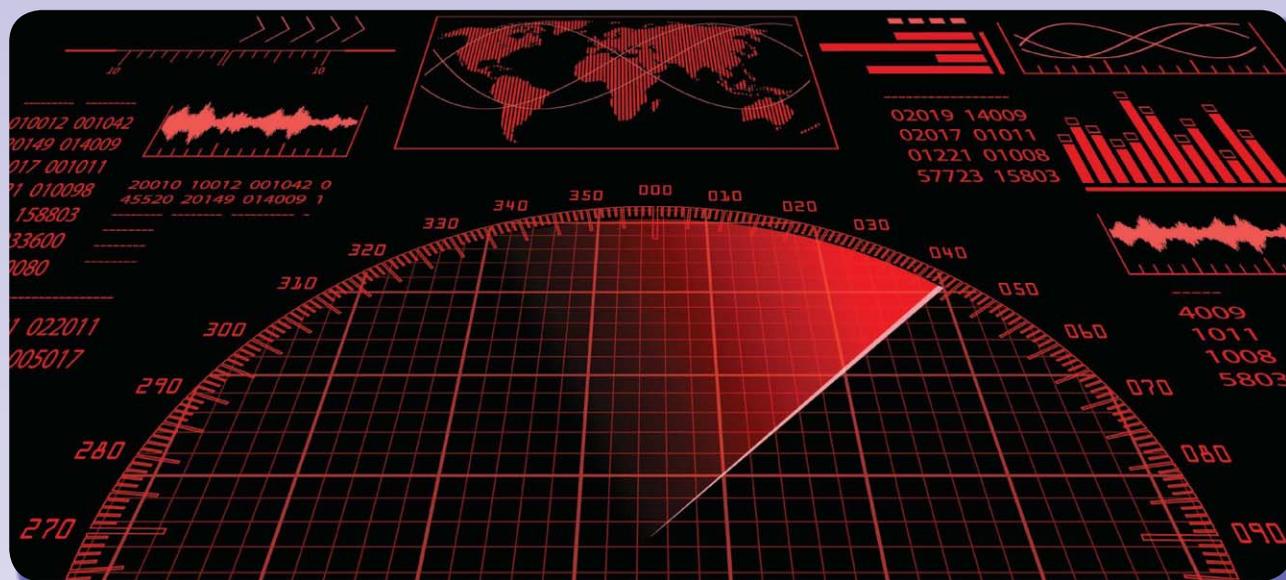
Engineering Practice and *IEEE Transactions on Intelligent Transportation Systems*. His research focuses on constrained optimal control and verification methods, applied to autonomous and semiautonomous mobile systems, cooperative driving, and intelligent vehicles. He is involved in a range of projects, in cooperation with industry, focusing on autonomous driving, cooperative driving, and vehicle dynamics control. He is a Member of the IEEE.

Henk Wymeersch (henkw@chalmers.se) received his M.S. degree in computer science in 2001 and his Ph.D. degree in electrical engineering/applied sciences in 2005, both from Ghent University, Belgium. He is a professor of communication systems with the Department of Signals and Systems, Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 to 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. He has been an associate editor of *IEEE Communication Letters* (2009–2013), *IEEE Transactions on Wireless Communications* (since 2013), and *IEEE Transactions on Communications* (since 2016). His current research interests include cooperative systems and intelligent transportation.

References

- [1] E.-H. Choi, "Crash factors in intersection-related crashes: An on-scene perspective," *National Traffic Highway Safety Association*, report no. HS-811 366, Washington, D.C., 2010.
- [2] "National motor vehicle crash causation survey," *National Traffic Highway Safety Association*, report no. DOT HS 811 059, 2008.
- [3] M. Campbell, M. Egerstedt, J. P. How, and R. M. Murray, "Autonomous driving in urban environments: Approaches, lessons and challenges," *Philos. Trans. Royal Soc. A*, vol. 368, no. 1928, pp. 4649–4672, 2010.
- [4] S. E. Shladover, "PATH at 20-history and major milestones," *IEEE Trans. Intell. Transport. Syst.*, vol. 8, no. 4, pp. 584–592, 2007.
- [5] R. Di Taranto, S. Muppisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Process. Mag.*, vol. 31, pp. 102–112, Nov. 2014.
- [6] H. Wymeersch, G. R. de Campos, P. Falcone, L. Svensson, and E. G. Ström, "Challenges for cooperative ITS: Improving road safety through the integration of wireless communications, control, and positioning," presented at the *Int. Conf. Computing, Networking and Communications*, 2015.
- [7] E. F. Camacho and C. B. Alba, *Model Predictive Control*. New York: Springer Science, 2013.
- [8] A. Colombo and D. Del Vecchio, "Efficient algorithms for collision avoidance at intersections," in *Proc. 15th ACM Int. Conf. Hybrid Systems: Computation and Control*, 2012, pp. 145–154.
- [9] S. Eichler, "Performance evaluation of the IEEE 802.11p WAVE communication standard," in *Proc. IEEE Vehicular Technology Conf.*, 2007, pp. 2199–2203.
- [10] J. B. Kenney, G. Bansal, and C. E. Rohrs, "LIMERIC: A linear message rate control algorithm for vehicular DSRC systems," in *Proc. 8th ACM Int. Workshop on Vehicular Inter-Networking*, 2011, pp. 21–30.
- [11] E. G. Ström, "On medium access and physical layer standards for cooperative intelligent transport systems in Europe," *Proc. IEEE*, vol. 99, no. 7, pp. 1183–1188, 2011.
- [12] I. Skog and P. Händel, "In-car positioning and navigation technologies survey," *IEEE Trans. Intell. Transport. Syst.*, vol. 10, no. 1, pp. 4–21, 2009.
- [13] J. B. Collins and J. Uhlmann, "Efficient gating in data association with multivariate Gaussian distributed states," *IEEE Trans. Aerospace Electron. Syst.*, vol. 28, no. 3, pp. 909–916, 1992.
- [14] K. Dresner and P. Stone, "Multiagent traffic management: A reservation-based intersection control mechanism," in *Proc. 3rd Int. Joint Conf. Autonomous Agents and Multiagent Systems*, 2004, pp. 530–537.
- [15] H. Kowshik, D. Caveney, and P. Kumar, "Provable systemwide safety in intelligent intersections," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 804–818, Mar. 2011.
- [16] R. Azimi, G. Bhatia, R. Rajkumar, and P. Mudalige, "Ballroom intersection protocol: Synchronous autonomous driving at intersections," in *Proc. IEEE 21st Int. Conf. Embedded and Real-Time Computing Systems and Applications*, 2015, pp. 167–175.
- [17] T.-C. Au, N. Shahidi, and P. Stone, "Enforcing liveness in autonomous traffic management," presented at the *25th Conf. Artificial Intelligence*, San Francisco, CA, August 2011.
- [18] T.-C. Au, M. Quinlan, and P. Stone, "Setpoint scheduling for autonomous vehicle controllers," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2012, pp. 2055–2060.
- [19] K.-D. Kim, "Collision free autonomous ground traffic: A model predictive control approach," in *Proc. ACM/IEEE Int. Conf. Cyber-Physical Systems*, 2013, pp. 51–60.
- [20] M. Ahmane, A. Abbas-Turki, F. Perronnet, J. Wu, A. E. Moudni, J. Buisson, and R. Zeo, "Modeling and controlling an isolated urban intersection based on cooperative vehicles," *Transport. Res. Part C*, vol. 28, no. 0, pp. 44–62, Mar. 2013.
- [21] A. Colombo and D. Del Vecchio, "Least restrictive supervisors for intersection collision avoidance: A scheduling approach," *IEEE Trans. Automat. Contr.*, vol. PP, no. 99, pp. 1–1, 2014.
- [22] G. R. Campos, F. D. Rossa, and A. Colombo, "Optimal and least restrictive supervisory control: safety verification methods for human-driven vehicles at traffic intersections," in *Proc. IEEE Conf. Decision and Control*, 2015, pp. 1707–1712.
- [23] R. Hult, G. R. Campos, P. Falcone, and H. Wymeersch, "An approximate solution to the optimal coordination problem for autonomous vehicles at intersections," presented at the *American Control Conf.*, Chicago, IL, 2015.
- [24] G. R. Campos, P. Falcone, H. Wymeersch, R. Hult, and J. Sjöberg, "A receding horizon control strategy for cooperative conflict resolution at traffic intersections," in *Proc. IEEE Conf. Decision and Control*, 2014, pp. 2932–2937.
- [25] 5G Infrastructure Association, "Automotive vertical sector," White Paper 5G-PPP, 2015.
- [26] H. Laurens, A. Festag, I. Llatser, L. Altomare, F. Visintainer, and A. Kovacs, "Enhancements of V2X communication in support of cooperative autonomous driving," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 64–70, 2016.
- [27] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robot. Automat. Mag.*, vol. 13, no. 2, pp. 99–110, 2006.
- [28] J. Ziegler, P. Bender, M. Schreiber, H. Lategahn, T. Strauss, C. Stiller, T. Dang, U. Franke, N. Appenrodt, C. G. Keller, et al., "Making Bertha drive an autonomous journey on a historic route," *IEEE Intell. Transport. Syst. Mag.*, vol. 6, no. 2, pp. 8–20, 2014.
- [29] S.-W. Kim, B. Qin, Z. J. Chong, X. Shen, W. Liu, M. Ang, E. Frazzoli, and D. Rus, "Multivehicle cooperative driving using cooperative perception: Design and experimental validation," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, pp. 663–680, Apr. 2015.
- [30] A. Howard, M. J. Matarić, and G. S. Sukhatme, "Putting the 'i' in 'team': an ego-centric approach to cooperative localization," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2003, vol. 1, pp. 868–874.
- [31] L. Riazuelo, J. Civera, and J. Montiel, "C2TAM: A cloud framework for cooperative tracking and mapping," *Robot. Auton. Syst.*, vol. 62, no. 4, pp. 401–413, 2014.

Lifan Zhao, Lu Wang,
Lei Yang, Abdelhak M. Zoubir,
and Guoan Bi



©ISTOCKPHOTO.COM/MAKHACH_M

The Race to Improve Radar Imagery

*An overview of recent progress
in statistical sparsity-based techniques*

The exploitation of sparsity has significantly advanced the field of radar imaging over the last few decades, leading to substantial improvements in the resolution and quality of the processed images. More recent developments in compressed sensing (CS) suggest that statistical sparsity can lead to further performance benefits by imposing sparsity as a statistical prior on the considered signal. In this article, a comprehensive survey is made of recent progress on statistical sparsity-based techniques for various radar imagery applications.

Introduction

The capability of operating in all-day, all-night, and all-weather conditions has allowed high-resolution radar imagery to play an important role in both civilian

Digital Object Identifier 10.1109/MSP.2016.2573847
Date of publication: 4 November 2016

and military remote-sensing applications. The principle of radar imagery is to utilize wideband transmitting signals and aperture synthesis to obtain the desirable slant-range and cross-range resolution, respectively [1]. However, the processing of wideband signals requires high-speed analog-to-digital converters at the receiver. Also, a long coherent processing interval (CPI) is required to obtain high cross-range resolution, which would inevitably introduce undesirable higher-order Doppler effects. To overcome the aforementioned limitations, sparsity-based techniques [2] have become increasingly important in radar imagery. Successful applications in recent years can be found in [3]–[7] and the references therein. In addition to achieving high-resolution radar images with limited data, sparsity-based techniques can offer additional advantages over conventional spectral analysis-based techniques, such as denoising and sidelobe suppression [3], [4].

In a nutshell, the theory of CS states that a high-dimensional signal can be accurately and robustly recovered from its low-dimensional projections if the signal is sparse or can be sparsely represented [8]. The success of CS techniques is due to the proper exploitation and utilization of sparsity. One obvious way to solve the problem is to enforce sparsity, i.e., to minimize the number of nonzero entries in the signal. However, solving an ℓ_0 -minimization problem, i.e., minimizing the number of its nonzero elements, is generally NP-hard, which requires an exhaustive search with intractable computational cost. Various algorithms have been developed in CS to obtain an approximate solution to this problem. In many radar imagery applications, CS theory can be conveniently applied due to the natural presence of sparsity. In particular, the target scene of interest is often parsimonious or can be parsimoniously represented by an appropriately chosen linear basis [3], [4]. Despite the diverse applications of sparsity-based techniques in radar imaging, these techniques can be summarized methodologically into the following three main categories.

1) The first category is based on the use of greedy algorithms that lead to a sparse solution in an iterative and greedy manner. Often, these algorithms require approximation of the signal's support and amplitude by refining the sparse signal estimation based on evaluation of the difference between the recovered signal and the observations. In [9], for example, greedy algorithm-based CS is used to operate wide-swath modes for radar imaging with reduced measurements. In [10], the authors have investigated the uniform and nonuniform target imaging problem with greedy approaches. Although greedy algorithms can be conveniently implemented and have desirable guarantees under some conditions, they generally result in a local optimum, which does not coincide with the sparsest solution.

In addition to achieving high-resolution radar images with limited data, sparsity-based techniques can offer additional advantages over conventional spectral analysis-based techniques.

2) Another major category of algorithms is a generalization of an ℓ_1 -regularized optimization method, which can be considered as the tightest convex relaxation of the ℓ_0 -minimization problem. Basis pursuit and basis pursuit denoising [11] are convex formulations to recover sparse signals in noiseless and noisy environments, respectively.

Various approaches have been proposed to solve the ℓ_1 -minimization problem, such as linear programming and the interior point method. Since the objective functions are convex, it is guaranteed that these algorithms lead to a global optimum. However, the solution does not necessarily coincide with the maximally sparse solution, except that the problem satisfies some specific conditions. In [12] and [13], the ℓ_1 -regularized optimization is applied for radar image reconstruction.

In [14] and [15], phase error correction and imaging are formulated within a convex optimization framework. The synthetic aperture radar (SAR) ground moving target imaging (GMTIm) problem is also formulated in a sparsity-driven manner and solved by convex optimization techniques [5], [16]. Remarkably, empirical results suggest that the ℓ_1 -regularized optimization provides substantial improvements over greedy approaches for some radar imaging problems [3]. However, the regularization parameter should be carefully tuned to obtain a desirable performance. But finding an optimal selection rule is still an open problem.

3) The third category of statistical sparsity-based methods [17] provides remarkable statistical advantages over conventional ones. Under certain conditions, the resulting algorithm guarantees that its global optimum coincides with the maximally sparse solution and smoothes the shallow local minimum [18]. Theoretical and empirical results show that enhanced performance can be achieved from Bayesian inference over conventional ℓ_1 -regularized optimization [17], [19]–[21]. This technique is particularly useful in overcoming some limitations of the previous two categories. The advantages for radar imagery applications include incorporation of flexible prior knowledge, estimation accuracy improvement, as well as estimation of error bars.

The main objective in radar imagery applications is to properly utilize one of these methods to obtain enhanced imaging performance, which is particularly useful in situations where the number of measurements is limited and the signal-to-noise ratio (SNR) is low. Our objective in this review article is to present the motivation and ways of utilizing statistical sparsity-based radar imaging techniques.

Recent overview articles [3]–[5] focus on sparsity-based radar imagery techniques from either a greedy or regularized perspective. These articles demonstrate the importance and effectiveness of sparsity in radar imagery applications. With the recent development of sparse Bayesian methods [17], [18], statistical sparsity-based techniques have become

a more promising research area for radar imagery applications. Compared to the deterministic sparsity-inducing framework, statistical sparsity-based techniques provide new opportunities to significantly improve the performance of radar imagery. This is mostly due to the capability of avoiding regularization parameter tuning, providing desirable statistical information, and allowing flexible modeling. These advantages are due respectively to the inherent advantage of the statistical framework, the desirable statistical information obtained from the estimation of the full posterior distribution, and the inherent flexibility of statistical sparsity-based modeling. To benefit from these advantages, sophisticated design is required. This article is a companion of recent tutorial articles on sparsity-based radar imagery [3]–[5], but with particular emphasis on sparsity-based radar imagery from a statistical perspective, which is missing in the recent literature.

We show how this design is to be performed and demonstrate how the various radar imagery problems can be formulated within a sparse Bayesian framework. We illustrate in detail why the statistical formulation greatly enhances the radar imaging performance in various practical problems. The introduced framework has much promise for future radar imaging systems, as it provides substantial improvements as well as new opportunities. The notations used in this article are summarized in Table 1.

Statistical sparsity formulation of radar imagery

We begin our treatment by reviewing the fundamentals of statistical sparsity-inducing models in radar imagery. We formulate the statistical sparsity-based framework and highlight from where the advantages arise along with the limitations of statistical sparsity-based methods.

Data modeling

In high-resolution radar imagery, the scattering response of a target of interest is often expressed as a sum of scatterers' responses. Without loss of generality, assuming that the radar emits successive pulses with time interval T_r and that there exist K strong scatterers in an imaging scene, the received radar signal can be given by

$$s_r(t, t_n) = \sum_{k=1}^K \sigma_k \cdot s\left(t - \frac{2R_k(t_n)}{c}\right) + n(t, t_n), \quad (1)$$

where σ_k represents the amplitude of the k th scatterer, c is the speed of light, and $R_k(t_n)$ represents the range from the radar to scatterer k in slow time t_n . The fast time and slow time of pulse n are denoted by $0 \leq t \leq T_r$ and $t_n = nT_r$, respectively. To achieve high-range resolution, the emitted pulses $s(t, t_n)$ are often chosen as linear frequency modulated (LFM) signals, but other waveforms such as sparse stepped-frequency signals [22], sparse probing-frequency signals [23], and adaptively optimized signals [24] can also be used for the purpose of improved imaging performance. To achieve high-cross-range resolution, a large aspect angle between the radar and the target is required

Table 1. Notation summary.

Notations	
$\mathbb{C}^{M \times N}$	The set of a complex $M \times N$ matrix
a, α	Scalar and vector
$A, A_i, A_{\cdot j}, A_{n,m}$	Matrix, the i th row, the j th column, and the (n, m) th entry of a matrix
$(\cdot)^T$ and $(\cdot)^H$	Matrix or vector transpose and conjugate transpose
A^{-1}	Matrix inverse
$ \cdot $	The absolute value of a scalar
$\ \cdot\ _p$	The ℓ_p norm of a vector
$\ \cdot\ _F$	The Frobenius norm of a matrix
$\exp(\cdot)$	The exponential function
$CN(\mu, \Sigma)$	The complex Gaussian distribution with mean μ and covariance matrix Σ
Beta(a, b)	The beta distribution with parameters a and b
$\Gamma(a, b)$	The gamma distribution with parameters a and b

during the CPI. Note that sparsity-based methods go beyond the convention in the sense that high-range resolution can be obtained with less bandwidth, and high-cross-range resolution can be obtained with a reduced CPI.

After preprocessing and arranging the cross-range measurements column-wise, a linear model is obtained [1] as

$$\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{B} + \mathbf{N}, \quad (2)$$

where $\mathbf{Y} \in \mathbb{C}^{M \times N}$ is the preprocessed radar echoes, $\mathbf{X} \in \mathbb{C}^{M \times N}$ is the unknown scattering coefficient, and $\mathbf{A} \in \mathbb{C}^{M \times M}$ and $\mathbf{B} \in \mathbb{C}^{N \times N}$ are the measurement matrices constructed from (1) for cross range and range, respectively. In general, \mathbf{A} and \mathbf{B} are Fourier matrices. There exist, however, other ways to construct the dictionary other than simply employing the Fourier matrix. These include the frame-based matrix [25] and the matched filter-based matrix [26]. Note that the model in (2) does not yet include the case of an undersampled or incomplete measurement of \mathbf{Y} . Modifications to capture these are straightforward. A more detailed discussion on this issue is presented in the ‘‘Superresolution Radar Imagery’’ section.

In (2), \mathbf{N} is assumed to be independently circularly symmetric complex Gaussian distributed, so that the received signal \mathbf{Y} follows a complex Gaussian distribution with a likelihood function given by

$$p(\mathbf{Y} | \mathbf{X}, \alpha_0) = \prod_{i=1}^M \prod_{j=1}^N CN(\mathbf{Y}_{ij} | \mathbf{A}_i \mathbf{X} \mathbf{B}_{\cdot j}, \alpha_0^{-1} \mathbf{I}). \quad (3)$$

The α_0 is the noise precision or the reciprocal of the variance, which is assumed to be random. For the sake of

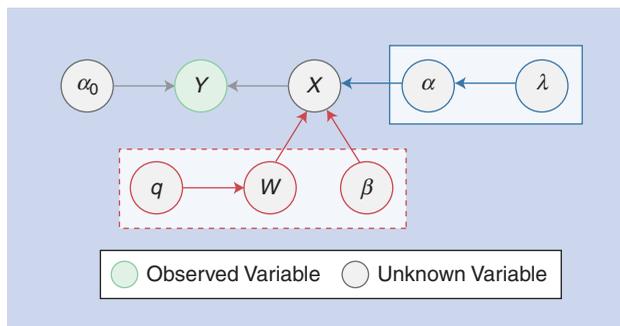


FIGURE 1. A graphical representation of the scaled Gaussian mixture model (in the solid-line frame) and the spike-and-slab model (in the broken-line frame).

convenient inference, we model the noise precision with a gamma distribution,

$$p(\alpha_0 | v_1, v_2) = \Gamma(\alpha_0 | v_1, v_2),$$

where v_1 and v_2 are hyperparameters, often set as trivial values to impose a noninformative prior on the noise level. The rationale of the gamma prior for the noise precision is due to the likelihood and prior conjugacy [27]. By this modeling, noise level estimation can be naturally incorporated into the signal estimation task.

To obtain a solution for the linear equation in (2), sparsity is often imposed to constrain the solution space. Unlike convex-based approaches constraining the solution space by regularization, statistical approaches seek posterior estimation to robustly estimate the signal \mathbf{X} by properly imposing sparse priors on the signals to be estimated.

Probabilistic sparse modeling

In sparsity-based radar imaging applications, we impose sparsity on the target scattering coefficients \mathbf{X} , i.e., some strong scatterers in the target scene. In other words, most of the coefficients in \mathbf{X} are zeros or nearly zeros due to the fact that the scatterers are sparsely distributed in the imaging scene. We show statistical ways of imposing sparsity on various radar imaging applications. To impose sparse priors and allow inference, the models need to be carefully designed. Essentially, the model should have two key characteristics of sparsity and model conjugacy. That is, the constructed model should not only induce sparsity but also allow convenient inference of the unknown parameters.

Toward this end, a hierarchical model is often utilized instead of a single-layer model. In what follows, two main classical models, the so-called scaled Gaussian mixture model and the spike-and-slab model, are briefly reviewed and shown in the solid line frame and the broken line frame in Figure 1, respectively. In the scaled Gaussian mixture model, the signal is assumed to follow a Gaussian distribution, and its variance is assumed to follow a particular distribution to induce sparsity and convenient inference. In contrast, in the spike-and-slab model, the signal is assumed to be a dot product of the support and the amplitude coefficients, where the support coefficient is hierarchically modeled for the sake of achieving sparsity.

The scaled Gaussian mixture model

Although there exist various models within this class, we choose to briefly review the three-stage hierarchical model [28] as an example. As shown in Figure 1, the sparse signal \mathbf{X} is hierarchically constructed.

- In the first stage, the sparse signal \mathbf{X} is modeled with a complex Gaussian distribution,

$$p(\mathbf{X} | \alpha) = \prod_{i=1}^M \prod_{j=1}^N \mathcal{CN}(X_{ij} | 0, \alpha_{ij}). \tag{4}$$

- In the second stage, we find the distribution of the variance α of the scattering coefficient \mathbf{X} . It is assumed to follow an independent gamma distribution since it is the conjugate prior of a Gaussian distribution, thus making inference tractable [27]:

$$p(\alpha_{ij} | \lambda) = \prod_{i=1}^M \prod_{j=1}^N \Gamma(\alpha_{ij} | \eta, \lambda). \tag{5}$$

Combining (4) and (5), it can be shown that the marginalized distribution of \mathbf{X}_{ij} is a complex Laplace distribution, which is known to be a suitable model for sparsity [2].

- In the third stage, we choose the gamma distribution

$$p(\lambda | v_3, v_4) = \Gamma(\lambda | v_3, v_4) \tag{6}$$

to infer that λ that controls sparsity of the prior during the learning from the data.

There are many variants of the hierarchical model, which can all be summarized as scaled Gaussian mixture models (Table 2).

Table 2. A summary of the scaled gaussian mixture for sparse modeling.

Number of Stages	Model Specification	Marginalized Distribution
Two stages	Gaussian–Jeffery [29]	No closed-form representation
	Gaussian–gamma [30]	Laplace distribution
	Gaussian–inverse gamma [17]	Student’s t distribution
	Gaussian–exponential [31]	Double exponential distribution
	Gaussian–half Cauchy [32]	No closed-form representation
Three stages	Gaussian–gamma–gamma [33]	Laplace distribution

The spike-and-slab model

In contrast, the spike-and-slab model is another popular approach to model the support and amplitude of sparse signals. In particular, the prior of the sparse signal is expressed as a mixture of a point probability distribution (spike) and a Gaussian distribution (slab):

$$p(\mathbf{X}|\mathbf{W},\boldsymbol{\beta}) = \prod_{i=1}^M \prod_{j=1}^N [(1 - \mathbf{W}_{ij}) \cdot \delta(\mathbf{X}_{ij}) + \mathbf{W}_{ij} \cdot \mathcal{CN}(\mathbf{X}_{ij}|0, \boldsymbol{\beta}_{ij})], \quad (7)$$

where δ is the point probability mass centered at 0, \mathbf{W} is the support coefficient, and $\boldsymbol{\beta}$ controls the amplitude of coefficient \mathbf{X} . The support coefficient \mathbf{W} determines the sparsity profile of the signal, and the amplitude coefficient $\boldsymbol{\beta}$ controls the amplitude of the signal.

■ In the first stage, the support coefficient \mathbf{W} is modeled by a Bernoulli distribution,

$$p(\mathbf{W} | \mathbf{q}) = \prod_{i=1}^M \prod_{j=1}^N q_{ij}^{\mathbf{W}_{ij}} (1 - q_{ij})^{(1 - \mathbf{W}_{ij})},$$

where q_{ij} is the probability of $\mathbf{W}_{ij} = 1$ and $1 - q_{ij}$ is the probability of $\mathbf{W}_{ij} = 0$. Note that in the spike-and-slab model, sparsity can be obtained by imposing a prior, such that most of the entries in \mathbf{W} are zeros. In other words, the probability of an entry being zero should be larger than the probability of an entry being nonzero. To conveniently estimate parameter q_{ij} , a beta distribution is imposed on q_{ij} :

$$p(q_{ij}) = \text{Beta}(q_{ij} | e, f),$$

where e and f are hyperparameters to be set as trivial values. The reason for selecting the beta distribution is for its conjugacy to the Bernoulli distribution [27]. In some applications, their values can be specified by some prior information.

■ In the second stage, the coefficient $\boldsymbol{\beta}$ that controls the amplitude follows a complex Gaussian distribution,

$$p(\boldsymbol{\beta}) = \prod_{i=1}^M \prod_{j=1}^N \mathcal{CN}(\boldsymbol{\beta}_{ij} | 0, \nu_0).$$

To allow for its inference, ν_0 can be modeled by a gamma distribution. In particular, unlike in the Gaussian mixture model, a single variance parameter ν_0 is assumed. This is because sparsity has already been captured by the Bernoulli-beta model in the first stage. Based on the two-stage model, this spike-and-slab model can therefore impose sparsity on the signal.

In summary, these two classical models are frequently utilized and modified in statistical sparsity-based radar imagery. We demonstrate ways to fully utilize these two models for desirable improvements in specific applications in the sections “Superresolution Radar Imagery,” “Enhanced Target Imagery by Exploiting Structured Sparsity,” “Statistical Sparsity-Based

Autofocus Techniques in Radar Imagery,” and “Statistical Sparsity-Based SAR GMTIm.”

Connections with convex optimization

The conventional sparsity regularization-based methods can be interpreted from a Bayesian perspective. The Laplace distribution, which is a popular choice as a sparse prior [2], is imposed on the signal \mathbf{X} . Then, the maximum a posteriori (MAP) technique is utilized for parameter estimation. It can be shown that the ℓ_1 -regularized method corresponds to the MAP estimation with a Gaussian likelihood and a Laplace prior [18]:

$$\begin{aligned} \hat{\mathbf{X}} &= \arg \max_{\mathbf{X}} \frac{p(\mathbf{Y} | \mathbf{X})p(\mathbf{X})}{p(\mathbf{Y})} \\ &= \arg \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 + \lambda \sum_{i=1}^N \|\mathbf{X}_{\cdot i}\|_1. \end{aligned} \quad (8)$$

This strategy, however, can provide only point estimation of \mathbf{X} , without any higher-order statistical information. In contrast, the full posterior, including higher-order statistical information, can be obtained in a statistical sparse framework due to the hierarchical model. This main difference allows statistical sparsity-based methods to perform better in many tasks. For example, incorporating a more sophisticated prior on the signals provides flexibility to the hierarchical Bayesian model. This is not the case in the regularized framework.

In summary, the statistical sparsity-based models, such as the scaled Gaussian mixture model and the spike-and-slab model, avoid the laborious tuning of the regularization parameter λ in (8). These methods are also flexible in view of choosing different priors and provide higher-order statistical information in the posterior distribution (due to the Bayesian inference). In the sections “Superresolution Radar Imagery,” “Enhanced Target Imagery by Exploiting Structured Sparsity,” “Statistical Sparsity-Based Autofocus Techniques in Radar Imagery,” and “Statistical Sparsity-Based SAR GMTIm” we will show ways to properly manipulate the statistical sparse model so as to make use of these desirable properties.

Bayesian inference

Based on the formulated probabilistic model, the remaining task is to infer the parameters. We recall the graphical representation in Figure 1, where all the unknown variables are required to be estimated. Based on the likelihood of \mathbf{Y} and a scaled Gaussian or a spike-and-slab prior, the posterior distribution can be expressed according to the Bayesian theorem

$$p(\boldsymbol{\Theta} | \mathbf{Y}) = \frac{p(\boldsymbol{\Theta})p(\mathbf{Y} | \boldsymbol{\Theta})}{p(\mathbf{Y})}, \quad (9)$$

where $\boldsymbol{\Theta}$ is a set of all the parameters to be estimated, i.e., the unknown parameters in Figure 1. However, one major difficulty is that the marginalized distribution cannot be explicitly calculated due to the intractability of the integral

$$p(\mathbf{Y}) = \int_{\Theta} p(\Theta) p(\mathbf{Y} | \Theta) d\Theta,$$

and thus the posterior distribution in (9) is not attainable. Although MAP estimation can be obtained from this model, the full posterior is more desirable, as it provides a more accurate description of the estimated parameter.

Because the exact inference is not attainable from this model, two classical inference techniques known as the *Markov chain Monte Carlo (MCMC) method* and the *variational Bayesian (VB) method* are often used to approximate the posterior from sampling and optimization, respectively. In this way, the approximated posterior can be obtained at the cost of increased computational complexity compared with other sparse signal recovery methods. The MCMC method is accurate when the number of samples becomes large, while the VB method provides a desirable approximation with a reasonable computational complexity.

The MCMC method

This method approximates the posterior by sampling. MCMC is a strategy for generating samples, while the equilibrium distribution of the Markov chain is the same as the desired probability distribution [34]. The most widely used MCMC algorithms are the Metropolis–Hastings and the Gibbs sampling algorithms [27]. Under the assumption that all the conditional distributions are available, Gibbs sampling is easily applicable. In fact, Gibbs sampling can be considered as a special case of the Metropolis–Hastings algorithm if the conditional distributions are provided [34].

Since the conditional distribution is available in our graphical model, as shown in Figure 1, we will briefly review Gibbs sampling. In this approach, sequential sampling of the conditional distribution, expressed as

$$\Theta_i \sim p(\Theta_i | \Theta_{k \neq i}, \mathbf{Y}), \quad (10)$$

is performed. Therefore, the algorithm iterates until the desirable posterior is obtained. A more detailed description of the algorithm can be found in [34] and the references therein.

The VB method

The main idea of this method is to approximate the true posterior by a factorizable form

$$q(\Theta) = \prod_{i=1}^k q(\Theta_i). \quad (11)$$

This is known as the *mean-field assumption* [27]. The objective is to find a factorizable $q(\Theta)$ that is as close as the true posterior $p(\Theta | \mathbf{Y})$. The closeness of the estimated posterior to the true one in the VB method is measured by the Kullback–Leibler (KL) divergence, and thus the optimal approximated posterior is obtained by minimizing the following KL divergence:

$$q^*(\Theta) = \operatorname{argmin}_{q(\Theta)} \int q(\Theta) \ln \frac{q(\Theta)}{p(\Theta | \mathbf{Y})} d\Theta. \quad (12)$$

Based on (11) and (12), it can be shown that the approximated posterior for each of the variables can be calculated as [27], [35]

$$q^*(\Theta_i) = \exp\{\langle \ln p(\Theta, \mathbf{Y}) \rangle_{q(\Theta_{k \neq i})}\}, \quad (13)$$

where $\langle \cdot \rangle_{q(\cdot)}$ represents expectation with respect to the probability density function $q(\cdot)$.

In the MCMC-based methods, sampling is required for each step during iterations. In contrast, the VB-based methods require matrix inversion in each step during iterations. Notably, sampling and matrix inversion would generally induce high computational complexity for MCMC and VB, respectively. Due to these reasons, statistical sparsity-based methods generally cost more computations than the nonstatistical approaches. The MCMC method can achieve better estimation accuracy than the VB method, but at a higher computational expense. In a practical application, one should choose the method according to its computational cost tolerance.

In summary, the key advantages of statistical sparsity-based methods are:

- They avoid regularization parameter tuning. Parameter tuning is not required in statistical sparsity-based methods, which will be demonstrated in all the applications reviewed in this article.
- They provide full posterior. With this capability, desirable improvements can be achieved by properly manipulating the statistical model, particularly as shown in the sections “Statistical Sparsity-Based Autofocus Techniques in Radar Imagery” and “Statistical Sparsity-Based SAR GMTIm.”
- They offer flexible modeling. Since the model is constructed probabilistically, encoding the prior can be carried out in a rather flexible way, which is demonstrated in the section “Enhanced Target Imagery by Exploiting Structured Sparsity.”

Despite their remarkable advantages, the key limitations of the statistical sparsity-based methods lie primarily in the following facts.

- They have high computational complexity. The generally required computational cost of statistical sparsity-based methods is higher than that required by greedy or regularized methods.
- They require sparsity assumption. The success of almost all sparsity-based methods depends on the existence of sparsity or compressibility. If the radar target scene does not exhibit sparsity, modifications should be made to allow a sparse representation [3], [5].

Superresolution radar imagery

In conventional Fourier-based radar imagery, the resolutions in cross range and slant range are bounded by the Rayleigh limit, which can be overcome by superresolution techniques. In general, superresolution radar imaging can be well formulated as an inverse problem, where the scattering field is required to be inversely estimated from the received radar

echoes. This problem is ill posed, since it requires estimation of high-dimensional signals from low-dimensional observations. Recent advances in superresolution radar imagery consider this problem from a sparsity perspective and obtain its solution by either greedy or regularized methods. A review of state-of-the-art works can be found in [3] and [5]. Although empirical results have demonstrated the superior performance of sparsity-based algorithms over conventional spectral analysis-based methods, their success often requires careful selection of key parameters.

To alleviate this drawback, superresolution imagery techniques have been developed more recently in a statistical sparsity-based framework. Herein, the radar returns and the target scattering coefficient are modeled probabilistically, and the formulated models have the advantage of being free of laborious parameter tuning. Assuming that preprocessing procedures have been carried out, the super-resolution radar imaging problem can be explicitly formulated to be underdetermined:

$$\mathbf{Y} = \Phi_1 \mathbf{A} \mathbf{X} \mathbf{B} \Phi_2 + \mathbf{N}, \quad (14)$$

where $\mathbf{Y} \in \mathbb{C}^{P \times Q}$ is the preprocessed data, $\mathbf{X} \in \mathbb{C}^{M \times N}$ is the unknown sparse scattering coefficient, $\mathbf{N} \in \mathbb{C}^{P \times Q}$ is the noise, and $\Phi_1 \in \mathbb{C}^{P \times M}$ and $\Phi_2 \in \mathbb{C}^{N \times Q}$ represent the cross-range and the slant-range undersampling matrices, respectively. The data arrangement is shown in Figure 2, with $P \leq M$ and $Q \leq N$, where each column represents the accumulated echoes from each range cell. An important message herein is that the matrices $\Phi_1 \mathbf{A}$ and $\mathbf{B} \Phi_2$ should be carefully designed to achieve desirable properties, such as a low mutual coherence or a certain restricted isometry property [2]. To achieve these goals, the emitting signal waveform or the undersampling patterns should be appropriately designed. Examples can be found in [13] and [36]. More specifically, the so-called Alltop sequences were utilized in [36] as emitting signals, which was proved to practically achieve the lower bound of maximum mutual coherence. In contrast, the authors in [13] investigated different undersampling patterns to obtain low maximal mutual coherence and achieve radar data compression.

■ In an attempt to obtaining super cross-range resolution for radar imaging, a scheme was proposed in [37]. By setting the undersampling matrix Φ_2 to identity and carrying out proper preprocessing, a degenerated model of (14) has been proposed:

$$\mathbf{Y} = \Phi_1 \mathbf{A} \mathbf{X} + \mathbf{N},$$

where $\mathbf{Y} \in \mathbb{C}^{P \times N}$ is the cross-range undersampled data. The matrix $\Phi_1 \mathbf{A}$ is constructed as a partial Fourier matrix, whose property was investigated empirically in [38].

■ Similarly, in an attempt to obtain super range resolution, probing frequency-based signals have been proposed to obtain undersampled data in range dimension. The mathematical model can be expressed as

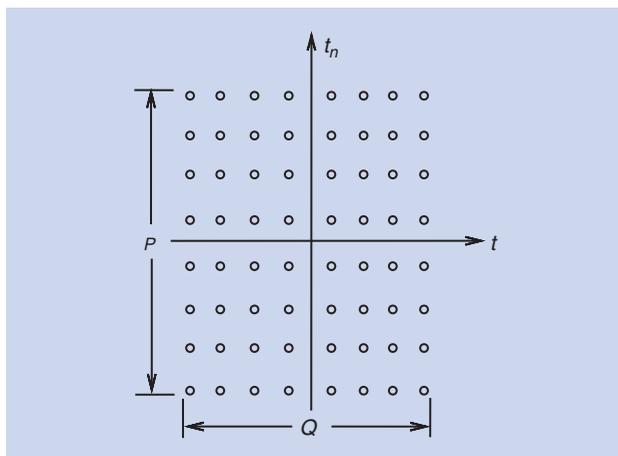


FIGURE 2. The obtained data arrangement.

$$\mathbf{Y} = \mathbf{X} \mathbf{B} \Phi_2 + \mathbf{N},$$

where $\mathbf{Y} \in \mathbb{C}^{M \times Q}$ is the range undersampled data. In particular, the maximum coherence condition of $\mathbf{B} \Phi_2$ was shown in [39]. Apart from the advantage of data compression, this strategy can greatly simplify the radar system design by avoiding emitting wideband signals.

■ In the scenario of obtaining both super cross-range and range resolution, the undersampling matrices for cross-range and range are Φ_1 and Φ_2 , respectively. The authors in [40] proposed a random sampling scheme for both slant-range and cross-range dimensions, where random selection is carried out in each dimension. This scheme is particularly useful for reducing data storage in various radar imagery applications.

To impose sparsity on the target scene, the works reviewed above employed a scaled Gaussian mixture model to induce sparsity as well as to avoid parameter tuning. Empirical experimental results found in [37], [40], and [41] demonstrate that the statistical sparsity model obtains cleaner images without parameter tuning, as compared to other sparse regularized-based methods [40]. Additionally, the statistical sparsity-based methods have been empirically shown to be less sensitive to noise and clutter in radar imaging.

In [37], [40], and [41], multiplicative speckle noise in radar imagery has not been considered. In the presence of speckle noise, the performance of sparsity-based methods would be compromised. For this particular problem, the authors in [3] and [13] proposed to use the ℓ_1 norm of the scene's gradient in addition to the ℓ_1 norm of the scene to obtain a despeckled radar image. To the best of our knowledge, despeckling in radar imaging has not yet been specifically considered under the statistical sparsity-based framework. However, similar ideas of exploiting gradient sparsity along with sparsity have been proposed in a statistical framework, an example of which can be found in [42]. An immediate advantage is the fact that the formulated statistical sparsity-based method is free of regularized parameter tuning process, and good performance is attainable.

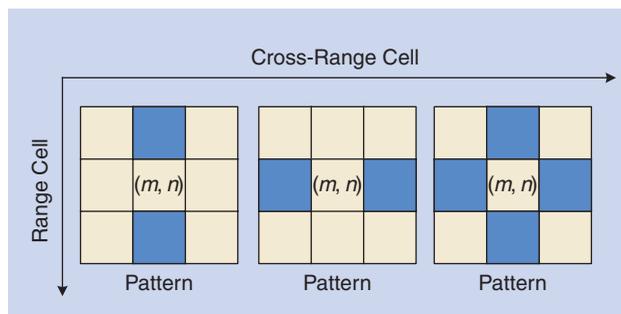


FIGURE 3. The first-order continuity patterns.

In summary, there are two important points to be noted here. First, the success of statistical sparsity-based methods depends on the proper selection of sampling patterns, which is also the case for conventional sparsity-based methods. In fact, statistical sparsity-based methods are less sensitive to highly correlated measurement matrices induced by some sampling patterns than sparsity regularized methods. A detailed analysis can be found in [43] and the references therein. Second, a tradeoff between design convenience and performance is to be properly balanced, depending on the performance requirements for superresolution in the problem at hand.

Enhanced target imagery by exploiting structured sparsity

In the “Superresolution Radar Imagery” section, we demonstrate that statistical sparsity could lead to improvements over deterministic sparsity in superresolution radar imagery. To carry out sparse estimation, the scaled Gaussian mixture is imposed on the scattering coefficients, which are assumed to be independently distributed. However, in practice, targets in radar images always exhibit strong spatial correlation due to the fact that a real target is physically continuous [39], [44]–[46]. For example, the radar returns from a tank or an airplane will often exhibit strong spatial correlation, i.e., nonzero-valued scatterers in the target region continuously residing in the range and/or cross-range dimensions. This phenomenon motivated the research in [39], [45], and [46], which modifies the statistical sparse model accordingly. In these works, continuity in the target scene is exploited by incorporating a correlated prior in a probabilistic framework. In what follows, we review methods that impose first-order and higher-order correlations on the sparse scattering coefficients.

First-order correlation

In [39], [44], and [45], a modification in spike-and-slab modeling was made so as to impose first-order spatial correlation of the coefficients. The reason for choosing the spike-and-slab prior rather than the scaled Gaussian mixture is because imposing correlation on the support of the sparse signal is more accurate and justifiable than imposing it on the amplitude of the sparse signal. As

discussed in the “Statistical Sparsity Formulation of Radar Imagery” section, the sparsity pattern of the signal is determined by \mathbf{W} in the spike-and-slab model in (7), where the parameter \mathbf{q} controls the probability of \mathbf{W} being nonzero. Therefore, a straightforward way to impose a continuity prior on the signal can be carried out directly on \mathbf{W} . However, this treatment deviates from the original intention to perform a flexible statistical modeling step. For this particular reason, it is suggested in [39] and [45] to encode the first-order structural information on \mathbf{q} in an intermediate way rather than straightforwardly on \mathbf{W} . The key modification is to replace the single beta prior for parameter \mathbf{q} by a set of beta priors that consist of three different sets of parameters, $\{e_k, f_k\}_{k=0,1,2}$, so as to capture strongly independent, strongly continuous, and noninformative priors, respectively.

The proposed sparsity patterns in [39] and [45] that both encourage continuity and preserve sparsity are summarized as follows.

- **Strong rejection:** If the first-order neighborhoods of \mathbf{X}_{mn} are all zero, it would be very likely that \mathbf{X}_{mn} is also zero, due to the continuity of the target scene. The prior $\text{Beta}(e_0, f_0)$, with $e_0 < f_0$, is utilized to make the probability q_{mn} of $W_{mn} = 0$ being large. This means that the absence of a first-order neighborhood implies the investigated scatterer being zero with a high probability. This rejection pattern can eliminate the undesired isolated speckles or artifacts in the radar image.
- **Strong acceptance:** If any of the continuity patterns for \mathbf{X}_{mn} in Figure 3 is observed, the prior that a nonzero-valued \mathbf{X}_{mn} arises with a high possibility should be imposed. This step imposes continuity of the target image. In this case, the prior $\text{Beta}(e_1, f_1)$, with $e_1 > f_1$, enforces the probability q_{mn} of $W_{mn} = 1$ to be large, and thus the scatterer under test can be accepted. This implies that the occurrence of any pattern in Figure 3 leads to one that is nonzero with a high probability. This pattern enforces first-order correlation of the scattering coefficient and therefore continuity of the target.
- **Weak rejection:** Apart from the scenario of strong rejection and strong acceptance patterns, a noninformative prior is imposed on any other neighborhood patterns for \mathbf{X}_{mn} . The prior $\text{Beta}(e_2, f_2)$, with $e_2 = f_2$, is used to impose a noninformative prior on q_{mn} . This appropriately allows the model to be effective in imposing the prior whenever necessary and to remain noninformative whenever no strong rejection or acceptance patterns appear.

By adaptively selecting from different beta hyperpriors, the statistical model can either encourage continuity or independence, apart from mere sparsity. In this manner, the structured information can be flexibly incorporated to obtain concentrated imagery results. A key component in incorporating the prior is that it is imposed on the parameter \mathbf{q} rather than directly on \mathbf{W} . The underlying motivation for this formulation is that it is more flexible to impose a probabilistic belief than a rigid support \mathbf{W} .

In Figure 4, the real Yak-42 data are used to test different algorithms, where the radar image obtained with all measurements is shown in Figure 4(a) for reference purposes. In general, there are two issues to be considered in the evaluation of radar images: first, how well the target is concentrated, i.e., more true scatterers preserved in the target region and less artifact recovered outside the target region, and second, how well the radar image is focused, i.e., lower sidelobe and noise. As shown in Figure 4(b), the radar image obtained by the range-Doppler algorithm (RDA) is highly corrupted by noise. Although the ℓ_1 -regularized approach can achieve better performance than RDA by exploiting sparsity, the obtained target image is not well concentrated, and artifacts around the target are not removed, as shown in Figure 4(c). Notably, the method in [39] that exploits first-order continuity patterns performs best, as shown in

With the recent development of sparse Bayesian methods, statistical sparsity-based techniques have become a more promising research area for radar imagery applications.

Figure 4(d). More specifically, the first-order continuity method in [39] outperforms the ℓ_1 -regularized approach in terms of radar target concentration and artifact removal, and is even clearer than the reference image obtained with all measurements as shown in Figure 4(a).

Higher-order correlations

A real radar image generally exhibits higher-order correlations than simply horizontal or vertical correlations, exploited in [39] and [45]. This motivates an extension of the first-order method. To formulate a generalized framework, a more sophisticated model is developed in [46] that captures higher-order correlations based on Markov random fields (MRFs). The MRF model is widely used in image processing for imposing structural constraints on the image. This work presents a unified framework of incorporating more complex structural information in the target scene, as

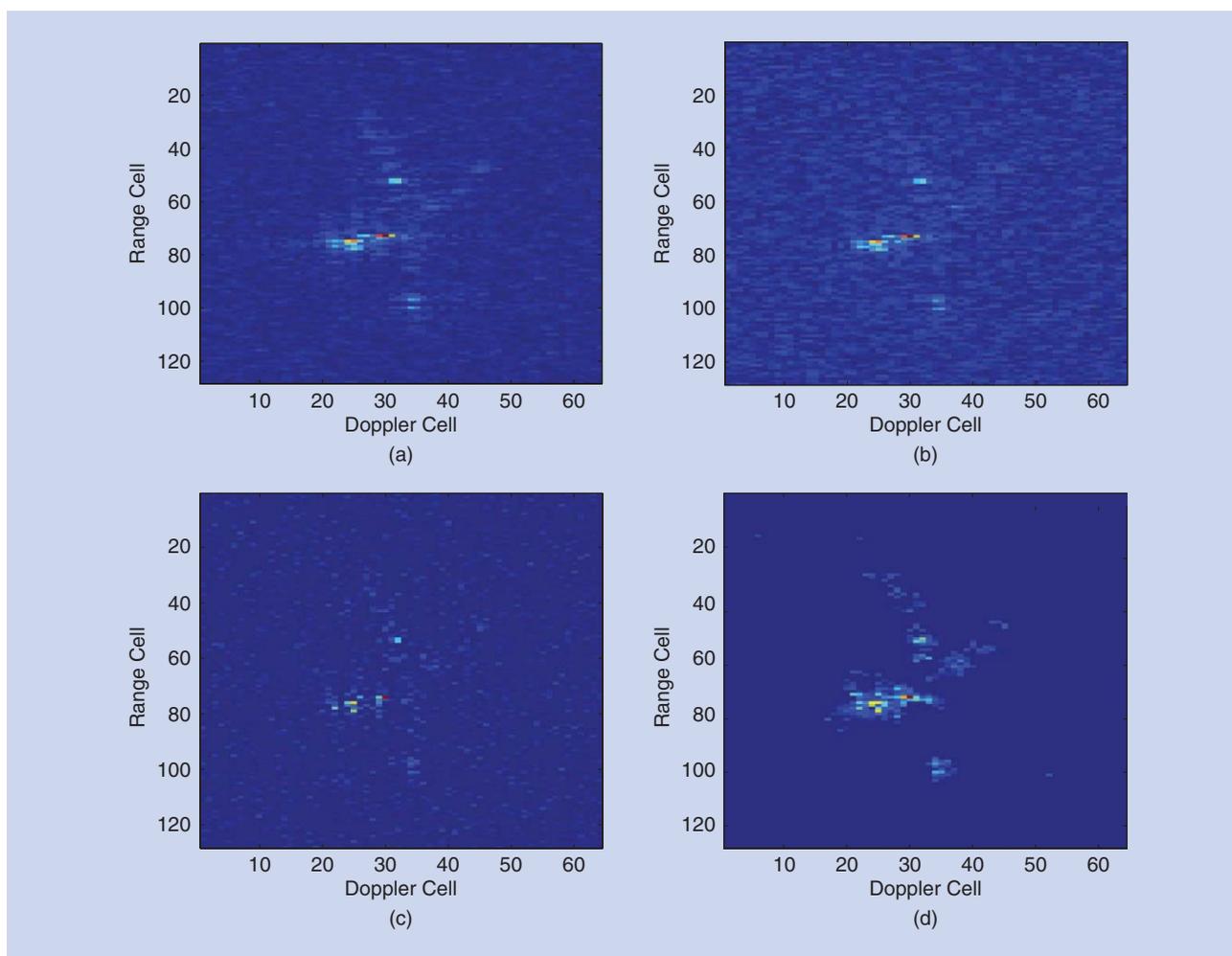


FIGURE 4. The radar images obtained by (a) RDA in the presence of noise, (b) RDA using one-half of the measurements, (c) an ℓ_1 -regularized method using one-half of the measurements, and (d) a first-order continuity method [39] using one-half of the measurements.

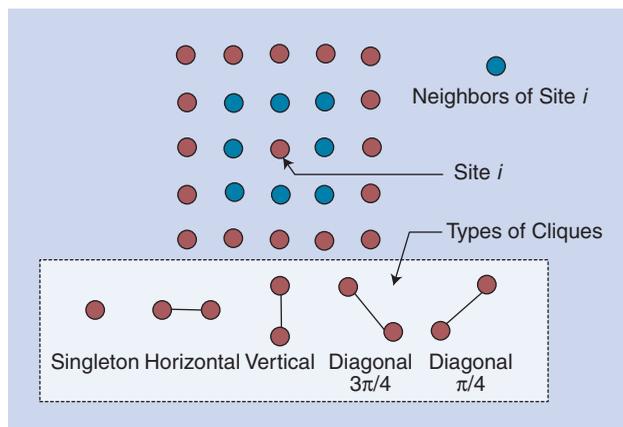


FIGURE 5. An example of a second-order MRF model, where eight neighbors are considered.

compared with the simple first-order approach. In Figure 5, the construction of the MRF model is presented, where a second-order neighborhood system is employed. This model allows continuity from four directions, i.e., horizontally, vertically, $3\pi/4$, and $1\pi/4$ diagonally. The authors argued that adopting second-order MRF enables the capture of correlations of the investigated scatterer with its nearest eight neighbors, as shown in Figure 5.

To impose continuity of the target scene, the authors in [46] considered a more general structured sparse prior as compared with that in [39] and [45]. A more complicated continuity prior has been proposed by modifying the spike-and-slab modeling to better preserve the weak scatterers. Moreover, the hyperparameter selection in [39] and [45] is avoided by adopting an MRF prior, since all the parameters can be automatically inferred. This is a very desirable feature for statistical inference. Based on this model, the authors employed a VB expectation maximization method for inference, where an improved rate of convergence can be obtained, as compared with the method in [39] and [45]. As commented earlier, the VB-based method generally requires less computational complexity than the MCMC-based one.

In Figure 6, we can observe that both the first-order continuity method in [39] and [45] and the second-order continuity method in [46] produce much-enhanced radar images in the sense of less noise and a better-concentrated target region, compared with the ℓ_1 -regularized method. As shown in Figure 6(b) and (c), the second-order continuity-based method performs much better than the first-order one in terms of removing the undesirable isolated artifacts and preserving weak scatterers outside and within the target region, respectively. More important, the computational time of the second-order continuity method is much less than that of the first-order one [46].

The above discussion demonstrates how, by incorporating structural priors in addition to sparsity, a statistical framework provides superior performance compared to a merely sparsity-based framework. The major advantage of these approaches is

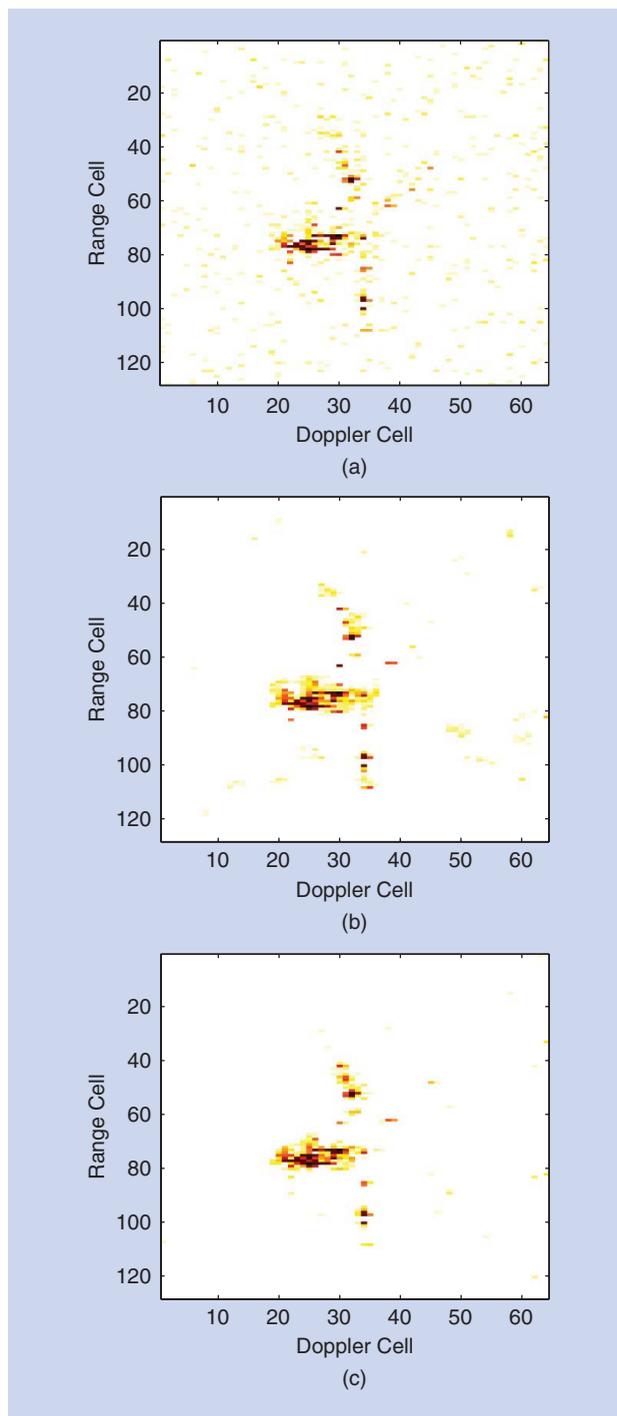


FIGURE 6. A comparison of radar images with a quarter of the full measurements and SNR = 5 dB obtained by (a) an ℓ_1 -regularized method, (b) a method in [39], and (c) a method in [46].

that they can statistically impose the structured sparsity on the signal in a rather flexible way, which allows the algorithm to adapt the structured sparse estimation in a data-driven manner. More specifically, in all the introduced models, the structural information is not directly imposed on the sparse signal itself but on the probability distribution that determines the sparsity profile.

Statistical sparsity-based autofocus techniques in radar imagery

The CS-based radar imagery techniques discussed in the previous sections generally depend on the premise that pre-processing procedures, such as range cell migration (RCM) correction and phase adjustment, have been perfectly conducted. Unfortunately, this is not a valid assumption from a practical viewpoint, since the motion of the target cannot be precisely compensated in coarse pre-processing stages. If these errors are not properly corrected or compensated for before carrying out any CS-based algorithms, the reconstructed radar image will not be well concentrated.

Recently, phase error correction has been considered by utilizing a sparse recovery technique, where alternating ℓ_1 -regularized approaches [14] were proposed to obtain more focused images. In these methods, the sparse scattering coefficient and the phase errors are iteratively estimated to induce sparsity and obtain a focused radar image. Although these methods have demonstrated remarkable improvements over conventional autofocus techniques, these regularization-based methods might converge to a shallow local minimum during the iterative procedure. The alternate optimization between the sparse scattering coefficient and the phase error would inevitably result in error propagation [47]. More concretely, the alternate optimization scheme would introduce errors, since the estimation accuracy of one parameter substantially influences that of another. This issue is particularly severe with undersampled data and in low SNR conditions. To appropriately overcome the aforementioned limitations, high-resolution imagery and phase error correction have been formulated in a statistical sparsity-based model [47]–[49]. In this formulation, probabilistic models are imposed on the signal to encode sparsity in a statistical way. Subsequent parameter estimation is conducted within a sparse Bayesian learning framework [19], [47].

Statistical sparsity-based autofocus

Assuming that the phase error in radar imagery exhibits range invariance [15], the mathematical model can be stated as

$$\mathbf{Y} = \mathbf{E}\Phi_1\mathbf{A}\mathbf{X} + \mathbf{N}, \quad (15)$$

where $\mathbf{E} = \text{diag}(e^{j\varphi_1}, \dots, e^{j\varphi_p})$ denotes the phase error matrix, which is a diagonal matrix representing cross-range variant phase errors. In [47], the authors utilized the scale Gaussian mixture model to impose sparsity on \mathbf{X} . The estimation of \mathbf{X} , α , and λ is obtained individually, since they are task-specific parameters, while estimation of α_0 and \mathbf{E} is performed in a global manner due to the task-invariant property.

According to the graphical model [47], the parameters \mathbf{X} , α , λ , and α_0 , can be estimated in a way similar to that

in the scaled Gaussian mixture model introduced previously. The most straightforward way to obtain an estimate of the phase error \mathbf{E} is to maximize the expected log-likelihood function as

$$\hat{\mathbf{E}} = \arg \min_{\mathbf{E}} \langle -\ln p(\mathbf{Y}, \mathbf{X}, \alpha, \lambda; \mathbf{E}) \rangle_{q(\mathbf{X})q(\alpha)q(\lambda)}. \quad (16)$$

Equation (16) is strictly convex with a close-form solution. By solving the optimization problem in (16), the updating formula can be obtained [49]. This updating rule for phase error is rather similar to that of the regularized approach in [14] and [15], because the updating formula could use only the first-order moment of \mathbf{X} to estimate \mathbf{E} , and the obtained covariance matrix Σ of \mathbf{X} does not appear in this updating rule. In other words, this formulation deviates from the original intention of utilizing higher-order statistical information in the first place.

By incorporating structural priors in addition to sparsity, a statistical framework provides superior performance compared to a merely sparsity-based framework.

To properly utilize the uncertainty information, the work in [47] proposed to incorporate the obtained covariance matrix Σ in the estimation of phase errors to obtain enhanced accuracy. Toward this end, the phase error is deliberately modeled as a complex parameter $a_i + jb_i$ rather than explicitly as $e^{j\varphi_i}$. By introducing this complex parameter instead of the angle parameter φ_i , we will see that the uncertainty information can be naturally incorporated in the algorithm to achieve enhanced estimation accuracy of \mathbf{E} in each iteration. In the derived updating formula in [47], it can be seen that Σ , which contains uncertainty information, can be incorporated into the estimation of the phase error parameter \mathbf{E} . It is demonstrated in [47] that by replacing the true phase error parameters with complex-valued error parameters, the resulting scheme could utilize the estimation uncertainty information and obtain a performance gain as compared with regularized sparsity-based autofocus techniques.

In Figure 7, an illustrative example is presented to evaluate the performance of the updating rule without and with high-order uncertainty information. In this simulation, a total of 11 scatterers are present in the imaging scene. In Figure 7(a), the random phase error is shown. Figure 7(c) and (d) shows that both updating rules lead to a more focused image compared with the RDA shown in Figure 7(b). In particular, the updating rule without utilizing the uncertainty information leads to a less focused image, where undesirable sidelobe effects exist for almost all the scatterers on the imaging scene. In contrast, the image obtained by the updating rule utilizing the uncertainty information is well focused, with substantially suppressed sidelobe effects. Quantitative evaluation also demonstrates that the radar image in Figure 7(d) provides a lower NMSE_x as well as MSE_φ than those obtained in Figure 7(c) due to the inherent ability to utilize the uncertainty information of estimation of \mathbf{X} . This validates the motivation of utilizing the uncertainty

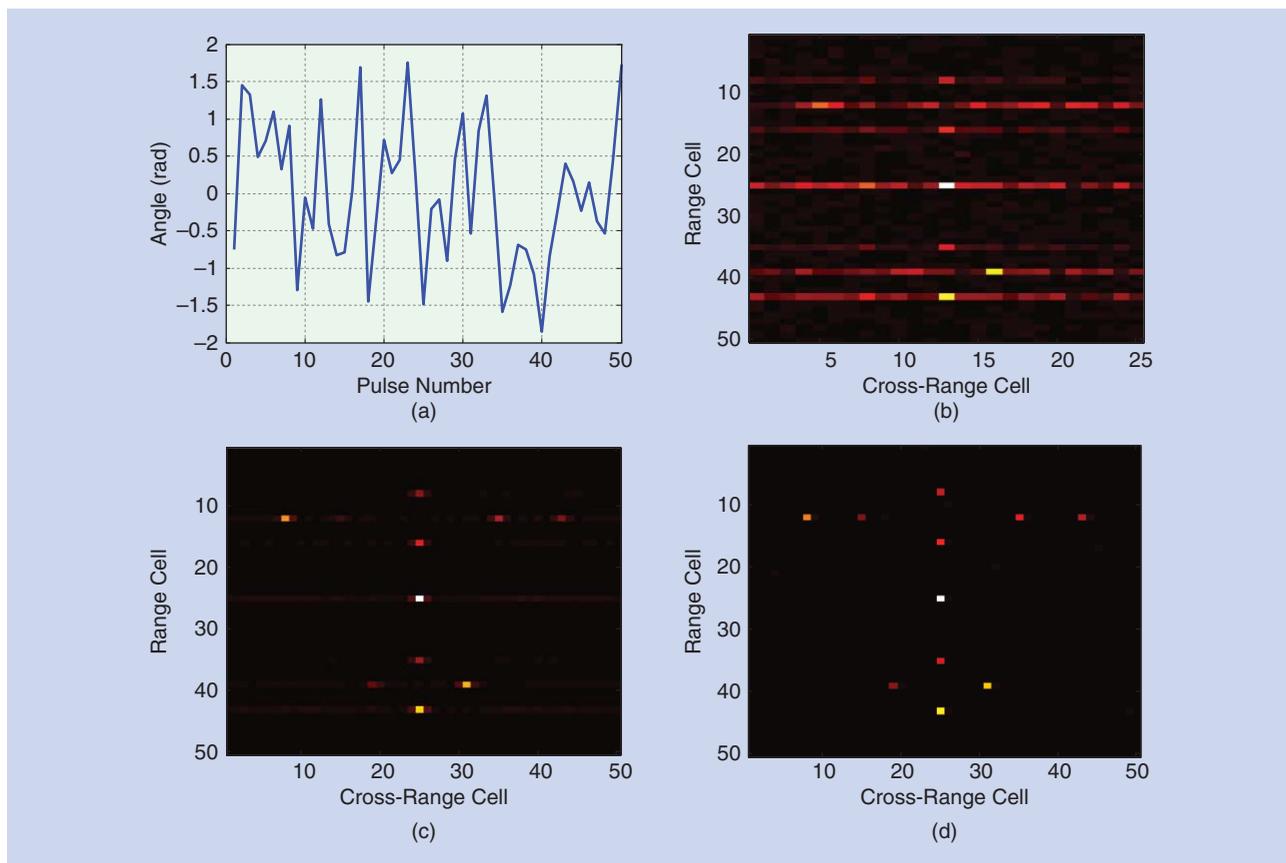


FIGURE 7. The radar imaging with 25 pulses (50% of the full measurements and SNR = 20 dB): (a) random phase error, (b) RDA method, (c) a method in [47] (without uncertainty information) ($NMSE_x = -4.0294$ dB, $MSE_\phi = 0.4019$), (d) a method in [47] (with uncertainty information) ($NMSE_x = -12.4399$ dB, $MSE_\phi = 0.0059$).

information to achieve higher estimation accuracy and therefore a better-recovered radar image. This work is also validated using the Yak-42 data. As observed in Figure 8(a), directly applying the RDA method can barely lead to a concentrated image. In Figure 8(b), the image obtained by ℓ_1 -minimization presents a reasonable profile of the airplane. However, it is still blurred, and some of the true scatterers are not recovered correctly. In contrast, the method in [47] obtains a better-concentrated image and removes most of the undesirable artifacts, as seen in Figure 8(c). With these comparisons, we conclude that the proper utilization of uncertainty information substantially enhances the performance in autofocus applications.

The statistical treatment of the sparsity-based algorithm can properly utilize uncertainty information during iterations to improve the estimation accuracy. In fact, we have demonstrated how this uncertainty information can be properly used throughout this particular application. Compared with the regularized approach, statistical sparsity-based algorithms do not require the time-consuming parameter tuning for improved performance as those in the ℓ_1 -regularized cases.

Autofocus meets structured sparsity

The basic idea in [47] is to iteratively estimate the sparse scatterer coefficients and the phase error to jointly induce sparsity. However, the objective of radar imaging is to obtain the most concentrated radar image rather than the sparsest one. Therefore, a merely sparsity-inducing scheme may result in undesirable image results, because it considers only sparsity as the performance measure. More precisely, a simple sparsity constraint cannot sufficiently preserve the weak scatterers or reduce background noise to a desirable

quality. A possible solution to obtain a more concentrated radar image rather than a mere sparse one is to exploit structured sparsity. In [49], the sparse Bayesian model is sophisticatedly modified to exploit structural sparsity. More specifically, the spatial consistency along range cells is exploited, and therefore the framework can cope simultaneously with structured sparse signal recovery and phase error correction in an integrated manner. The focused high-resolution radar image can be obtained by iteratively estimating the sparse scatterer coefficients and phase errors to jointly obtain a structured sparse solution.

The statistical treatment of the sparsity-based algorithm can properly utilize uncertainty information during iterations to improve the estimation accuracy.

Due to the utilization of the structured sparse constraint, the proposed method preserves the target region and alleviates the overshrinkage problem, as compared to the previously presented sparsity-driven autofocus approaches. The superior performance of the structured sparsity-based technique is shown in Figure 9. Compared with other approaches, the structured sparsity-based autofocus method achieves a better-concentrated image, with more coefficients recovered in the target region with different under-sampling ratios.

Statistical sparsity-based SAR GMTIm

Imaging ground moving targets in SAR has become increasingly important. Conventionally, in imaging a potentially moving target, hypotheses of the target motion were constructed to match the signal by a filter bank [50]. In the scenario of closely located targets, however, their responses cannot be well distinguished from each other. Recent advances in sparsity-based SAR GMTIm [5], [16], suggest that sparsity can be properly exploited to enable multitarget processing and higher accuracy. This application is rather different from the previously introduced ones, since the received radar echoes can no longer be simply modeled as a sum of harmonics but rather as multicomponent LFM signals with unknown chirp rates. Therefore, the key challenge in SAR GMTIm is to properly formulate a mathematical model that allows a sparse representation of the images for moving targets. In [16], the signal model was constructed as a sparse linear model, where an overcomplete dictionary was constructed by using a discretized velocity grid. Although empirical results demonstrate the success of the method, its performance is inhibited by the discretization errors in the dictionary. In this section, we briefly review two recent works based on statistical sparsity from different angles.

In [51], a statistical framework was formulated to obtain the moving target image, which could avoid the construction of a large overcomplete dictionary. In particular, this work considers a K channel SAR system with F passes, collecting data from P azimuth angles and Q range cells. The complex-valued raw SAR image is decomposed as [51]

$$\begin{aligned} \mathbf{Y}_{p,f} &= \mathbf{E}_{p,f} \cdot (\mathbf{L}_{p,f} + \mathbf{S}_{p,f} + \mathbf{N}_{p,f}), \\ p &= 1, \dots, P \text{ and } f = 1, \dots, F, \end{aligned} \quad (17)$$

where $\mathbf{Y}_{p,f} \in \mathbb{C}^{Q \times K}$ denotes the raw SAR image at azimuth p and pass f , $\mathbf{E}_{p,f} \in \mathbb{C}^{Q \times K}$ is the corresponding spatial-temporal calibration error, $\mathbf{L}_{p,f} \in \mathbb{C}^{Q \times K}$ represents background clutter, $\mathbf{S}_{p,f} \in \mathbb{C}^{Q \times K}$ models the moving target, and $\mathbf{N}_{p,f} \in \mathbb{C}^{Q \times K}$ models the noise. Since the number of parameters to be estimated is much larger than the number of observations Y , proper priors must be selected for each of these parameters. Interested readers are referred to [51] for more details. Here, we only highlight the key statistical models in this formulation.

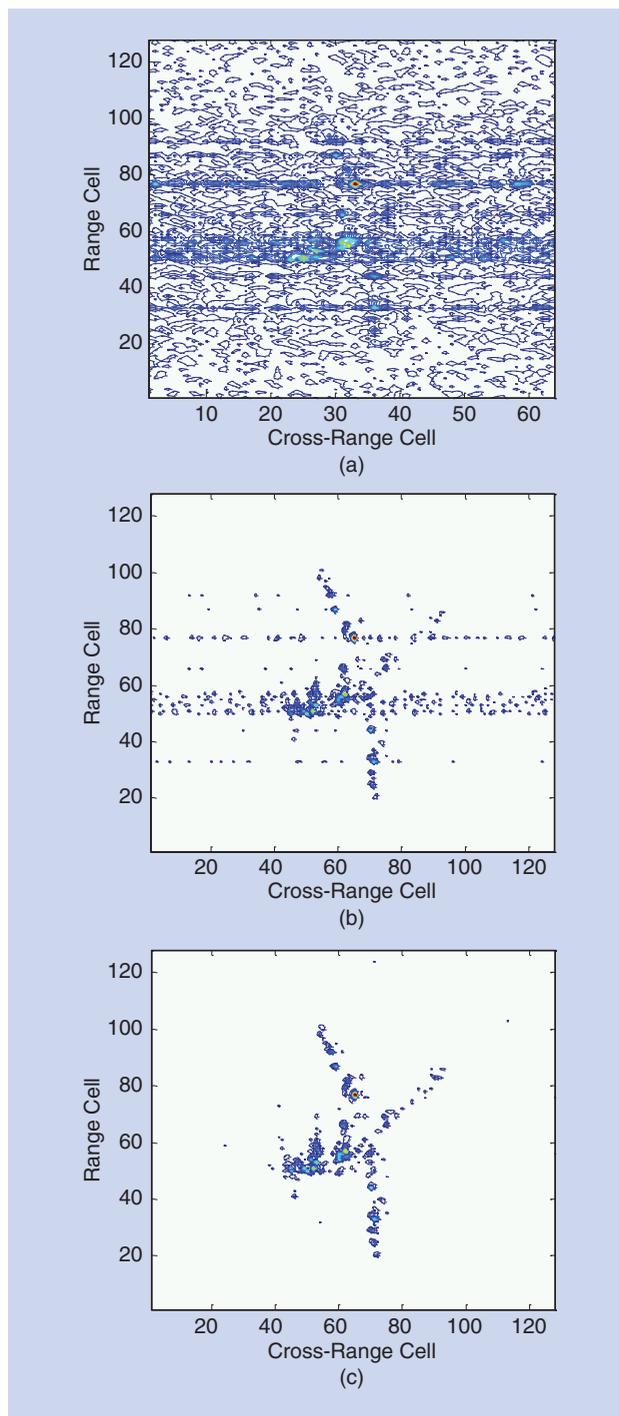


FIGURE 8. The radar imagery results with one-half of the measurements, obtained by (a) RDA, (b) a method in [14], and (c) a method in [47].

- The clutter $\mathbf{L}_{p,f}$ is decomposed into a sum of a pass-invariant background term \mathbf{B}_p and a pass-specific speckle term $\mathbf{X}_{p,f}$. Assuming that the background clutter \mathbf{B}_p can be represented by one of several classes, such as a road or buildings, a complex Gaussian prior is used for \mathbf{B}_p with a set of unknown covariance matrices that account for different classes, where each covariance matrix follows an inverse Wishart distribution.

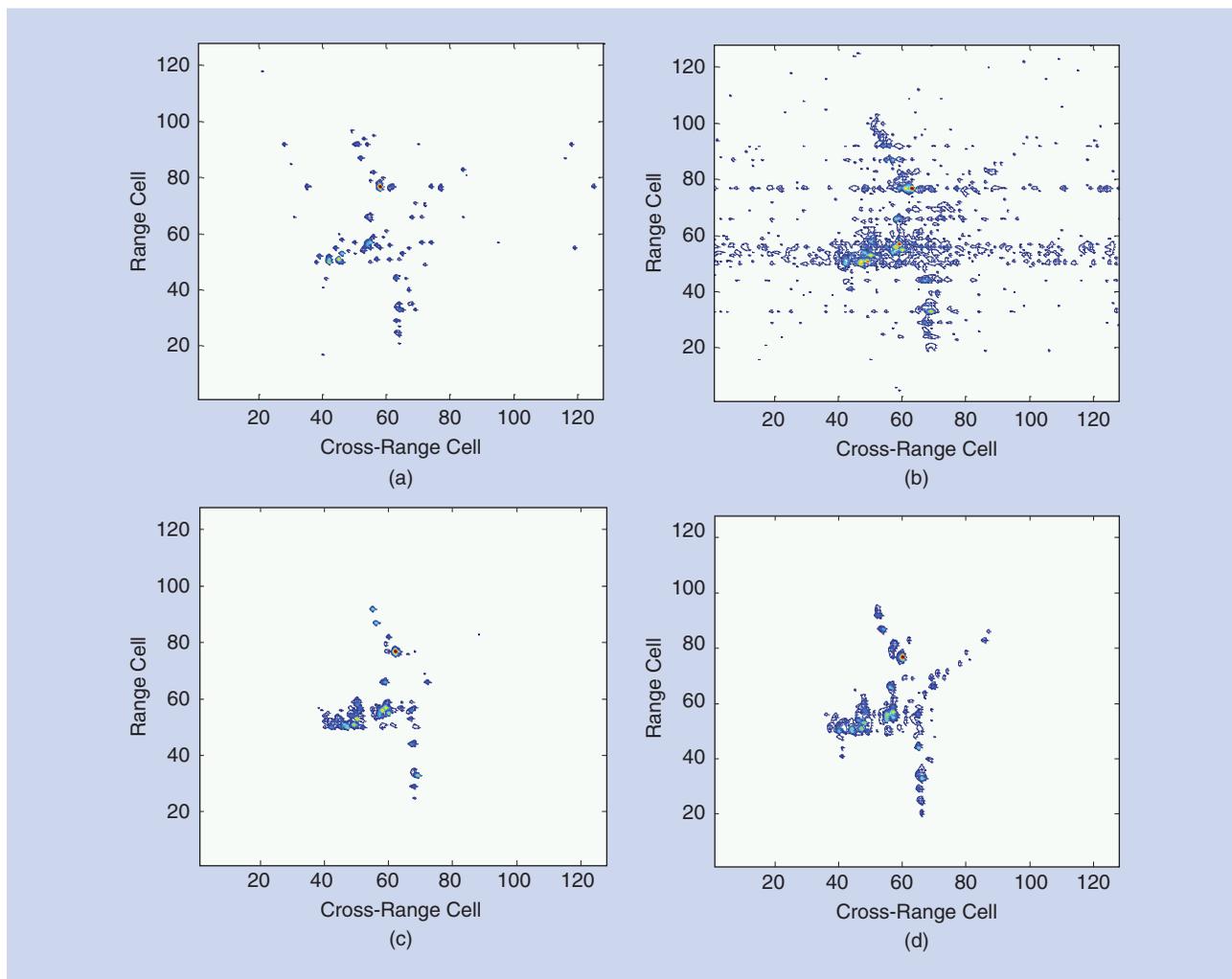


FIGURE 9. The radar imagery results with SNR = 5 dB using (a) an ℓ_1 -regularized method, (b) a TV + ℓ_1 -regularized method, (c) the statistical sparsity-based method in [47], and (d) the statistical structured sparsity-based method in [49].

Similarly, the pass-specific speckle term $\mathbf{X}_{p,f}$ is modeled probabilistically.

- Since the moving target is assumed to be sparse in the raw SAR image domain, a modified spike-and-slab model captures sparsity as well as moving target signatures. More specifically, the sparsity is modeled by a Bernoulli–beta distribution, and the moving target signature is modeled as a complex Gaussian-inverse Wishart distribution. The rationale of this modeling is to allow a rather tractable inference.
- An additional constraint can be imposed on the hyperparameters of the sparse moving target to encourage a smooth trajectory. It is noted that this smooth prior is constructed by modifying the beta distribution instead of the support parameter directly, where this manipulation can be rather

The theory of CS states that a high-dimensional signal can be accurately and robustly recovered from its low-dimensional projections if the signal is sparse or can be sparsely represented.

flexible in encoding the prior information in a probabilistic sense.

Since the work in [51] utilized the decomposition of a raw SAR image, the construction of the dictionary as in [16] could be avoided. Since this method is formulated in a statistical framework, the algorithm could utilize the uncertainty information obtained in one parameter to subsequently enhance the estimation of other parameters. These desirable properties of the statistical sparsity-based method have led to substantial improvements over conventional methods.

Another approach for SAR GMTIm is based on formulating a parametric model, where statistical sparsity is enforced [52]. In this work, the clutter was assumed to have been suppressed by off-the-shelf methods, and an initial representation of the received signal was first carried out by utilizing Lv’s distribution (LVD) [53], which is a novel time–frequency

representation for representing LFM signals. However, the resolution of the LVD is constrained by the CPI of the target and its discretized grid [53]. It should be noted that the limited accuracy of the LVD may cause an unfocused target response and, thus, a degraded target image. To deal with this challenge within a statistical sparsity-based framework, a dynamical refinement was suggested for an accurate estimation of the chirp rate in [54]. In particular, this dynamic refinement iteratively refines the initialized γ_i by the LVD and the sparse target coefficient. In this way, the estimation accuracy can be improved in a statistical sparsity framework,

and therefore a concentrated moving target image can be obtained. Considering P azimuth and Q range cells, the clutter-suppressed signal model can be formulated as in [54]

$$\mathbf{Y} = \mathbf{E}\mathcal{A}(\gamma_1, \dots, \gamma_K)\mathbf{X} + \mathbf{N}, \quad (18)$$

where $\mathbf{Y} \in \mathbb{C}^{P \times Q}$ is the clutter-suppressed data, $\mathbf{E} \in \mathbb{C}^{P \times P}$ represents the unknown phase errors, $\mathbf{X} \in \mathbb{C}^{KN \times Q}$ models the sparse moving target to be estimated, $\gamma_i, i = 1, \dots, K$, is a set of parameters in the dictionary to be estimated,

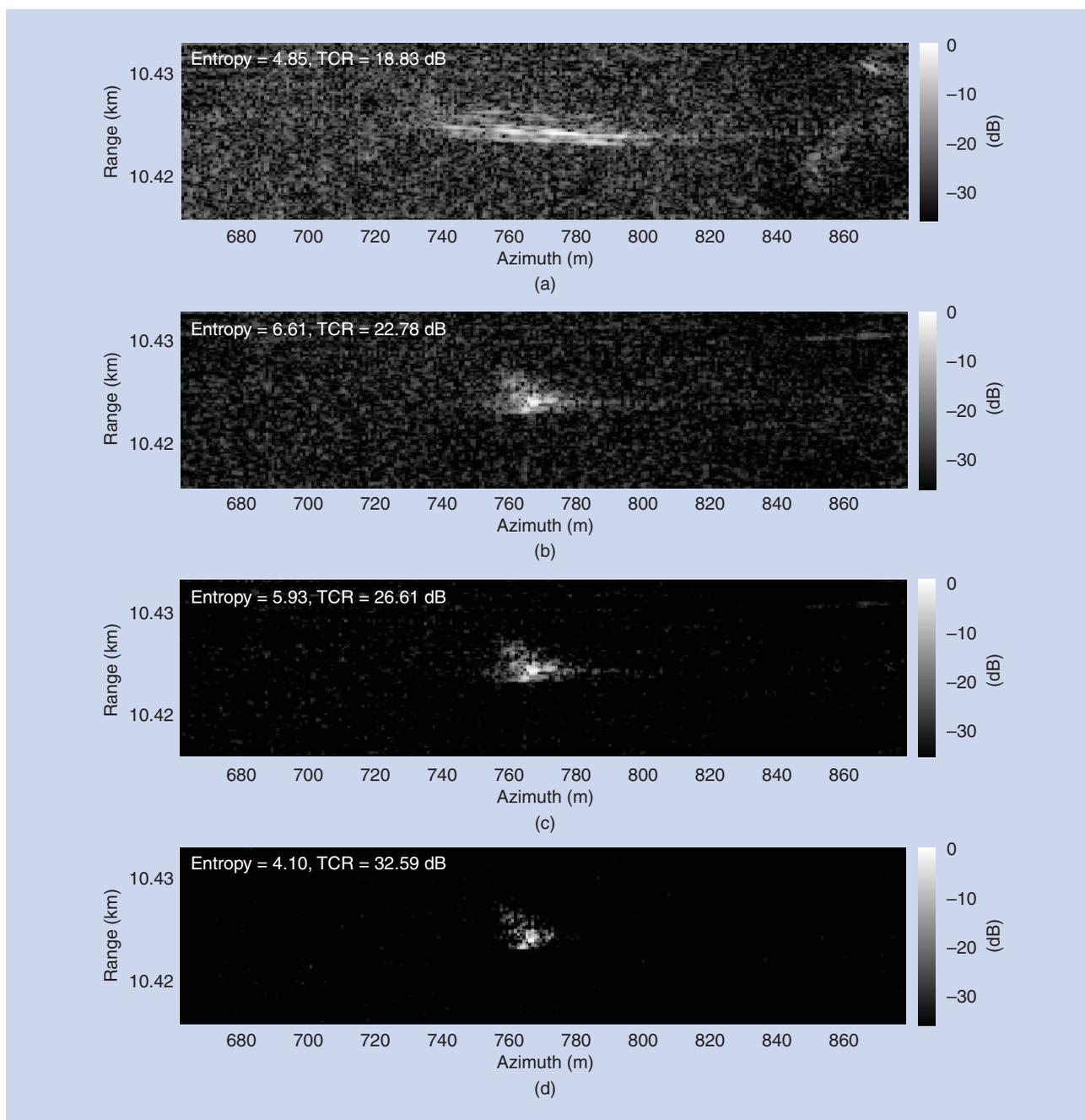


FIGURE 10. The Durango target image: (a) the original image, (b) the image obtained by ℓ_1 -norm regularization, (c) the image obtained by a conventional statistical sparsity-based method, and (d) the image obtained by the parametric and dynamic statistical sparsity-based method in [54].

and K is the number of moving targets. As described previously, the dictionary $\mathcal{A}(\gamma_1, \dots, \gamma_K) \in \mathbb{C}^{P \times KN}$ is an overcomplete one. It is constructed by concatenating K subdictionaries, where each subdictionary is constructed by an LFM matrix with chirp rate γ_i . In [54], a scaled Gaussian mixture distribution is used to model sparsity. Similar to the work covered in the “Statistical Sparsity-Based Autofocus Techniques in Radar Imagery” section, statistical information is utilized to estimate the error parameter \mathbf{E} and the chirp rate γ_i , where the error propagation problem during iteration is reduced [47], [54].

In Figure 10, the canonical Gotcha data set is used for validation, and an example of the Durango target image is given to demonstrate the performance. Due to the movement of the target, the original image is substantially blurred, as observed in Figure 10(a). After representing the received signal by the LVD, the ℓ_1 -norm regularization method and the conventional sparse Bayesian method are applied to obtain the moving target images, as shown in Figure 10(b) and (c), respectively. The ℓ_1 -regularized method and the sparse Bayesian method cannot properly focus the target image due to the representation error in the LVD. In contrast, the statistical sparsity-based method with refinement leads to the best imaging performance in terms of better concentration and desirable noise suppression, as shown in Figure 10(d). The superior performance of the statistical sparsity-based method is also evaluated quantitatively by the calculated entropy and target-to-clutter ratio (TCR) as shown in Figure 10. In particular, the target image is focused within a $5 \text{ m} \times 5 \text{ m}$ area that is in accordance with the Durango truth with a size of $5 \text{ m} \times 2 \text{ m}$.

Summary and future directions

Summary

Sparsity-based techniques have been reviewed from a statistical perspective, along with their recent advances in radar imagery. Various applications show that improved performance can be obtained by adequately utilizing a statistical sparse model. The improvements obtained in the reviewed applications were largely dependent on the following core ingredients:

- Probabilistic modeling by incorporating flexible priors in the signal is one of the most remarkable advantages over deterministic approaches. The advantage of the statistical framework is its flexibility. In this way, the formulation could model a particular structure in a probabilistic way and also allows for a fitting of the likelihood.
- The utilization of uncertainty information during parameter estimation is important for a performance gain. Particularly, in conventional approaches, the error

Compared to the deterministic sparsity-inducing framework, statistical sparsity-based techniques provide new opportunities to significantly improve the performance of radar imagery.

estimated in one stage can lead to a degraded performance in the subsequent stages. In the discussed framework, the signal estimation is conducted in a statistical manner, where the obtained statistics indicate the uncertainty in the signal estimation. Therefore, the estimation could be more accurate.

By properly manipulating the statistical sparsity models, a performance gain can be obtained.

Future directions

Since the statistical sparsity-based methods are quite attractive, it would be most interesting to investigate the following problems in the future.

- *Computational complexity.* The statistical sparsity-based methods operate in an iterative manner, where the number of iterations and the computational cost of each iteration determine the total computational cost. Compared to the conventional Fourier-based approach for radar imaging, the computational complexity is much higher. It is therefore imperative to develop fast algorithms that could decrease the computational complexity or obtain fast convergence. The fast algorithms would be particularly useful for many radar applications requiring real-time processing.
- *Motion compensation errors.* In high-resolution radar imaging, a large CPI is required. Then the target movement becomes a problem as the radar line-of-sight dramatically changes. In such a scenario, even after carrying out coarse motion compensation, RCM and phase error would still be present in the radar echoes. Then, the dictionary allowing sparse representation would become more complicated, where the proposed imaging algorithm should also be able to correct RCM and phase errors. The main challenge is to properly obtain the approximated solution in the presence of a more complicated model. Toward this end, it would be particularly suitable to exploit statistical sparsity to limit error propagation. One possible way of coping with this challenge is to encode priors on the error parameters to properly regularize the solution space.
- *Temporal correlation in SAR GMTIm.* Conventionally, most SAR GMTIm algorithms focus on image formation of the moving target at one particular time instant. However, it is important to also monitor the movement of the moving target. Since the target's motion and imaging background are time-varying, simply generating a single-frame image cannot provide time-varying characteristics of the moving target. Therefore, it is necessary to develop temporal SAR GMTIm based on the statistical sparsity-based framework, which is a promising research direction in SAR GMTIm technology. In fact, Sandia Laboratory has successfully realized Video-SAR GMTIm, where the processed results have been released on their official website. In particular, the

temporal SAR GMTIm is a good candidate for applications in complicated urban scenes, where improved performance is valuable. The statistical sparsity-based framework for moving target imaging in urban environments could be formulated to include the temporal smoothness constraint during the radar passes and to cope with a complicated background.

- **Improved classification performance.** An important objective of radar imagery is to classify different types of targets automatically and accurately. One should capture more structural features during target imaging by utilizing the training information obtained from the recognition stage, which will in turn greatly benefit automatic target recognition. More precisely, radar imagery should be discriminative enough for target recognition purposes. One promising future work direction is to incorporate appropriate priors in a statistical framework to perform discriminative radar imagery.

In summary, statistical sparsity-driven techniques have been shown to be very promising for radar imagery due to their flexibility and good statistical properties. It is expected that these applications will immensely benefit from the more recent theoretical advances in this area.

Acknowledgment

This work is partially supported by project fund MOE2014-T2-1-079 and RG103/14, Singapore.

Authors

Lifan Zhao (zhao0145@e.ntu.edu.sg) received his B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2010, and his Ph.D. degree from the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore, in 2016. He is currently a research fellow in the School of Electrical and Electronic Engineering, NTU. His research interests include statistical signal processing, compressed sensing, machine learning, Bayesian methods, and their applications in radar and communications. He has published more than 20 papers in these areas.

Lu Wang (wangle@nwpu.edu.cn) received her B.S. degree in electrical and electronic engineering from Xidian University, Xi'an, China, in 2007; her M.Eng. degree in signal processing from the National Key Laboratory of Radar Signal Processing, Xidian University, in 2010; and her Ph.D. degree from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2014. She is currently an assistant professor in the School of Marine Science and Technology, Northwestern Polytechnic University, Xi'an. Her major research interests include sparse signal processing, radar imaging [synthetic aperture radar (SAR)/inverse SAR], and sonar signal processing.

One promising future work direction is to incorporate appropriate priors in a statistical framework to perform discriminative radar imagery.

Lei Yang (yanglei@ntu.edu.sg) received his B.S. degree in electronic engineering and his Ph.D. degree in signal and information processing, both from Xidian University, Xi'an, China, in 2007 and 2012, respectively. He is currently working as a research fellow in the School of Electrical and Electronic

Engineering, Nanyang Technological University, Singapore. His main research interests include high-resolution airborne and spaceborne synthetic aperture radar (SAR) imaging, data-driven motion error compensation (autofocusing) algorithm for unmanned aerial vehicle SAR, and real-time processing systems for SAR. He is a Member of the IEEE.

Abdelhak M. Zoubir (zoubir@spg.tu-darmstadt.de) received his Dr.-Ing. degree from Ruhr-Universität Bochum, Germany. He was with Queensland University of Technology, Australia, from 1992 to 1998. He then joined the Curtin University of Technology, Australia, as a professor of telecommunications and was interim head of the School of Electrical and Computer Engineering from 2001 to 2003. Since 2003, he has been a professor of signal processing at the Darmstadt University of Technology, Germany. He is an IEEE Distinguished Lecturer (class of 2010–2011), past chair of the IEEE Signal Processing Society's Signal Processing Theory and Methods Technical Committee, and past editor-in-chief of *IEEE Signal Processing Magazine* (2012–2014). His research interest lies in statistical methods for signal processing applied to telecommunications, radar, sonar, car engine monitoring, and biomedicine. He has published more than 400 papers in these areas. He is a Fellow of the IEEE.

Guoan Bi (egbi@ntu.edu.sg) received his B.S. degree in radio communications from the Dalian University of Technology, China, in 1982, and his M.S. degree in telecommunication systems and his Ph.D. degree in electronics systems from Essex University, United Kingdom, in 1985 and 1988, respectively. Between 1988 and 1990, he was a research fellow at the University of Surrey, United Kingdom. Since 1991, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He was the chair of the IEEE Signal Processing Society, Singapore Chapter. His research interests include digital signal processing algorithms and hardware structures and signal processing for various applications, including sonar, radar, and communications. He is a Senior Member of the IEEE.

References

- [1] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar*. New York: Wiley, 1991.
- [2] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [3] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, no. 6, pp. 1006–1020, June 2010.
- [4] J. H. Ender, "On compressive sensing applied to radar," *Signal Processing*, vol. 90, no. 5, pp. 1402–1414, 2010.
- [5] M. Cetin, I. Stojanovic, N. O. Onhon, K. R. Varshney, S. Samadi, W. C. Karl, and A. S. Willisky, "Sparsity-driven synthetic aperture radar imaging:

- Reconstruction, autofocusing, moving targets, and compressed sensing," *IEEE Signal Processing Mag.*, vol. 31, no. 4, pp. 27–40, July 2014.
- [6] M. Leigsnering, M. Amin, F. Ahmad, and A. M. Zoubir, "Multipath exploitation and suppression for SAR imaging of building interiors: An overview of recent advances," *IEEE Signal Processing Mag.*, vol. 31, no. 4, pp. 110–119, July 2014.
- [7] X. X. Zhu and R. Bamler, "Superresolving SAR tomography for multidimensional imaging of urban areas: Compressive sensing-based tomoSAR inversion," *IEEE Signal Processing Mag.*, vol. 31, no. 4, pp. 51–58, July 2014.
- [8] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–121, 2007.
- [9] T. A. Mariví, L. Paco, and J. J. Mallorquí, "A novel strategy for radar imaging based on compressive sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 12, pp. 4285–4295, 2010.
- [10] G. Li, H. Zhang, X. Wang, and X. Xia, "ISAR 2-D imaging of uniformly rotating targets via matching pursuit," *IEEE Trans. Aerospace Electron. Syst.*, vol. 48, no. 2, pp. 1838–1846, Apr. 2012.
- [11] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.
- [12] X. Zhu and R. Bamler, "Tomographic SAR inversion by L1-norm regularization: The compressive sensing approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, 2010.
- [13] V. M. Patel, G. R. Easley, D. M. Healy, Jr., and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Select. Topics Signal Processing*, vol. 4, no. 2, pp. 244–254, 2010.
- [14] N. Onhon and M. Cetin, "A sparsity-driven approach for joint SAR imaging and phase error correction," *IEEE Trans. Image Processing*, vol. 21, no. 4, pp. 2075–2088, 2012.
- [15] X. Du, C. Duan, and W. Hu, "Sparse representation based autofocusing technique for ISAR images," *IEEE Trans. Geoscience Remote Sensing*, vol. 51, no. 3, pp. 1826–1835, 2013.
- [16] I. Stojanovic and W. C. Karl, "Imaging of moving targets with multi-static SAR using an overcomplete dictionary," *IEEE J. Select. Topics Signal Processing*, vol. 4, no. 1, pp. 164–176, Feb. 2010.
- [17] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2346–2356, June 2008.
- [18] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Processing*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.
- [19] L. Zhao, G. Bi, L. Wang, and H. Zhang, "An improved auto-calibration algorithm based on sparse Bayesian learning framework," *IEEE Signal Processing Lett.*, vol. 20, no. 9, pp. 889–892, 2013.
- [20] L. Wang, L. Zhao, G. Bi, and C. Wan, "Hierarchical sparse signal recovery by variational Bayesian inference," *IEEE Signal Processing Lett.*, vol. 21, no. 1, pp. 110–113, Jan. 2014.
- [21] L. Zhao, L. Wang, G. Bi, L. Zhang, and H. Zhang, "Robust frequency-hopping spectrum estimation based on sparse Bayesian method," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 781–793, Feb. 2015.
- [22] L. Zhang, Z. Qiao, M. Xing, Y. Li, and Z. Bao, "High-resolution ISAR imaging with sparse stepped-frequency waveforms," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 11, pp. 4630–4651, Nov. 2011.
- [23] H. Wang, Y. Quan, M. Xing, and S. Zhang, "ISAR imaging via sparse probing frequencies," *IEEE Geosci. Remote Sensing Lett.*, vol. 8, no. 3, pp. 451–455, 2011.
- [24] J. Zhang, D. Zhu, and G. Zhang, "Adaptive compressed sensing radar oriented toward cognitive detection in dynamic sparse target scene," *IEEE Trans. Signal Processing*, vol. 60, no. 4, pp. 1718–1729, 2012.
- [25] Y. Huang, X. Wang, X. Li, and B. Moran, "Inverse synthetic aperture radar imaging using frame theory," *IEEE Trans. Signal Processing*, vol. 60, no. 10, pp. 5191–5200, Oct. 2012.
- [26] W. Rao, G. Li, X. Wang, and X. Xia, "Parametric sparse representation method for ISAR imaging of rotating targets," *IEEE Trans. Aerospace Electron. Syst.*, vol. 50, no. 2, pp. 910–919, Apr. 2014.
- [27] C. M. Bishop, *Pattern Recognition and Machine Learning*, vol. 1. New York: Springer, 2006.
- [28] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Trans. Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.
- [29] M. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 9, pp. 1150–1159, 2003.
- [30] N. L. Pedersen, C. N. Manchón, M. Badiu, D. Shutin, and B. H. Fleury, "Sparse estimation using Bayesian hierarchical prior modeling for real and complex linear models," *Signal Processing*, vol. 115, pp. 94–109, Oct. 2015.
- [31] N. G. Polson and J. G. Scott, "Shrink globally, act locally: Sparse Bayesian regularization and prediction," *Bayesian Statistics*, vol. 9, pp. 501–538, Dec. 2010.
- [32] C. M. Carvalho, N. G. Polson, and J. G. Scott, "The horseshoe estimator for sparse signals," *Biometrika*, vol. 97, no. 2, pp. 465–480, 2010.
- [33] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Trans. Image Processing*, vol. 19, no. 1, pp. 53–63, Jan. 2010.
- [34] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine Learning*, vol. 50, no. 1–2, pp. 5–43, 2003.
- [35] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Processing Mag.*, vol. 25, no. 6, pp. 131–146, Nov. 2008.
- [36] M. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Processing*, vol. 57, no. 6, pp. 2275–2284, 2009.
- [37] H. Liu, B. Jiu, H. Liu, and Z. Bao, "Superresolution ISAR imaging based on sparse Bayesian learning," *IEEE Trans. Geosci. Remote Sensing*, vol. 52, no. 8, pp. 5005–5013, Aug. 2014.
- [38] L. Zhang, M. Xing, C. Qiu, J. Li, J. Sheng, Y. Li, and Z. Bao, "Resolution enhancement for inversed synthetic aperture radar imaging under low SNR via improved compressive sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 10, pp. 3824–3838, 2010.
- [39] L. Wang, L. Zhao, G. Bi, C. Wan, and L. Yang, "Enhanced ISAR imaging by exploiting the continuity of the target scene," *IEEE Trans. Geosci. Remote Sensing*, vol. 52, no. 9, pp. 5736–5750, 2014.
- [40] J. Xu, Y. Pi, and Z. Cao, "Bayesian compressive-sensing in synthetic aperture radar imaging," *IET Radar, Sonar Navigation*, vol. 6, no. 1, pp. 2–8, Jan. 2012.
- [41] B. Wang, S. Zhang, and W. Wang, "Bayesian inverse synthetic aperture radar imaging by exploiting sparse probing frequencies," *IEEE Antennas Wireless Propagat. Lett.*, vol. 14, pp. 1698–1701, Apr. 2015.
- [42] M. Luessi, S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian simultaneous sparse approximation with smooth signals," *IEEE Trans. Signal Processing*, vol. 61, no. 22, pp. 5716–5729, Nov. 2013.
- [43] D.P. Wipf, "Sparse estimation with structured dictionaries," in *Proc. 24th Advances in Neural Inform. Processing Syst.*, 2011, pp. 2016–2024.
- [44] L. Wang, L. Zhao, G. Bi, and L. Zhang, "ISAR imaging by exploiting the continuity of target scene," in *Proc. 2014 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6072–6076.
- [45] Q. Wu, Y. Zhang, F. Ahmad, and M. Amin, "Compressive-sensing-based high-resolution polarimetric through-the-wall radar imaging exploiting target characteristics," *IEEE Antennas Wireless Propagat. Lett.*, vol. 14, pp. 1043–1047, May 2015.
- [46] L. Wang, L. Zhao, G. Bi, and C. Wan, "Sparse representation-based ISAR imaging using Markov random fields," *IEEE J. Select. Topics Appl. Earth Observations Remote Sensing*, vol. 8, no. 8, pp. 3941–3953, 2015.
- [47] L. Zhao, L. Wang, G. Bi, and L. Yang, "An autofocus technique for high-resolution inverse synthetic aperture radar imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 52, no. 10, pp. 6392–6403, Oct. 2014.
- [48] Q. Wu, Y. Zhang, M. Amin, and F. Ahmad, "Autofocus Bayesian compressive sensing for multipath exploitation in urban sensing," in *Proc. 2015 IEEE Int. Conf. Digital Signal Processing (DSP)*, July 2015, pp. 80–84.
- [49] L. Zhao, L. Wang, G. Bi, and C. Wan, "Structured sparsity-driven autofocus algorithm for high-resolution radar imagery," *Signal Processing*, vol. 125, pp. 376–388, Aug. 2016.
- [50] J. K. Jao and A. Yegulalp, "Multichannel synthetic aperture radar signatures and imaging of a moving target," *Inverse Problems*, vol. 29, no. 5, p. 054009, 2013.
- [51] G. Newstadt, E. Zelnio, and A. Hero, "Moving target inference with Bayesian models in SAR imagery," *IEEE Trans. Aerospace Electron. Syst.*, vol. 50, no. 3, pp. 2004–2018, 2014.
- [52] L. Yang, G. Bi, M. Xing, and L. Zhang, "Airborne SAR moving target signatures and imagery based on LVD," *IEEE Trans. Geosci. Remote Sensing*, vol. 53, no. 11, pp. 5958–5971, Nov. 2015.
- [53] X. Lv, G. Bi, C. Wan, and M. Xing, "Lv's distribution: Principle, implementation, properties, and performance," *IEEE Trans. Signal Processing*, vol. 59, no. 8, pp. 3576–3591, 2011.
- [54] L. Yang, L. Zhao, and G. Bi, "SAR ground moving target imaging algorithm based on parametric and dynamic sparse Bayesian learning," *IEEE Trans. Geosci. Remote Sensing*, vol. 54, no. 4, pp. 2254–2267, Apr. 2016.

Siavash Bayat, Yonghui Li,
Lingyang Song, and Zhu Han

Matching Theory

Applications in wireless communications

Matching theory is a powerful tool to study the formation of dynamic and mutually beneficial relations among different types of rational and selfish agents. It has been widely used to develop high performance, low complexity, and decentralized protocols. In this article, a comprehensive survey of matching theory, its variants, and their significant properties appropriate for the demands of wireless communications and network engineers is provided. Recent research progress in applying matching theory to wireless communications to address major technical opportunities and challenges is presented. A novel classification of matching models from the practical perspective is provided, and the properties and structure of each model are explained. This will enable a network designer to select an appropriate matching model for a specific application in wireless communications. Finally, the application of different matching models to various emerging wireless networks is discussed.

Introduction

Game theory offers a formal analytical framework to study the complex and dynamic interactions among interdependent selfish and rational users in a wide

Digital Object Identifier 10.1109/MSP.2016.2598848
Date of publication: 4 November 2016

1053-5888/16©2016IEEE

IEEE SIGNAL PROCESSING MAGAZINE | November 2016 |

103

number of disciplines ranging from economics, politics, and philosophy to sociology [1]–[3].

Wireless networks are made up of diverse interacting selfish and rational agents with a natural propensity to solicit their maximum benefit from the network without caring about other agents. To model such competitive behaviors of agents in networks, game theory has been widely used in open literature to facilitate autonomous network management and dynamic resource allocation. However, conventional game theory-based approaches can only deal with simple scenarios. In many complex networks, different types of agents with various characteristics and requirements want to interact with each other, and conventional game theoretical models can hardly be utilized.

As a powerful tool to study the formation of dynamic and mutually beneficial relations among different types of rational and selfish agents, matching theory [4], [5] is particularly effective in developing high performance, low complexity, decentralized, and practical solutions in these complex networks. In particular, it can effectively deal with the high dynamics of networks; selfish, competitive, and distributed nature of network elements; limited radio resources; and the dynamic quality of service (QoS) constraints of different elements.

Recently, there has been significant progress in intensive research work that uses matching theory to handle resource allocation problems in wireless networks, such as in cognitive radio (CR) networks [6], [7], heterogeneous cellular networks [8], physical layer security systems [9], distributed orthogonal frequency-division multiple access (OFDMA) networks [10], routing, and queuing systems [11], [12].

There are some short surveys in the literature that introduce matching theory [13]–[15] and its applications in wireless communications [16]. The traditional marriage problem when discussing the matching model was the focus of [13]. It introduced some concepts of the marriage model and its general properties. More definitions and interpretations of the marriage problem and discussions about its stability were provided in [15]. Signaling and data exchange in matching theory were considered in [14]. A more relevant survey that discussed matching theory in wireless communications was presented in [16]. It introduces some definitions in matching theory and its applications in wireless communications. However, these existing surveys only covered some basics and focused on specific aspects of matching theory. They did not provide a comprehensive framework of this diverse topic and major technical opportunities and challenges using matching models.

Although these works present general concepts of matching theory, some important matching models, analytical foundations of matching theory, and significant applications in wireless communications have not been covered in these articles. A variety of matching theories and models, such as matching algorithm design, modeling, matchings with transfer, assignment matching

games, convergence of the matching algorithms, stability analysis of the matchings, and optimality discussions of matching variants, have not been covered. Besides, some core concepts of matching theory for applications in wireless communications, such as competitive equilibrium, stability variants, incentive compatibility, and autonomous mechanism design, have not been discussed. Finally, the applications of matching theory in wireless and signal processing have only been briefly discussed.

This tutorial constitutes a comprehensive introduction of matching theories that will enable the communication and signal processing engineers to fully exploit the potential of matching theory. In addition to presenting the fundamental matching concepts, a new taxonomy is introduced to classify matching games. Various applications of matching games in wireless communications networks and signal processing are introduced. Corresponding matching theories, matching algorithms, and their significant properties are provided in the article.

This gives readers a clear picture of the underlying strengths and challenges of this powerful analytical tool. In summary, this article fills a void in existing communications literature by applying matching theory in communication networks through comprehensive theoretical and technical details as well as through detailed examples drawn from both game theory and wireless communications.

Matching without transfer

Utility function

In game theory, utility is a measure of motivation of a player over a set of actions. To evaluate the overall satisfaction of a player in matching games, the utility function, denoted by \mathcal{U} , is considered. It combines all the multiple related parameters to a single number to represent the net losses and gains [3]. These parameters can be of different types. Utility functions have been widely used in wireless literature to model various radio resource management problems [17], [18].

Preference list

The main goal of matching is to optimally match two sets of agents together, given their individual utilities. [Although the main matching models in the literature are between the two sets of agents, it should be noted that there are matching models that are among the agents in one set only (that are basic cooperative game models), and matching models among three sets of agents [19] known as three-dimensional matching. It is notable that three-dimensional matchings are still under study and are beyond the scope of this tutorial.] If there are two finite and disjoint sets of agents, $\Theta = \{\theta_i\}_{i=1}^{|\Theta|}$ and $\mathcal{E} = \{\xi_j\}_{j=1}^{|\mathcal{E}|}$, then each agent θ_i ranks the agents of the opposite set, \mathcal{E} , using a preference relation \succ_{θ_i} that is a complete and transitive binary relation between the set of agents of the opposite set. The notation $\xi_j \succ_{\theta_i} \xi_k$ implies agent θ_i prefers

Selfishness implies that agents compete with each other to maximize their individual utility value, without caring about another agent's utility, while rationality implies that users always make decisions that can increase their utilities.

agent ξ_j over ξ_k , and similarly, $\theta_i >_{\varepsilon_j} \theta_z$ implies that agent ξ_j prefers agent θ_i over θ_z . To put it simply, the preference of an agent over other agents can be shown by the utility value that quantifies the performance of each agent in relation with other agents.

Transfer

In the matching literature, transfer indicates any type of transaction between two different agents. The nature of transfer could be any type of entity in which the transfer affects the utility value of an agent. It can be real money, fictitious money or credit, service, or goods. For example, in some certain matching markets, the transferable parameter is real money, and through the money transfer procedure, the utility of the payer is reduced and the utility of the payee is increased. In wireless networks, a user may provide some of its resources such as power, spectrum, time slot, and so forth to another user as a transfer. Based on whether there is any transfer between agents, matching models are divided into two main categories, matching with transfer and matching without transfer.

Incentive

A common and realistic assumption in matching theory is that the agents on both sides of the matching are selfish and rational. Selfishness implies that agents compete with each other to maximize their individual utility value, without caring about another agent's utility, while rationality implies that users always make decisions that can increase their utilities. It is therefore important to always provide incentive for the agents to participate in the proposed matching algorithm by taking into account the selfish and rational nature of the agents; otherwise, the agents will not autonomously participate in the matching procedure.

Classification and definitions

The simplest matching model is the marriage problem (one-to-one matching), which was first introduced by Gale and Shapley in [20]. It is an interesting and highly practical framework that discusses the matching among men and women. In the marriage problem, men have preferences over women, and women have preferences over men. The outcome of the marriage problem needs to be a set of marriages such that there are no two people of opposite sex who would both prefer each other over their current partners. In other words, the marriages need to be stable.

In addition to the classical one-to-one matching, in reality there are many practical scenarios in which the agents in

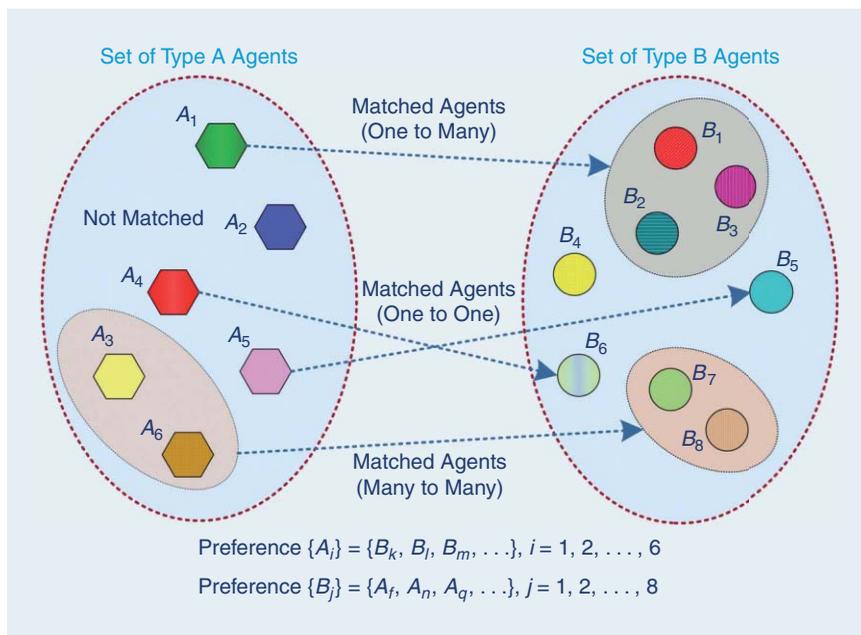


FIGURE 1. The general configuration of the matching structure.

one side of the matching are allowed to be matched with a number of other agents from the other side of the matching (one-to-many matching), such as college admission where the students are admitted to a college or cellular user (CU) associations where the mobile users are allocated to a base station (BS). In these scenarios, multiple students/users can be allocated to each college/BS. In this matching game, the concept of group stability is introduced to ensure a stable matching is achieved. Finally, if the number of the allowable matches for both sides of the matching is unrestricted, then it becomes a many-to-many matching problem. In many-to-many matching, the notion of pairwise stability is introduced. Here the *without transfer* indicates that there is not any kind of transactions or transfers between the agents in the matching. In the sequel, different types of matching are discussed in detail. Figure 1 shows a general matching structure that demonstrates three types of matching configurations, one-to-one, one-to-many, and many-to-many matchings. For example, agent A_4 from the set of type A agents is matched to $\{B_6\}$ from the set of type B agents and forms a one-to-one matching configuration. Agent A_1 from the set of type A agents is also matched to $\{B_1, B_2, B_3\}$ from the set of type B agents and forms a one-to-many matching configuration. Some agents also are not matched. More details about these matching types will be discussed in the following sections.

One-to-one matching

In the classical marriage problem, there are two sets of agents, $\mathcal{M} = \{m_i\}_{i=1}^M$ and $\mathcal{W} = \{w_j\}_{j=1}^N$, which are called *men* and *women*, respectively, where each agent has ranked all members of the other set by a unique preference number. The outcome of the marriage problem is a one-to-one

Table 1. The Gale–Shapely algorithm.
Step 1: Initialization

- 1) Set $\mathcal{M} = \{m_i\}_{i=1}^M$, $\mathcal{W} = \{w_j\}_{j=1}^N$.
- 2) Construct m_i^{list} and w_j^{list} .
- 3) Construct the list of all men that are not matched, denoted by $\mathcal{M}_{\text{MatchList}}$.

Step 2: Dynamic Proposals

- 1) Each $m_i \in \mathcal{M}_{\text{MatchList}}$ makes a proposal to w_j that is the first in its preference list.
 - If the ranking of proposing m_i is moving higher in w_j 's list than w_j 's current partner then
 - w_j accepts m_i 's proposal and rejects her current partner m_i .
 - m_i will be removed from $\mathcal{M}_{\text{MatchList}}$, and m_i will be added to $\mathcal{M}_{\text{MatchList}}$.
 - Else if w_j does not prefer the proposing m_i to her current partner, m_i , then
 - rejects m_i 's proposal and holds her current partner m_i .
 - m_i updates his preference list, m_i^{list} , by removing w_j .

Step 3: End

- 1) If m_i^{list} is not empty, go to step 2; otherwise, the algorithm ends.

matching of men and women. The one-to-one matching is denoted by [5].

■ **Definition 1:** Given two disjoint sets \mathcal{M} and \mathcal{W} , a one-to-one matching, μ , is defined as an allocation from $\mathcal{M} \cup \mathcal{W}$ to $\mathcal{M} \cup \mathcal{W}$ such that if $\mu(m) \neq m$, then $\mu(m) \in \mathcal{W}$ and if $\mu(w) \neq w$, then $\mu(w) \in \mathcal{M}$. The partner of w is referred to as $\mu(w)$ if $\mu(w) = m$.

Note that $\mu(m) = m$ implies that man m is matched to itself, which means it is not really matched. Besides, $\mu(m) = w$ implies that if man m is matched to woman w , then woman w is also matched to man m ($\mu(w) = m$). An important property of each matching is how the matchings will change during the time. An obvious question is, what happens when the agents realize that they can be matched with better partners and achieve a higher utility value—specifically, how stable will a matching behave? The concept of stability for a marriage problem matching is stated as follows:

■ **Definition 2:** If there is not any couple comprising one man and one woman who both prefer each other over their current partners, all the marriages are called *stable*.

To implement a stable matching between two sets of agents, matching algorithms that converge to stable outcomes should be used. A well-known algorithm, which leads to a stable matching in the marriage problem, is the Gale–Shapely algorithm that is shown in Table 1. The original Gale–Shapely algorithm for a traditional marriage problem was presented in [20]. In this tutorial, we would like to present the Gale–Shapely algorithm in a more general form. The Gale–Shapely algorithm is expressed by a series of proposals from one side of the matching to the other side.

In the Gale–Shapely algorithm, both women and men can propose. When the algorithm is executed, at any instance each person is either engaged or single. When the persons of gender A are proposing, each member of gender B may change

between engaged and single status. Because the persons of the gender B are only accepting or rejecting the offers, once a member of gender B is engaged, he/she will not go back to single status again, although his/her partner may change. This is intuitive, as some gender A persons could be rejected by some gender B persons and become available for engagement with other gender B persons. If a person of gender A is rejected by a person of gender B and then becomes engaged to another partner of gender B, his/her new partner is less desirable to him/her.

When an available person of gender B receives an offer, he/she will immediately accept it and become engaged to the first proposer. When an engaged person of gender B receives another offer, he/she compares the second proposer with his/her current partner and rejects the less preferred person of gender A. To facilitate the description, in the following, gender A is assumed as male and gender B is assumed as female. This can also be assumed oppositely. In the marriage matching game, each man proposes to the women based on the order of his preference list, until he becomes engaged. If his engagement is broken by a woman, then he becomes available again. He continues his proposals to the next woman on his list. The algorithm terminates when no new proposal is being made. Furthermore, it is shown by Theorem 1 that, on termination, the engaged couples constitute a stable matching [21]:

■ **Theorem 1:** At any instance in the marriage problem when the Gale–Shapely algorithm terminates, the matched pairs form a stable matching.

In a marriage problem when the men/women are the proposers, there is a very interesting property in the result of the matching. Every man/woman receives the best partner that he/she can possibly have in any stable matching. This means that all the men/women who are competing with each other to get matched with their most preferred woman/man can agree on a stable matching that is simultaneously optimal for all of them. This result is represented in [21]:

■ **Theorem 2:** The outcome of all possible scenarios of the Gale–Shapely algorithm results in the same stable matching. In the resulted stable matching, each man/woman gets matched with the best partner that he/she can have in any stable matching.

This theorem states that if each man/woman is given his/her best stable partner, then the result of that matching is a stable matching. The men/women stable matching resulting from the Gale–Shapely algorithm is called *men/women optimal*.

There is a possibility that the man/woman optimal stable matchings are identical in a matching, but this is not a general case. It is notable that if a matching is optimal for the members of one gender, it is the worst matching for the members of the opposite gender. Specifically, in the man-optimal/women-optimal matching, each woman/man has the worst partner that she/he can have in any stable matching, and this will be valid for men/women in the woman/man optimal stable matching.

One-to-many matching

There are many practical scenarios in which the agents from one side of the matching are allowed to be matched with a number of other agents from the other side of the matching, such as when students are allocated to a college or when doctors are allocated to a hospital. One-to-many matching indicates that each student can be matched with one college, although there could be multiple colleges. To proceed with explaining one-to-many matching, a well-known example in matching literature called *college admission* is used. Let's assume there are two finite and disjoint sets, $\Theta = \{\theta_i\}_{i=1}^{|\Theta|}$ and $\mathcal{E} = \{\xi_j\}_{j=1}^{|\mathcal{E}|}$, which represent the set of colleges and students, respectively. Each student has preferences over each college, and each college has preferences over each student.

The difference between the college admission and the marriage model is that associated with each college θ there is a positive integer, $q_\theta \in \mathbb{N}$, called its *quota*, which indicates the maximum number of positions the college may fill. An outcome of the college admission problem is a matching of students to colleges such that each student is matched to at most one college, and each college is matched to at most its quota of students. A student who is not matched to any college will be matched to himself/herself as in the marriage model, and a college that has a certain number of unfilled positions will be matched to itself in each of the unfilled positions. It is notable that matching is bilateral, in the sense that a student is admitted at a given college if and only if the college admits that student. This matching is stated formally as

- **Definition 3:** Given two disjoint finite sets of players, $\Theta = \{\theta_i\}_{i=1}^{|\Theta|}$ and $\mathcal{E} = \{\xi_j\}_{j=1}^{|\mathcal{E}|}$, let two disjoint finite sets of $\phi^\Theta = \{1, \dots, |\Theta|\}$ and $\phi^\mathcal{E} = \{1, \dots, |\mathcal{E}|\}$, then a one-to-many matching function $\Phi: \{\Theta\} \cup \{\mathcal{E}\} \rightarrow \{\Theta\} \cup \{\mathcal{E}\}$ is defined such that for all $i \in \phi^\Theta$ and $j \in \phi^\mathcal{E}$
 - 1) $\Phi(\theta_i) \subseteq \{\xi_{j \in \phi^\mathcal{E}} \in \mathcal{E}\}$ and $|\Phi(\theta_i)| \leq q_\theta$
 - 2) $\Phi(\xi_j) \in \{\theta_{i \in \phi^\Theta} \in \Theta\}$ and $|\Phi(\xi_j)| \in \{0, 1\}$
 - 3) $\Phi(\theta_i) = \xi_j \Leftrightarrow \Phi(\xi_j) = \theta_i$.

This matching function Φ is denoted with a three-tuple $\Phi: (\Theta, \mathcal{E}, q)$. Condition 1 implies that each member of Θ can be matched to multiple members of \mathcal{E} , condition 2 implies that each member of \mathcal{E} can be matched to at most one member of Θ , and condition 3 implies that if θ_i is matched to ξ_j , then ξ_j is also matched to θ_i . For notational purposes, θ_0 and ξ_0 are defined as the dummy members of Θ and \mathcal{E} , respectively. The dummy members θ_0 and ξ_0 can be matched to multiple members in \mathcal{E} and Θ , respectively.

To formally define a stable matching, let's first define a matching that is blocked by an individual and a matching that is blocked by a pair. The parameters $\theta_i, i \in \phi^\Theta$ and $\xi_j, j \in \phi^\mathcal{E}$ are considered, such that $\Phi(\theta_i) \neq \xi_j$. A matching Φ is blocked by an individual θ_i (ξ_j) if θ_i (ξ_j) prefers to be unmatched over being matched with its current partner under Φ [21]. Mathematically, because the agents have to receive nonnegative utilities, for θ_i , this implies that $\mathcal{U}_{\theta_i}(\Phi_1(i)) < 0$, while for ξ_j , this implies that

$\mathcal{U}_{\xi_{\Phi_1(j)}}(\Phi_1(j)) < 0$, where $\Phi_1(i) \triangleq j$ when $\Phi(\theta_i) = \xi_j$ and $\Phi_1(j) \triangleq i$ when $\Phi(\xi_j) = \theta_i$.

A matching Φ is blocked by a pair (θ_i, ξ_j) if 1) the matching is not blocked by individual θ_i and ξ_j , and 2) both θ_i and ξ_j can achieve a higher utility value if they match together, as opposed to their current matching under Φ [21]. This implies that $\mathcal{U}_{\xi_j}(\theta_i) > \mathcal{U}_{\xi_j}(\Phi_1(j))$ and $\mathcal{U}_{\theta_i}(\xi_j) > \mathcal{U}_{\theta_i}(\Phi_1(i))$. The stability of a matching can be defined as

- **Definition 4:** A matching Φ is defined as stable if it is not blocked by any individual or any pair.

It should be noted that to achieve a stable matching, there is no need for $|\mathcal{E}| = |\Theta|$. In one-to-many matching where a member of Θ is matched with a group of \mathcal{E} members, the stability is different from one-to-one matching. For the one-to-many matchings the concept of group stability is used. To define group stability, a coalition C that consists of at least one member of Θ is assumed first. A matching Φ is blocked by a coalition C if there exists another matching Φ' such that $\forall i, j \in C$, 1) $\Phi'(j) \in C$; 2) $\mathcal{U}_{\xi_{\Phi'(j)}}(\Phi'(j)) > \mathcal{U}_{\xi_{\Phi(j)}}(\Phi(j))$; 3) $\mathcal{U}_{\theta_i}(\Phi'(i)) > \mathcal{U}_{\theta_i}(\Phi(i))$; and 4) if $u \in \Phi'(\phi^\Theta)$, then $u \in C \cup \Phi(\phi^\Theta)$.

Condition 1 states that all the members of Θ in C are matched to a member of Θ in C , conditions 2 and 3 state that all the members of Θ and members of \mathcal{E} in C prefer their current matches in Φ' to their matches in Φ , and condition 4 states that every member of Θ in C can be matched to a combination of new member of Θ in C or the member of Θ that was matched under Φ . Therefore, Φ is blocked by some coalition C of the member of Θ and \mathcal{E} , if the member of Θ and the member of \mathcal{E} in C all find a matching preferable to Φ . Given the above conditions group stability is defined as

- **Definition 5:** A matching Φ is defined as group stable if it is not blocked by any coalition.

Many-to-many matching

If the number of allowable matches for the agents in both sides of the matching is unrestricted, it is a many-to-many matching problem.

- **Definition 6:** Given two disjoint finite sets of players, $\Theta = \{\theta_i\}_{i=1}^{|\Theta|}$ and $\mathcal{E} = \{\xi_j\}_{j=1}^{|\mathcal{E}|}$, let two disjoint finite sets of $\phi^\Theta = \{1, \dots, |\Theta|\}$ and $\phi^\mathcal{E} = \{1, \dots, |\mathcal{E}|\}$, then a many-to-many matching function is defined $\Phi: \{\Theta\} \cup \{\mathcal{E}\} \rightarrow \{\Theta\} \cup \{\mathcal{E}\}$ such that for all $i \in \phi^\Theta$ and $j \in \phi^\mathcal{E}$
 - 1) $\mu(\xi_j)$ is contained in Θ and $\mu(\theta_i)$ is contained in \mathcal{E}
 - 2) $\Phi(\theta_i) \subseteq \{\xi_{j \in \phi^\mathcal{E}} \in \mathcal{E}\}$ and $|\Phi(\theta_i)| \leq q_\theta$
 - 3) $\Phi(\xi_j) \in \{\theta_{i \in \phi^\Theta} \in \Theta\}$ and $|\Phi(\xi_j)| \leq q_\xi$
 - 4) $\Phi(\theta_i) = \xi_j \Leftrightarrow \Phi(\xi_j) = \theta_i$,

where $q_\xi \in \mathbb{N}$ denotes the quota of \mathcal{E} members.

In many-to-many matching models, many stability concepts can be considered depending on the number of players who can improve their utility by matching to new partners. However, for a large number of agents, it is difficult to identify large coalitions by considering all the possible combinations of players. Therefore, the notion of pairwise stability is presented and used in many-to-many matching

in the following way [22]. To explain the pairwise stability concept, the student–college example is used. In a pairwise stable matching, the matching is not blocked by an individual student, or an individual college, or a student–school pair. However, 1) each student θ_i prefers his/her matched college, rather being unmatched; 2) each college prefers not rejecting some of the matched students; and 3) there is no student–college pair who are not matched but prefer to be matched together.

Matching with transfer—assignment game

Each matching structure (one to one, one to many, and many to many) can be further classified into a matching with transfer or a matching without transfer depending on if there is any transfer between the matching sides. Matchings with transfer are very common in the real world. A simple example is the buyer–seller scenario where buyers pay sellers in exchange for goods. After the transactions between the sellers and buyers end, a matching outcome is resulted. In wireless networks, the transferred item can be the network's resources, such as subchannels, power, time slots, and so forth. One applicable and general matching framework with transfer is the assignment game that discusses the matching and transactions among buyers and sellers. The assignment game has different variants used for modeling various complex scenarios in wireless networks, for example, the scenario where primary users (PUs) and secondary users (SUs) negotiate on spectrum sharing in CR networks. When the items that the sellers are selling are of different types [one has an apple, another has an orange, etc. (simple and more tangible examples are used here to simplify understanding the differences of the matching models)], this is called the *multiple objects assignment game*. If each seller has a number of items of the same type (e.g., seller one has three apples, seller two has three apples, etc.), this is called the *multiple unit assignment game*. If each seller sells a combination of different objects (e.g., seller one has two apples, two oranges, zero bananas), this becomes a heterogeneous multiple unit assignment game. The problem becomes more interesting if each buyer also needs a different combination of different objects. The buyers also can be limited in the amount of money they can spend. To solve different variants of assignment games, iterative and dynamic matching algorithms can be developed such that the assignment game converges to the competitive equilibrium outcome. In the competitive equilibrium, there are no two agents who can form a matching pair in such a way that benefits both of them more than that of their current status, and also there is no matched agent who may prefer to be unmatched. Competitive equilibrium is a key concept in matchings with transfer and will be discussed in more detail in the following sections. Because in the matching with transfers there is some sort of transaction among the sides of the matching, sometimes in the literature these matchings are referred to as *matching markets*.

General assignment game model

In the general assignment game model, there could be multiple agents in both sides of the matching, and agents from one side have transactions with agents in the opposite side. A well-known example for the assignment game that clearly reveals the details of the matching procedure is the firms–workers market in which the firms would like to find the best-matched workers [23].

To model this market it is assumed that there are two finite disjoint sets of players, workers, denoted by $\mathcal{P} = \{p_1, p_2, \dots, p_m\}$, and firms, denoted by $\mathcal{S} = \{s_1, s_2, \dots, s_n\}$, containing m and n players, respectively. Associated with each possible partnership (p_i, s_j) in $(\mathcal{P} \times \mathcal{Q})$ there is a nonnegative real number α_{ij} that represents the gain of this matching.

If each worker has a reservation price of zero for working in a firm, then r_{ij} represents firm p_i 's reservation price for the service offered by worker s_j to firm p_i . In this case if firm p_i employs worker s_j with a salary of β_{ij} , and if no other monetary transfers are made or received by firm p_i and worker s_j , then the resulting utilities of the two agents are $u_i = r_{ij} - \beta_{ij}$ and $v_j = \beta_{ij}$, respectively.

More generally, if each worker s_j has a reservation price c_j and each firm p_i has a reservation price r_{ij} for the service by worker j , α_{ij} is assumed to be the potential gains from trade between p_i and s_j ; that is, $\alpha_{ij} = \max[0, r_{ij} - c_j]$. In this case if firm p_i pays for service provided from worker s_j at a salary β_{ij} , and if no other monetary transfers are made, the utilities of firm p_i and worker s_j are $u_i = r_{ij} - \beta_{ij}$ and $v_j = \beta_{ij} - c_j$, respectively. For technical convenience, one virtual worker, s_0 , and one virtual firm, p_0 , are introduced. Several firms may be matched with the virtual worker and several workers may be matched with the virtual firm.

It should be noted that in wireless networks there are cases where the two sets of agents do not really trade a good, instead they exchange services. For example, a set of agents share their resources to other agents. In a similar way these two kinds of agents in this scenario can be regarded as firms and workers. To facilitate the mathematical notations, for all general types of matching, the matching index x_{ij} is defined as

$$x_{ij} = \begin{cases} 1, & \text{if } \Phi(\xi_j) = \theta_i \text{ for } i \in \phi^\theta \cup 0, j \in \phi^\varepsilon \cup 0, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The parameter X is defined as an $|\Theta| \times |\mathcal{E}|$ matching matrix with the (i, j) th element denoted by x_{ij} . In the matchings without transfer there is no transaction between the two sides of the matching, so the preferences of the members in one side over the members of the other side do not change. To generalize the model, in the matching literature the agents can also be defined as buyers and sellers. For example, in the firms–workers model, the firms who employ the workers can be regarded as the buyers who are purchasing the workers' services. The workers who are providing services to the firms are assumed as sellers who are selling the services to firms. It should be mentioned that the buyers–sellers model

in matching theory is essentially different from the normal auction models.

Normally, in matching, the aim is to form stable matching pairs between the firms and workers. However, in the auctions, the aim is to sell objects to the buyers. Moreover, in more complex matching models, there are multiple buyers and multiple sellers, and each seller has a different object to sell. Each buyer also has a specific preference over each object. However, in the auction models normally one seller auctions one object to multiple buyers. In some more general auctions, multiple sellers auction the same objects to the buyers. In matching theory, the multiple seller–multiple buyer case is traditionally classified as an assignment game.

Optimization problem

The maximization problem to determine the matching in the assignment game is called the *optimal assignment problem* or simply the *assignment problem*. In this part, a one-to-one assignment problem is formulated as follows:

$$\begin{aligned} & \max \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} x_{ij} \\ & \text{s.t. : 1) } \sum_{i=1}^m x_{ij} \leq 1, \forall j \in \{1, 2, \dots, n\} \\ & \quad 2) \sum_{j=1}^n x_{ij} \leq 1, \forall i \in \{1, 2, \dots, m\} \\ & \quad 3) x_{ij} \in \{0, 1\}, \forall i \in \{1, 2, \dots, m\} \text{ and } \forall j \in \{1, 2, \dots, n\}. \end{aligned} \quad (2)$$

Condition 1 guarantees that each buyer will be matched with only one seller, condition 2 states that each seller can be matched with only one buyer, and condition 3 guarantees that the value of x will be one or zero. The optimization problem in (2) is a binary linear programming problem (BLP), which is a special case of integer programming in which the variables can only be zero or one. The optimization problem (2) is also in the standard form of BLP problems with the set of inequality conditions on the binary variables that is shown to be NP-hard. It has been shown that there exists a solution of this optimization problem for which values of x will be one or zero. A matrix X is a feasible solution for the assignment problem in (2), if it satisfies conditions 1–3. Moreover, among all the feasible solutions, the solution X is called the *optimal solution* if

$$\sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} x_{ij} > \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} x'_{ij},$$

where x'_{ij} belongs to another solution X' .

A feasible outcome X is stable if the two following conditions are satisfied:

- 1) $u_i \geq 0, v_i \geq 0$
- 2) $u_i + v_i \geq \alpha_{ij}, \forall i, j$.

Condition 1 is called the *individual rationality condition* and reflects that a player may remain unmatched. Condition 2 requires that the outcome is not blocked by any pair. If condition 2 is not satisfied for some agents i and j , then they will break up their present partnership and form a new partnership, because both can receive higher utility value.

General distributed assignment game algorithm

The general assignment game algorithm (GA-*alg*) that is a highly practical model is discussed in this section. The purpose of the proposed algorithm is to obtain a solution to the optimization problem in (2) in a distributive way. The algorithm in Table 2 is a general algorithm and can be customized for a wide range of complex scenarios in matching games.

The specific details of the algorithm are described in Table 2, with the notations defined below. The GA-*alg* consists of an initialization stage in step 1, followed by multiple iterations. At this part a brief description of the algorithm during iteration t is given.

In GA-*alg*-Step 2, each seller that is not matched makes a price-allocation number offer to the unmatched buyers (GA-*alg*-Step 2-1). This price-allocation number is denoted by $\beta_{i,j}^t$ for the j th seller and i th buyer. Each buyer then determines the seller that provides the highest positive utility (GA-*alg*-Step 2-2-a). For the i th buyer, this is denoted by the demand set [the term *demand set* is used as it is also used in economics literature (see, e.g., [5] and [24])]:

$$\Omega_i(b_i^t) = \begin{cases} \operatorname{argmax}_{j \in S} U_{i,j}^p, & \text{if } \max_{j \in S} U_{i,j}^p \geq 0; \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where $b_i^t = [\beta_{1,i}^t, \beta_{2,i}^t, \dots, \beta_{n,i}^t]$ is the price-allocation vector at stage t and $U_{i,j}^p$ is the utility function of buyer p_i when it is matched with seller s_j . The demand set for p_i is equal to the index of the seller that provides the maximum utility. If the demand set is not empty, the buyer will bid for the seller in its demand set; otherwise, the buyer will not bid and remains unmatched. The bid from p_i to s_j is denoted by $g_{i,j}^t$, with $g_{i,j}^t$ taking on the values

$$g_{i,j}^t = \begin{cases} 1, & \text{if } p_i \text{ bids for } s_j; \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

In GA-*alg*-Step 3, sellers will decide if they want to match with the buyers. There are three possibilities. The first possibility (GA-*alg*-Step 3-1) is that the seller receives no bids from buyers after increasing its price-allocation number, but in the previous iteration, the seller had received bids from multiple buyers. The seller will choose one of these buyers randomly to be matched with. The second possibility is that the seller has bids from multiple buyers (GA-*alg*-Step 3-2-a). The seller will increase its price-allocation number by ϵ for the next iteration, that, the $t + 1$ st iteration. Note that $\epsilon \in \mathbb{R}^+$ represents the price-step number, which indicates the price-allocation number offer from the sellers to the source increases at each step. The third possibility (GA-*alg*-Step 3-2-b) is that the seller has

Table 2. The general assignment game algorithm (GA-alg).
Step 1: Initialization

- 1) Set $t = 1$, $\beta_{i,j}^t = \beta_{\min,i}$, $\forall i \in \mathcal{P}, \forall j \in \mathcal{S}$.
- 2) Set $g_{i,j}^t = 0$, $\forall i \in \mathcal{P}, \forall j \in \mathcal{S}$.
- 3) Construct the list of all the sellers that are not matched denoted by $\text{MATCHLIST}_s = \{s_j\}_{j=1}^n$ (implying $\Phi^t(s_j) = (p_0, 0)$, $\forall s_j \in \mathcal{S}$).

Step 2: Buyer Demand of Goods or Service

- 1) Each $s_j \in \text{MATCHLIST}_s$ announces its price-allocation number $\beta_{i,j}^t$ to all the unmatched buyers.
- 2) Set $\text{SLIST}_p = \emptyset$.
- 3) Each unmatched buyer p_i , $\forall i \in \mathcal{P}$:
 - Determines its demand $\Omega_i(b^t)$.
 - If $\Omega_i(b^t) \neq \emptyset$, then p_i is added to SLIST_p and bids for s_j ($g_{i,j}^t = 1$), where $j = \Omega_i(b^t)$, $\forall s_j \in \mathcal{S}$.
 - Otherwise, p_i does not bid, ($g_{i,j}^t = 0$, $j \in \mathcal{S}$) and $\Phi^t(p_i) = (s_0, 0)$.

Step 3: Sellers' Decision Making

- 1) For all $s_j \in \mathcal{S}$
 - If $\Phi^t(s_j) = p_0$ and $\sum_{i=1}^m g_{i,j}^t = 0$ and $\beta_{i,j}^t > \beta_{\min,i}$, $\forall i \in \mathcal{P}$
 - $\Phi^t(s_j) = (p_{i^*}, \beta_{i^*,j}^{t-1})$, where $i^* = \text{random}(\{i^* \mid g_{i^*,j}^{t-1} = 1, \forall i^* \in \mathcal{P}\})$.
 - Remove s_j from MATCHLIST_s .
 - Set $g_{i^*,j}^t = 0$, where $j^* = \Omega_{i^*}(b^t)$.
- 2) For all $s_j \in \mathcal{S}$
 - If $\sum_{i=1}^m g_{i,j}^t > 1$
 - Set $\beta_{i,j}^{t+1} = \beta_{i,j}^t + \epsilon$.
 - If $s_j \notin \text{MATCHLIST}_s$
 - s_j is added to MATCHLIST_s .
 - Else $\sum_{i=1}^m g_{i,j}^t = 1$
 - $\Phi^t(s_j) = (p_i, \beta_{i,j}^t)$, where $i = \arg \{i \in \mathcal{P} \mid g_{i,j}^t = 1\}$.
 - Remove s_j from MATCHLIST_s .
- 3) For all $s_j \in \mathcal{S}$
 - If $(\Phi^t(s_j) \neq p_0$ and $\sum_{i=1}^m g_{i,j}^t = 0)$ or $\Phi^t(s_j) = (p_0, 0)$
 - $\beta_{i,j}^{t+1} = \beta_{i,j}^t$, $i \in \mathcal{P}$.
 - Set $t = t + 1$. If $\text{SLIST}_p \neq \emptyset$, go to step 2; otherwise, go to step 4.

Step 4: End of the Algorithm

received a bid from only one buyer. In this case, the seller will be matched with this node. The last iteration of the GA-alg is denote by t_{End} .

Competitive equilibrium

The requirements of a stable matching in previous sections are defined with the general implication that none of the two matched agents has a desire to leave their current matching and make a new match. In the wireless networks, if users are assumed as the agents in matching then the focus of the network is to maximize the agents' satisfaction level. Although the final matching may be stable, it is an important factor to investigate whether any of the matched buyers prefer to be matched with another seller.

In this regard, it is convenient to present the competitive equilibrium, defined as

■ **Definition 7:** The matching matrix X , and price-allocation vector b , that are produced by the matching algorithms are in competitive equilibrium if the following conditions are satisfied:

- 1) for each $s_j, j \in \mathcal{S}$, if $\Phi_1(s_j) \neq p_0$, then $\beta_{ij} \geq c_j$
- 2) for each $p_i, i \in \mathcal{P}$, $\Phi_1(p_i) \in \Omega_i(b_i)$
- 3) for each $s_j, j \in \mathcal{S}$, if $\Phi_1(s_j) = p_0$, then $\beta_{ij} = c_j$, where $\Omega_i(b_i)$ is defined in a similar way to (3) but dropping parameter t .

The conditions can be interpreted as follows. Condition 1 states that a matched seller always receives a nonnegative utility; condition 2 states that each buyer will be matched with the seller that provides the highest positive utility; and condition 3 states that if a seller is not matched, then its price-allocation number results in a zero utility. Note that based on Definition 7, the competitive equilibrium concept is a special case of the stable matchings.

However, the concept of competitive equilibrium needs to be distinguished from stable matching. In stable matching there are no buyer–seller pairs who both prefer each other over their current match, but in competitive equilibrium there is no buyer who prefers any seller over his/her current match. Competitive equilibrium is a feasible matching outcome that guarantees that 1) there are no two agents who are able to form a matching pair such that both would benefit better than their current state, and 2) there is no matched agent who prefers to be unmatched. Competitive equilibrium has been widely studied in the literature [25], [26]. It has been shown that a competitive equilibrium always exists [26]. A competitive equilibrium can be stated by the matching matrix and the price vector. Based on this definition, the following theorem and lemma are presented [8].

■ **Theorem 3:** The algorithm in Table 2 produces matching and price-allocation matrices, which are in competitive equilibrium for sufficiently small values of ϵ .

Theorem 3 provides an important property of the matching algorithm, that when ϵ is small enough, the outcome of the algorithm will be in the competitive equilibrium. Decreasing ϵ will increase the number of iterations of the matching algorithm in Table 2, so there is always a tradeoff between the optimality and the complexity of the matching algorithm in Table 2. The parameter ϵ can be considered as a controlling parameter for the network designer to balance between the convergence time and performance of the matching algorithm.

■ **Lemma 1:** An algorithm that results in a competitive equilibrium also results in a stable matching.

Lemma 1 implies that competitive equilibrium covers the stable matching. However, there are some outcomes of the matching that are not competitive equilibrium.

Matching with externalities

In all the matching models discussed previously, the preferences of agents over each other is fixed over time. In such a

conventional model of two-sided matchings, the preference of each agent only depends on with whom the agent is being matched. This matching is called *matching without externalities*. It means that the agents do not care about whom the other agents are going to match with. However, there are scenarios where it is important for an agent to know who is matched to other agents because they may share the same resources. This matching is called *matching with externalities*. For example, let's consider a user subchannel matching in an ad hoc network with subchannels can be accessed by multiple users. At the beginning user A may choose subchannel C as its most preferred subchannel. However, when the network gets congested with more users allocated to subchannel C, the interference level in this subchannel increases, and user A may change its most preferred subchannel. In other words, the preference of user A over subchannel C depends on the choices of other users. In another example, in the college admission problem, a student may not only care about the quality of the college that he/she is going to apply for but also who else is applying for the same college.

A matching game with externalities is defined as [26]:

■ **Definition 8:** A matching game with externalities is represented as a tuple $G = (\Theta, \mathcal{E}, \mathcal{U})$, where (Θ, \mathcal{E}) is the set of agents and \mathcal{U} is a real valued function such that $\mathcal{U}(z | \eta)$ is the utility of agent z when matching η forms.

In the assignment game [13], $\mathcal{U}(\theta_i, \xi_i)$ denotes the utility value of θ_i when θ_i and ξ_i are matched. Now with externalities, a generalized assignment game that includes externalities is assumed such that for each matching η and every $(\theta_i, \xi_j) \in \eta$, the amount $\mathcal{U}(\theta_i, \xi_j | \eta)$ denotes the utility from matching of the pair (θ_i, ξ_j) at A. This matching is a nontransferable utility game, where the utility of each agent depends on the underlying coalition structure. Definition 8 represents the most general form of a matching game with externalities when utility is nontransferable:

■ **Definition 9:** A matching game with additive externalities is represented as a tuple $G = (\Theta, \mathcal{E}, \mathcal{U})$, where (Θ, \mathcal{E}) is the set of agents and \mathcal{U} is a utility function such that $\mathcal{U}(z | \theta_i, \xi_j)$ is the utility that agent z receives from the formation of pair (θ_i, ξ_j) . Given a matching η , the utility of agent z in η is

$$\mathcal{U}(z, \eta) = \sum_{(\theta_i, \xi_j) \in \eta} \mathcal{U}(\theta_i, \xi_j | z).$$

In a matching game with additive externalities, the agent's total utility is the sum of all utilities the agent gains from its own matches, plus the sum of all the externalities that the agent will gain from the matchings of other agents. It is interesting to know whether or not there exists any stable matching for a given matching game with externalities and to know how to find such stable matching if it exists. Like the general matching definition, a matching is stable if there is no subset of agents that have incentive to reorganize and form new matchings among themselves [27].

Given a matching game $G = (\Theta, \mathcal{E}, \mathcal{U})$, a matching η is corewise stable if there is no blocking coalition $\mathcal{B} \subseteq \Theta \cup \mathcal{E}$, such that at least one agent in \mathcal{B} can receive a higher utility value and no agent in \mathcal{B} receives a lower utility value by forming new matches among themselves.

In the matching without externalities, the actions chosen by other agents do not have any effect on other agents because their utility depends only on whom the agent is matched with. However, in matchings with externalities the utility of the agents does not only depend on an agent's own action. For example, the utility of agent θ_i , may depend on the other matches involving agent ξ_j even if $(\theta_i, \xi_j) \notin \eta$. Therefore, it is clear that the stability concepts need to take into account the actions chosen by other agents in $\mathcal{K} = \mathcal{H} \setminus \mathcal{B}$, $\mathcal{H} = \Theta \cup \mathcal{E}$ after a coalition \mathcal{B} deviates. This is because these actions will affect the utility value of the agents in \mathcal{B} . However, computing all the possible reactions of other agents that cause deviation is very complex because there might be a big number of possible combinations.

Matchings with incomplete preference lists and with ties

Another variant of matching is when an agent's preference list is incomplete, that is, an agent may exclude some agents that it does not want to be matched with from its preference list. In this case, the agent θ_i is acceptable to agent ξ_j if θ_i appears on the preference list of ξ_j and unacceptable otherwise. This matching is stable if there are no two agents that prefer the other to its current partner, or if one agent is unmatched, the other agent has been unacceptable for it.

In some cases, ties are allowed in the preference lists, that is, one agent can include two or more agents with the same preference value as a tie. For this type of matching, three notions of stability are defined: weakly stable, strongly stable, and superstable. A more complex extension in matching that also can happen in wireless networks is to allow both incompleteness and ties in preference lists. In the following variants, the marriage problem is used as an example to clarify the matching models.

Matching with incomplete preference lists

In this type of matching game, each agent's preference list could be incomplete. In other words, some options that the agent does not want to be matched with can be excluded from its preference list.

If an agent θ_i 's list includes an agent ξ_j , then ξ_j is acceptable to θ_i . When there are incomplete lists, a matching is not necessarily perfect. Hence, in matchings with incomplete lists the definition of a blocking pair is extended. As was mentioned earlier for a matching μ , (θ_i, ξ_j) is a blocking pair if all the following three conditions are met:

- 1) $\mu(\theta_i) \neq \xi_j$, but θ_i and ξ_j are acceptable to each other
- 2) $\xi_j \succ_{\theta_i} \mu(\theta_i)$ or θ_i is single
- 3) $\theta_i \succ_{\xi_j} \mu(\xi_j)$ or ξ_j is single.

An important property of stable matching with an incomplete list is that, for example, in the marriage

problem, after the matching, the set of men (women) can be partitioned into two different sets; one is the set of men (women) who have partners in all stable matchings, and the other is the set of men (women) who are single in all stable matchings [13]. This implies that all stable matchings for a single instance are of the same size. Furthermore, the Gale–Shapley algorithm with a slight modification can be applied to find a stable matching in the matchings with incomplete lists.

Matching with preference lists with ties

Another variant of matching games is when agents have ties in their preference lists. In other words, one agent can include two or more agents with the same preference in a tie. If agent ξ_k and ξ_z are in the same tie of agent θ_i 's preference list, this is denoted by $\xi_k =_{\theta_i} \xi_z$. $\xi_k \succeq_{\theta_i} \xi_z$ means $\xi_k \succ_{\theta_i} \xi_z$ or $\xi_k =_{\theta_i} \xi_z$.

In stable matching with ties, there are three stability concepts, superstability, strong stability, and weak stability. In superstability, a blocking pair is defined as a pair (θ_i, ξ_j) such that $\mu(\theta_i) \neq \xi_j$, $\xi_j \succeq_{\theta_i} \mu(\theta_i)$ and $\theta_i \succeq_{\xi_j} \mu(\xi_j)$. In strong stability, (θ_i, ξ_j) is a blocking pair if $\mu(\theta_i) = \xi_j$, $\xi_j \succeq_{\theta_i} \mu(\theta_i)$ and $\theta_i \succeq_{\xi_j} \mu(\xi_j)$. Finally, in weak stability, a blocking pair is defined as (θ_i, ξ_j) such that $\mu(\theta_i) = \xi_j$, $\xi_j \succ_{\theta_i} \mu(\theta_i)$ and $\theta_i \succ_{\xi_j} \mu(\xi_j)$. Note that a superstable matching is strongly stable, and a strongly stable matching is weakly stable [13].

A weakly stable matching always exists and can be found in polynomial time. However, there are some cases where superstable or strongly stable matchings may not exist. Nevertheless, there are some heuristics algorithms that are able to find out the existence of a superstable matching [13]. The weak stability definition is similar to a condition based on tie breaking, as is represented by [27]

■ **Theorem 4:** Let μ be a matching in an instance I of stable matching with ties. Then μ is weakly stable in I if and only if μ is stable in some instance of the stable matching algorithm obtained from I by breaking the ties.

In general, there are many ways of breaking the ties. Different ways of tie breaking may result in stable matchings, while some may not. It has been shown [27] that the problem of deciding whether a stable matching with ties admits a weakly stable matching is NP-complete. A matching μ is superstable if there are no two agents x and y that both strictly prefer each other to their partner in μ or are indifferent between them. This is obvious that a superstable matching is weakly stable. The superstability criterion gives rise to the following analogue of theorem 4 [27]:

■ **Theorem 5:** Let μ be a matching in an instance I of stable matching with ties. Then μ is superstable in I if and only if μ is stable in every instance of stable matching algorithm obtained from I by breaking the ties.

In [27], a linear time algorithm has been formulated for finding a superstable matching if one exists, given an instance of stable matching with ties. The existence of an efficient algorithm for stable matching with ties under superstability has remained as an open problem so far.

Applications of matching theory in signal processing and wireless communications

The classification method and the matching models proposed in this article cover the main varieties of matching models with applications in wireless communications and signal processing. This helps wireless engineers achieve a clear understanding of this diverse field, identify the applications of matching games, and utilize suitable matching models for modeling complex problems in communications networks and signal processing. To provide a better understanding, in the following some of the major applications of matching theory in signal processing and wireless communications will be introduced. The works chosen in this section are current research challenges in wireless networks where matching theory has played an effective role in addressing the problems. Some applications of diverse matching models such as matching with transfer, without transfer, one to one, one to many, and matching with externalities have been discussed to show how these tools can be used in network optimization and practical problems.

Radio resource allocation in CR networks

A general multiple PU/multiple SU wireless CR network CR is a promising technology that improves the spectral efficiency in wireless networks. This is achieved by allowing the unlicensed SUs to coexist with licensed PUs in the same band. This coexistence is managed by the spectrum access schemes that set up an agreement between PUs and SUs on spectrum usage. In these networks, PUs are motivated to lease spectrum bands to SUs in exchange for a certain form of compensations. These compensations can be in different types, such as monetary or services applicable in wireless networks.

Although many works on PU/SU spectrum sharing are based on monetary transaction, there are some other forms of transactions. To reach higher data rates, the use of cooperative relaying is a well-known technique to increase the user diversity and provide higher capacity in wireless networks [28]. In [7], the SU cooperatively relays the PU's data in exchange for spectrum access and monetary compensation. When the rate is important, instead of monetary compensation, the matching algorithm is flexible in terms of prioritizing either PUs or SUs, by a manipulation of controlling parameters.

Figure 2 shows the SUs and PUs spectrum-access model that considers an overlay CR wireless network, comprising multiple PU transmitter–receiver pairs, with each pair occupying a unique spectrum band of constant size. In this network, there are multiple SU transmitter–receiver pairs seeking to obtain access to one spectrum band occupied by a PU pair. There are T time slots per transmission frame, and each SU pair has access to a monetary value C . Each primary transmitter (PT) attempts to grant spectrum access to a unique SU pair, as determined by the matching algorithms, in exchange for 1) the secondary transmitter (ST) cooperatively relaying the PT's data to the corresponding primary receiver and 2) monetary compensation. In particular, without loss of generality, let us

consider (PT_1, PR_1) , whose transmission is relayed by ST_2 during a fraction $\beta_{1,2}$ ($0 \leq \beta_{1,2} \leq 1$) of T , while also receiving a fraction $\kappa_{1,2}$ ($0 \leq \kappa_{1,2} \leq 1$) of C from ST_2 . The parameters $\kappa_{1,2}$ and $\beta_{1,2}$ are referred to as the *price* and *time-slot allocation numbers*, respectively, whose exact values will be determined by the matching algorithms.

During the cooperative relaying stage in the initial slot 1, a fraction $\tau_{1,2}$ ($0 < \tau_{1,2} < 1$) is first allocated for PT_1 to broadcast its signal to ST_2 and PR_1 , occurring in the first $\beta_{1,2}\tau_{1,2}T$ time slots. In the subsequent slot 2, ST_2 cooperatively relays the signal from PT_1 to PR_1 . PR_1 then applies maximum ratio combining to the signal received from PT_1 in the first $\beta_{1,2}\tau_{1,2}T$ time slots (slot 1), and the signal received from ST_2 in the subsequent $\beta_{1,2}(1 - \tau_{1,2})T$ time slots (slot 2). After this cooperative relaying stage, PT_1 ceases transmission, allowing ST_2 to transmit to SR_2 over the spectrum occupied by (PT_1, PR_1) in the final $(1 - \beta_{1,2})T$ time slots. It is notable if the amplify-and-forward relaying protocol is considered, $\tau_{1,2} = 1/2$.

To summarize the matching algorithm, each PT will first make an offer to the ST that is first in its preference list. The ST will then check if the offering PT is in its preference list. If it is, and the ST is already matched with another PT, the ST has two choices.

- 1) If the offering PT can provide a better utility than the ST's current matching, then the ST will reject its current matching in favor of the new matching.
- 2) If the offering PT cannot provide a better utility than the ST's current matching, the ST will reject the PT's offer.

If the ST is not matched, then the ST will be matched with the offering PT. If the offering PT is not in the ST's preference list, the ST will reject the offering PT.

Note that if the ST rejects a PT, then the PT will update its proposal, and the PT will either 1) decrease its price-allocation number by a price-step number ϵ or 2) decrease its time-slot allocation number by a time-slot step number, depending on which option maximizes the PT's utility and assuming a positive price and time-slot allocation number and the minimum data rate requirement for the PT is satisfied. The algorithm will then repeat this procedure with each PT until no more matchings are possible.

As discussed previously, a network management mechanism needs to be incentive compatible, that is, the users are motivated to participate and accept the results. However, in the majority of current game theory techniques used in CR, only one type of users is considered as the primary decision makers. In other words, either the PUs or the SUs are involved in the game to determine access to the spectrum resources [29]. In these approaches, the decision makers do not take into account

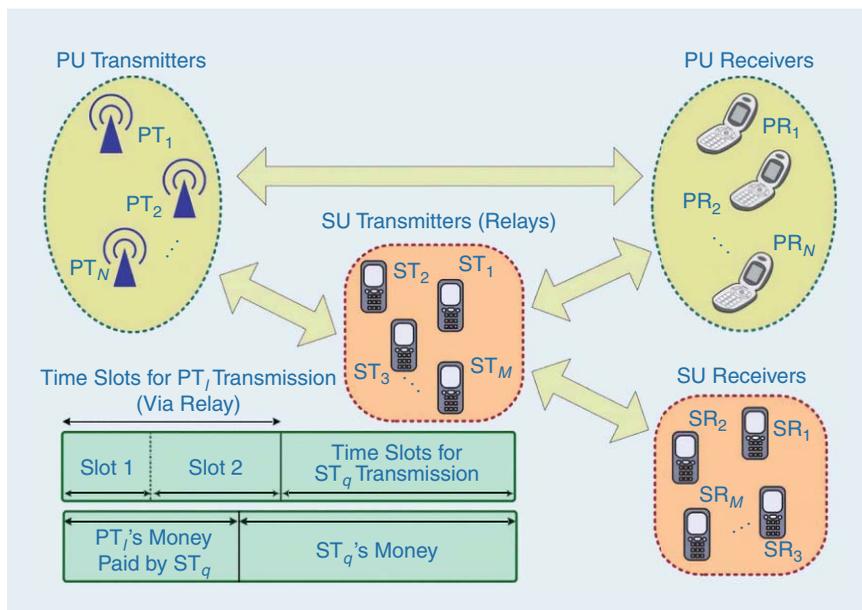


FIGURE 2. The SU and PU spectrum-access model.

the performance metric of the nondecision users, which will lead to unacceptably low performance for these users. Here matching theory approaches are able to obtain a better performance, one that is desirable for the spectrum access strategy to get all users involved in the decision-making process.

In [30] and [7], the proposed distributed spectrum access framework is based on the interaction of both sides, both the PUs and the SUs. Moreover, most of the work in CR literature [31]–[33] considers a scenario with only one PU, while matching-based models can cover multiple players from both PUs and SUs with or without transfer.

Channel allocations in opportunistic spectrum access

There are many works in the literature that discuss opportunistic channel selection in CR networks; however, most of them are based on two assumptions. The first assumption is that when a channel is available, all the SUs can have similar performance using that channel. The second assumption is that the full knowledge of channel parameters or information about other users is always available. However, these assumptions are not realistic in all scenarios.

Due to the dynamic environment in PUs, spectrum availability may be different for different SUs even in the same channel. Moreover, available spectrum spaces are time varying due to the random traffic of PUs. When there are multiple PUs and SUs with different channel information and different performance for each channel, modeling such scenarios is considered challenging without using matching theory. In [34], a channel allocation problem without the limiting assumptions is considered, where each SU only considers its individual history and its local observation and independently makes a decision to access a channel. To improve the network's throughput, it is desirable to design a dynamic and distributed spectrum-sharing strategy. In [34], by using matching theory, a decentralized

optimal learning Gale–Shapley scheme is designed to achieve this goal.

As the Gale–Shapley algorithm always reaches a stable one-to-one matching, multiuser contention under this interference model will be avoided. Moreover, the Gale–Shapley theorem has a unique stable matching when the preference matrix is different with no ties. The Gale–Shapley theorem is always a suitable tool to solve the spectral allocation problem in asymmetrical networks. The Gale–Shapley theorem was applied to effectively reduce multiuser conflicts, and when channels are not available and information exchange is limited, they exploit the order-optimal learning algorithm to achieve stable matching. This proposed distributed algorithm does not require prior knowledge of primary traffic and information exchange about SUs' actions that is an important property for applying to real networks.

In [35] a wireless CR network consisting of multiple PUs and multiple SUs was considered. The PUs and SUs aim to optimize their utilities in terms of transmission rate and power consumption. The combined Stackelberg game and matching theory was proposed to optimize the utility gain of both PUs and SUs. In their model they assumed the PTs can work as decode-and-forward relays for the PUs.

A competitive market model was presented in [36], where the channels are associated with prices and SUs have a monetary budget. SUs demand a set of channels based on their utility function and current channel prices. A many-to-one stable matching was used to model this scenario. The conditions under which the stable matching is unique and is optimal for PUs has been discussed. It shows that the proposed algorithm can achieve a Walrasian equilibrium that maximizes the sum utility of the users and at the same time balances the channel demand and supplies.

The association between SUs in the network and frequency bands licensed to PUs was investigated in [37]. The problem was modeled as a matching game between SUs and PUs where SUs use hypothesis testing to detect PUs' signals and rank them based on the function of probability ratio. A distributed algorithm that enables both SUs and PUs to interact and self-organize into a stable and optimal matching was presented.

Multiple-input, multiple-output techniques for wireless networks

Multiple-input, multiple-output (MIMO) has been a technology that is widely used in wireless communications networks. However, some limitations of mobile devices do not allow installing multiple antennas on these devices. Virtual MIMO (VMIMO) is an effective solution that only requires one antenna in one device and can provide spatial diversity through antenna sharing among mobile devices.

In VMIMO, at least two mobile devices need to be grouped to form a virtual antenna array to communicate with a BS.

User pairing in VMIMO is a critical issue. In some works, pairing schemes such as random pairing scheduling and orthogonal pairing scheduling have been proposed but with low performance.

Due to the existence of multiple numbers of users, different preferences of users, and conflicting interests of users, matching theory is a promising tool for antenna sharing [38]. The stability of the pairing is of high importance in a dynamic VMIMO system, as the communications need to be stable for data transfer. Moreover, the stable pairing motivates the selfish users to participate in the VMIMO cooperation to increase their own spatial diversity. If the pairings are unstable, users have to change their cooperative groups repetitively, leading to excessive control overhead required to update their partners. In this case, the system performance is degraded significantly. The algorithm designed in [38] is distributed and shows a good performance in pairing of users in a VMIMO.

Joint uplink/downlink resource allocation in OFDMA wireless networks

The increase in resource-demanding wireless applications such as online multimedia streaming and many other online services requires an efficient radio resource allocation in wireless networks. The need of the new wireless services for high QoS requirements forces the system designers to

use approaches such as OFDMA schemes. The OFDMA resource allocation has been widely studied in the literature for both the downlink and the uplink using a variety of tools such as classical optimization, auction theory, and game theory. Most of the existing works have mainly focused on optimizing the uplink and downlink users independently. However, there are many applications for which the performance of the users needs to be optimized in uplink and downlink at the same time, and an end-to-end QoS for the user in both directions should be guaranteed, for example,

in scenarios such as video conferencing where user satisfaction depends on both uplink and downlink performance.

To deal with a joint uplink–downlink resource allocation problem, it is essential to develop a distributed solution where the users can make distributed actions on the channels. Designing such a distributed algorithm in OFDMA networks is complex, as multiple users are competing to select multiple channels, and each user has a specific preference over the channels.

Matching theory is a suitable tool for modeling and finding efficient solutions for joint uplink–downlink resource allocation. In [39], a distributed resource allocation algorithm by using matching theory is proposed with both uplink and downlink QoS requirement guarantees. This work modeled the problem as a two-sided stable matching game and proposed an algorithm for performing the matching within an OFDMA system in the absence of channel state information

The outcome of the problem needs to be a set of marriages such that there are no two people of opposite sex who would both prefer each other over their current partners. In other words, the marriages need to be stable.

at the transmitter side. The proposed algorithm allocates the channels to the users in such a way as to maximize a utility function that captures the joint uplink–downlink QoS requirements of the users.

Physical layer security in wireless networks

The security of a wireless communication link has always been of great importance. The broadcast nature of the wireless transmission medium makes eavesdropping easy, and anyone within communication range can receive and possibly decode private transmission signals. Moreover, higher-layer security key distribution and management may be difficult to implement and may be vulnerable to attacks in some environments, such as ad hoc or relay networks, in which transceivers may join or leave randomly [40]. Alternatively, physical layer-based security explores the characteristics of the wireless channel to improve wireless transmission security.

The use of a friendly jammer to facilitate the degradation of the source to an eavesdropper channel has been considered (see, e.g., [41]). This is achieved by a friendly jammer transmitting a jamming power signal, which has the effect of decreasing the signal-to-noise ratio at the eavesdropper. This approach is often referred to as *cooperative jamming*. With the additional degrees of freedom provided by multi-antenna systems, friendly jammers can generate artificial noise to degrade the channel condition of the eavesdropper while maintaining little interference to the source nodes.

Most existing literature assumes perfect channel knowledge from the source to the eavesdropper. This is valid if the eavesdropper is part of the communication system. For example, in [42], a scenario was considered in which the receivers eavesdrop on the message intended for other receivers. However, in some cases, the eavesdropper may not be a user of the system, and obtaining the eavesdropper channel information is difficult. The algorithm that facilitates using the eavesdropper needs to have acceptable performance without perfect channel knowledge.

Another requirement is to select the best friendly jammer for each source–destination pair when there are multiple friendly jammers and multiple source–destination pairs. The final selection needs to be stable to guarantee a long-term secure communication. There are some centralized [43], [44] and distributed [45], [46] methods based on game theory that have been proposed to handle the interaction between the source nodes and the friendly jammer. However, they did not cover a general scenario with multiple source–destination pairs and multiple friendly jammers.

In [9], a matching theory-based distributed algorithm is proposed to

establish secure communications for multiple source–destination pairs from a malicious eavesdropper, through the assistance of multiple friendly jammers. To enhance the overall network secrecy rate [43], source nodes are willing to provide a certain amount of monetary compensation to jammers in exchange for jamming power. The matching algorithm in [9] determines both 1) the matched source node–friendly jammer pairs and 2) the exact amount of money transfer that motivates both source nodes and friendly jammers to cooperate, such that the final matching is stable and maximizes the sum of utilities of all source nodes and friendly jammers, referred to as the *network social welfare*. The framework is a dynamic matching with transfer and the proposed matching algorithm converges to a competitive equilibrium.

Figure 3 shows the physical layer security wireless network model comprising multiple source node S –destination node D pairs, with each pair using one spectrum band with the same size. There is a set of multiple friendly jammers J and one malicious eavesdropper node in the network. For simplicity, the friendly jammers are all equipped with multiple antennas, while the eavesdropper, source nodes, and destination nodes are equipped with a single antenna each.

Friendly jammers compete with each other to assist a particular source–destination pair by creating sufficient interference to the eavesdropper. In exchange, the source–destination pair will pay a monetary amount to the jammer. For simplicity, it is assumed that during each time period, each source node will be assisted by only one friendly jammer.

First, access points (APs) that are not matched make a price offer to the unmatched source nodes. Each source node then chooses the jammer that provides it with the highest positive utility. Then, each jammer will decide if it wants to match with the source nodes. There are three possibilities. The first

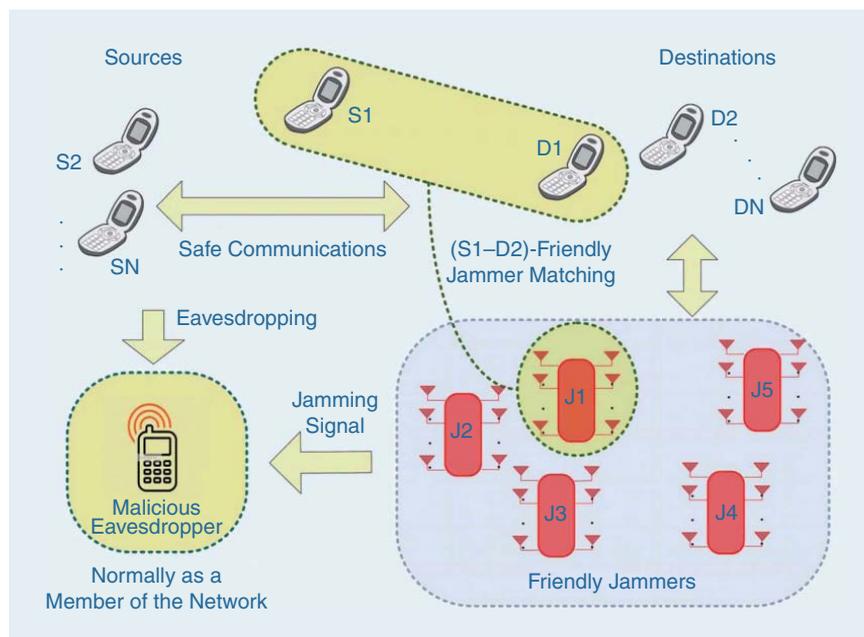


FIGURE 3. The physical layer security wireless network model with multiple friendly jammers equipped with multiple antennas.

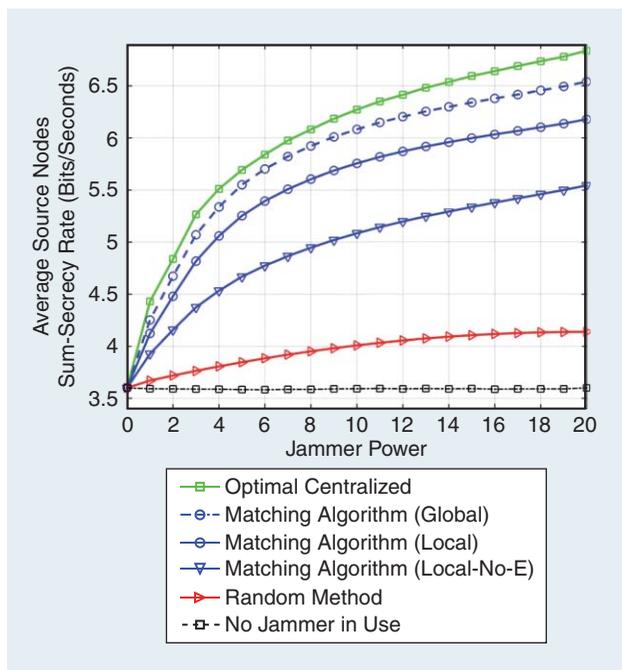


FIGURE 4. The average source nodes sum-secrecy rate versus jamming power.

possibility is if the jammer receives no offer by source nodes after increasing its price-allocation number, while in the previous iteration the jammer had received bids from multiple source nodes. The jammer will choose one of these source nodes at random to be matched with. The second possibility is if the jammer has offers from multiple source nodes. The jammer will increase its price-allocation number by ϵ for the next iteration, that is, $(t + 1)$ st iteration. Note that here also $\epsilon \in \mathbb{R}^+$ represents the price-step number, which indicates the amount that the price-allocation number offer from the jammers to the source increases. The third possibility is if the jammer has an offer from only one source node. In this case, the jammer will be matched with this node. This procedure will continue until there is no new offer.

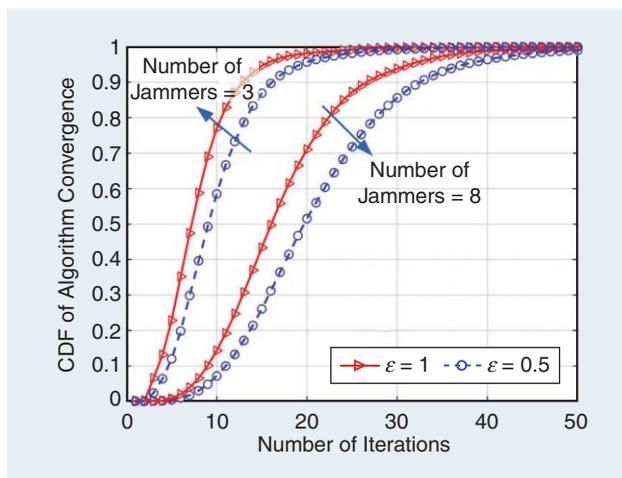


FIGURE 5. The CDF of the matching algorithm convergence versus average number of iterations [9].

Figure 4 shows the average source nodes sum-secrecy rate over all source nodes versus the friendly jammers' power, where the average is taken with respect to all channels. It is observed that the matching algorithm achieves a significant proportion of the social welfare of the centralized method. It is also obvious that the matching algorithm performs significantly better than random matching, and even more so when there is no jamming power.

Another interesting scenario is discussed in [9]. In this scenario, the source nodes and jammers have local knowledge of the channel state information (CSI) and source nodes have access to the distribution of the jammers' CSI. A closed-form expression is derived for the achievable secrecy rate, the scenario where local CSI is known but no channel information regarding the eavesdropper is known.

The matching algorithm in [9] is distributed, incurring significantly less overhead and complexity compared to centralized algorithms. An exact analysis of the amount of overhead and complexity is difficult, due to the dependency on a number of system parameters, such as the friendly jammers' minimum price-allocation number $\beta_{\min,q}$, the price-step number ϵ , and the number of source nodes and friendly jammers. However, an expression for the upper bound on the maximum number of communication packets between the source nodes and the friendly jammers is given as:

■ **Theorem 6:** The number of communication packets between source nodes and friendly jammers required in the SJMA is upper bounded by

$$\begin{aligned} \mathcal{N}_{\max} &= (N + F + MN) \mathcal{I}_{\max} \\ &= \frac{(N + F + MN)}{\epsilon} \max_{q \in \mathcal{J}} (\beta_q^{\max} - \beta_{\min,q}), \end{aligned} \quad (5)$$

where N is the number of source-destination pairs, M is the number of friendly jammers, $F = \min\{N, M\}$,

$$\mathcal{I}_{\max} = \frac{1}{\epsilon} \max_{q \in \mathcal{J}} (\beta_q^{\max} - \beta_{\min,q}),$$

where $\beta_q^{\max} = \max_{\ell \in \mathcal{S}} \beta_{\max,\ell,q}$, and $\beta_{\max,\ell,q}$ is the friendly jammer's maximum price-allocation number.

The complexity of the centralized algorithm is given by [9]

$$\begin{cases} \mathcal{O}\left(\frac{N!}{(N-M)!} 2^{2N+M}\right), & \text{if } N \geq M; \\ \mathcal{O}\left(\frac{M!}{(M-N)!} 2^{2N+M}\right), & \text{if } N < M. \end{cases} \quad (6)$$

The complexity of the matching algorithm is $\mathcal{O}(N + F + NM)$ and for the random algorithm is $\mathcal{O}(F)$. It is observed that the centralized method has significantly higher complexity than both the proposed and randomized algorithm and increases exponentially in the number of source nodes and jammers. In contrast, the complexity of the proposed algorithm only increases linearly with the number of source nodes or jammers. Figure 5 also depicts the cumulative distribution function (CDF) of the matching algorithm convergence versus average number of iterations. It shows that, in the worst case scenarios, the algorithm will converge after almost 40–50

iterations. It is observed in (5) that the amount of overhead, and the number of iterations, decreases with ϵ .

User/BS association in heterogeneous wireless networks

To meet the increasing demands for reliable wireless services and the need for high data rates, service providers (SPs) are challenged to improve their next-generation heterogeneous wireless access networks (HetNet) by utilizing small, low-cost, and low-power APs, known as femtocell APs (FAPs) [47], in addition to their macrocell APs (MAPs). FAPs can be overlaid on any existing wireless technology (e.g., 2G, 3G, LTE, WiMAX, etc.) and are envisioned as a cost-effective solution for improving the performance of wireless networks while enabling resource-demanding applications and personalized wireless services [48].

To reap the benefits of FAP deployment, numerous technical challenges such as interference management, optimal association of user equipment (UE) to APs, and optimal allocating of the FAPs to the SPs must be addressed [49], [50]. Associating the UE to their serving APs for uplink transmission is a key challenge in FAP networks that remains relatively unexplored, as outlined in [48] and [51]. Inherently, within an overlaid HetNet, the problem of assigning the UE to their serving APs faces a number of challenges that are significantly different from classical CU association problems [52]. Unlike MAPs, the FAPs are typically randomly deployed in the network without predimensioning or planning and are resource constrained in nature, which can significantly impact the UE association problem. Moreover, although some operators might use dedicated FAPs, in general, the FAPs are likely to be owned by private entities such as SPs or home users [51]. To improve the QoS for their UE and to ensure that every UE can receive satisfactory wireless service, the SPs would like to deploy FAPs in their networks [53]. More formally, the SPs are willing to deploy the FAPs in their network to reduce their churn rate. The competition among the SPs motivates the unsatisfied subscribers to change their SP, which is called churning. It is important for the SPs to manage the network resources based on the expectation of the UE; otherwise, the UE may churn to a different provider.

In [8], a distributed framework based on the matching theory is developed that associates UE with APs and then establishes an optimal and stable cooperation between multiple competitive SPs and competitive FAPs. Under this framework, the SPs, FAPs, and UE are assumed to be selfish and rational entities that merely care about their own interests. To achieve a higher data rate, the UE compete to connect to a FAP that provides the highest data rate. The limited FAP resources and competition among the UE motivate the UE to offer a certain amount of monetary compensation to the FAPs in exchange for receiving the wireless service. The proposed algorithm models the competition among the UE and FAPs and determines both 1) the serving AP and the set of subchannels allocated to each UE and 2) the exact amount of money transfer that motivates both the UE FAPs to cooperate

such that the final matching is stable and the total UE sum rate is maximized.

In summary, the distributed mechanism in [8] jointly performs associating the UE to the APs and allocating the FAPs to the SPs. The proposed mechanism motivates SPs and FAPs to cooperate with each other such that the total UE satisfaction level is maximized. Then they proposed a distributed subchannel allocation algorithm for the uplink OFDMA networks, which is beyond the scope of this tutorial.

Figure 6 represents a general configuration of an OFDMA uplink wireless communication network, comprising multiple SPs and multiple FAPs. It is assumed that each SP owns multiple MAPs and multiple subscribed UE denoted. Each UE is required to connect to either a MAP or a FAP to receive wireless services. The term AP is used to refer to a MAP or a FAP. The FAPs utilize an open access control mechanism that allows arbitrary nearby CUs to use the FAP's wireless services [51]. Each SP can utilize multiple FAPs to serve its subscribed UE. For simplicity, it is assumed that each FAP can be used by only one SP at each transmission frame. Note that the inter-FAP interference occurs when two FAPs allocate the same subchannel to two different UE; however, due to the MAP's fixed configuration to MAPs, the inter-MAP interference is negligible. Each AP has access to multiple subchannels. To guarantee fairness among the UE, a maximum number of subchannels allocated to each UE was set. Each AP can serve a maximum number of UE that will be referred to as the *quota* of an AP. Note that the quota of the MAPs is much larger than the quota of the FAPs.

First each UE will form a descending order preference list in terms of its utility over all the potentially available APs, that is, the first AP in its list corresponds to $AP^{\dagger} = \text{argmax } \mathcal{U}_{UE}$. The MAPs are able to accommodate a large number of the UE, and it is assumed a bid from UE to a MAP will always be accepted and the matchings are permanent. However, because a FAP has a limited quota, in case the number of its received bids is more than its quota, the FAP only selects the most preferred UE, which results in the highest FAP's utility and rejects the others. Because the UE are subscribed to different SPs, each FAP has different lists of temporarily matched UE, and at the end of the association algorithm they determine which set of temporarily matched UE will be served by each FAP.

To deploy the FAPs in their networks, the SPs should define new tariffs within their existing billing system and charge the UE that ask for more resources (e.g., requesting for more subchannels) or ask for FAP access. The exact amount of this charge is determined by another matching algorithm in [8] that is a matching with transfer that was explained previously.

Finally, to match the FAPs to each SP an iterative matching algorithm with transfer was developed. The purpose of the proposed algorithm is to obtain a solution to the optimization problem (7) in a distributed way by utilizing the outcome of the FM. The optimization problem that maximizes the total satisfaction level of the UE subscribed to each SP can be written as

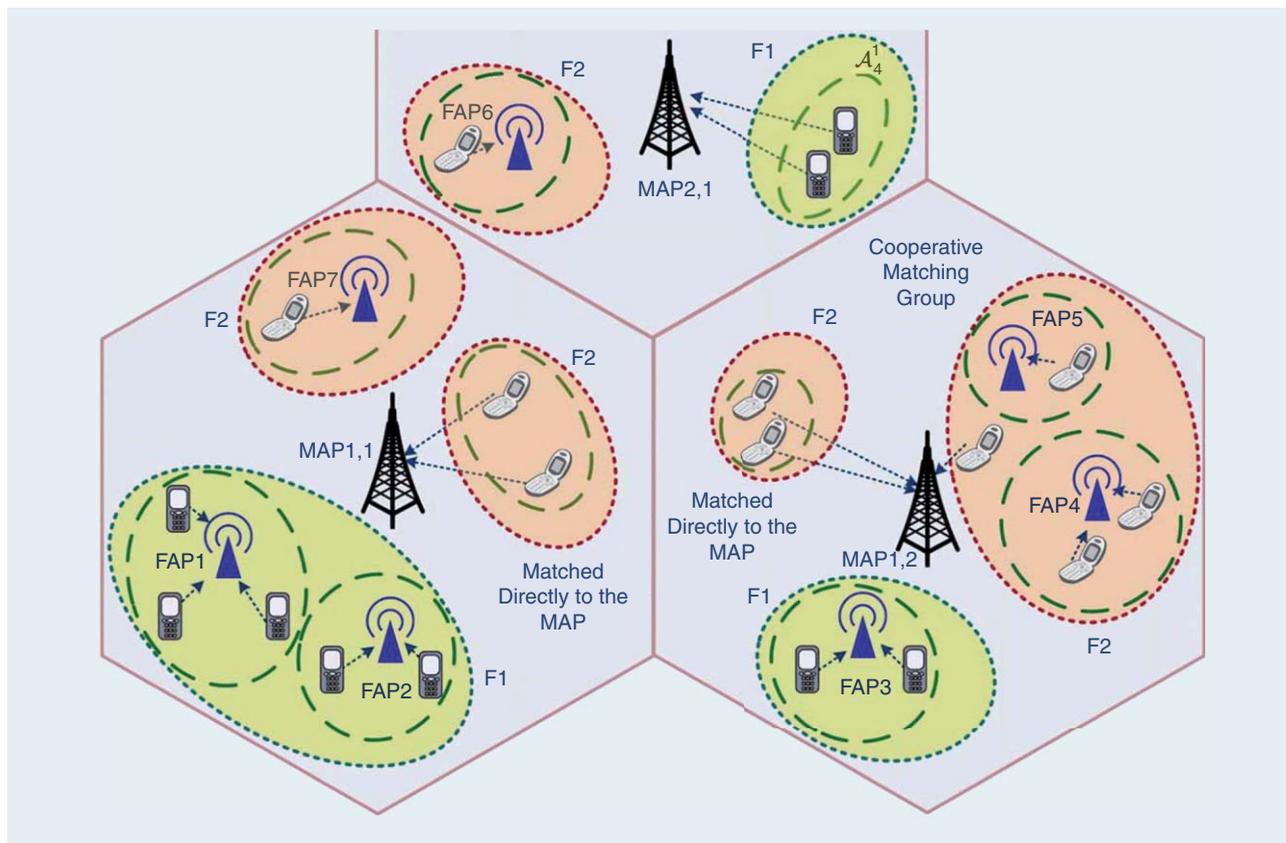


FIGURE 6. The HetNet with multiple SPs and multiple FAPs [8].

$$\begin{aligned}
 & \max_{X, \{\mathcal{F}_i\}_{i=1}^N} \zeta \sum_{i=1}^N \sum_{j=1}^F x_{ij} [Q_{SP^i}(\mathcal{F}_i, \mathcal{A}_{\mathcal{F}_i}^i) - c] \\
 & \text{s.t. : } 1) \sum_{i \in \mathcal{W}} x_{ij} \leq 1, \quad \forall j \in \mathcal{F}, \\
 & \quad 2) \sum_{j \in \mathcal{F}} x_{ij} \leq F, \quad \forall i \in \mathcal{W}, \\
 & \quad 3) x_{ij} \in \{0, 1\}, \quad \forall i \in \mathcal{W}, \forall j \in \mathcal{F}, \quad (7)
 \end{aligned}$$

where X is the matrix made up of elements $x_{ij}, i \in \mathcal{M}_i, j \in \mathcal{F}$, which represents if the i th SP and j th FAP are matched or not. Condition 1 states that each FAP can be matched with one SP, condition 2 states that at most F FAPs can be matched with one SP, and condition 3 states that the values of x_{ij} can be only 0 and 1.

First, suitable and simple utility functions for the FAPs and SPs should be defined. The utility of each SP is defined as the difference between the revenue it gains through its subscribed UE and its associated cost that is paid to FAPs. The utility function for SP can be written as

$$\mathcal{U}_{SP^i}(\mathcal{F}_i, \{p_{ij}^{SP}\}) = \zeta Q_{SP^i}(\mathcal{F}_i) - \sum_{j \in \mathcal{F}_i} p_{ij}^{SP}, \quad (8)$$

where Q_{SP^i} is the satisfaction level of the UE subscribed to SP^i and is proportional to the subscriber's rate, QoS, and so forth, $\zeta \in \mathbb{R}^+$ is a fixed coefficient with unit revenue (price) per user satisfaction, \mathcal{F}_i is the set of FAPs hired by SP_i , and p_{ij}^{SP} is the amount of money given to $FAP_i \in \mathcal{F}_i$ by SP_i .

The FAPs are willing is to maximize their revenue, so the utility function for the FAPs should reflect their monetary gain, and the utility function for FAP_j matched with SP_i is given by

$$\mathcal{U}_{FAP_j}(p_{ij}^{SP}) = p_{ij}^{SP} - c, \quad (9)$$

where c is the minimum required cost for a FAP to cooperate with SPs. The network's social welfare is now defined as the sum of the utilities of all the SPs and the FAPs in the network, given by

$$\begin{aligned}
 \text{S Welfare} &= \sum_{i=0}^N \sum_{j=0}^F x_{ij} [\mathcal{U}_{SP^i}(\mathcal{F}_i, \{p_{ij}^{SP}\}) + \mathcal{U}_{FAP_j}^{SM}(p_{ij}^{SP})] \\
 &= \zeta \sum_{i=0}^N \sum_{j=0}^F x_{ij} [Q_{SP^i}(\mathcal{F}_i, \mathcal{A}_{\mathcal{F}_i}^i) - c], \quad (10)
 \end{aligned}$$

where N and F are the number of SPs and FAPs in the network, respectively. Interestingly the optimization (7) maximizes the network's social welfare in (10). By noting the fact that the satisfaction of the UE with FAP utilization is always greater than that without FAP utilization, the "0" dummy index summation terms can be removed from (10) without changing the problem objective.

First each FAP that is not matched makes a price-allocation number offer to the unmatched SPs. Each SP then determines the set of FAPs that provide the highest positive utility.

The i th SP firms its demand set. The demand set for SP_i is equal to the index of the FAPs that provides the maximum utility. If the demand set is not empty, the SP will bid for the FAP in its demand set; otherwise, the SP will not bid and remains unmatched.

Then, each FAP will decide if it wants to match with the SPs. There are three possibilities. The first possibility is if some FAPs have bids from multiple SPs. The FAP will increase its price-allocation number by ϵ for the next iteration, that is, the $(t + 1)$ 'st iteration. Note that here also $\epsilon \in \mathbb{R}^+$ represents the price-step number, which indicates the increment of the price-allocation number offer. The second possibility is if a group of the FAPs has a bid from only one SP. In this case, this group of the FAPs will be matched with this SP. The third possibility is that some of the FAPs do not receive any bid from the SPs. In this case they do not change their price-step number and wait for the next round. The algorithm ends when there is no new offer.

Cellular machine-to-machine communications network

Machine-to-machine (M2M) communications is a new type of communications that enables pervasive connectivity between one or more autonomous machine-type devices (MTD) without or with minimal human intervention [54]. M2M communication is the key technology to facilitate the new paradigm of applications, such as smart city, intelligent transportation, building automation, eHealth, security, and surveillance.

Cellular networks are currently one of the most widely deployed wireless networks with wide coverage, and they have played an important role in M2M communications development. To fully use the widely deployed communications infrastructure, it is beneficial and economical to support M2M communications by using the existing cellular networks. In M2M networks, there is a massive number of MTD devices, significantly higher than the number of users in existing cellular networks, each has only a small amount of traffic, and they are mostly distributed in the uplink [55].

Although some existing cellular technologies such as LTE-Advanced incorporated some limited M2M features in their design, there is little progress in addressing these challenges [56], such as excessive overhead, delays, and control channel congestion.

To guarantee a minimum required QoS for a large number of MTDs with diverse QoS characteristics in congested areas, [57], [58] developed a 3GPP access mechanism. However, the effects

of this mechanism on the CUs and the CUs' performance degradation have not been discussed. All these works are centralized approaches. For the M2M networks with a large number of MTDs, the centralized schemes will be highly complex and introduce excessive overhead. In [59], a distributed resource allocation algorithm for cellular M2M communication networks is proposed, where all MTD traffic is sent through the channels allocated to some CUs. They assumed that different MTDs have different requirements. Then they developed a distributed matching algorithm based on matching theory, which negotiates between CUs and MTDs and determines matching from MTDs to CUs. The MTDs will provide monetary compensation to the CUs in exchange for the subchannel access. Under this framework, the CUs and MTDs are assumed to be selfish and rational entities, as they are in practical scenarios, and their negotiation results in a stable equilibrium point such that the sum weighted data rate of all the MTDs is maximized. In the sum weighted data rate, the data rate of an MTD is weighted according to its priority, and more urgent MTDs will be given a higher weight.

They considered an OFDMA uplink single-cell scenario with multiple CUs and one eNB located at the center of the cell as depicted in Figure 7. All the CUs, MTDs, and the eNB are equipped with a single omnidirectional antenna. It is assumed that each MTD has a minimum required data rate and a maximum tolerable delay. A set of orthogonal subchannels are

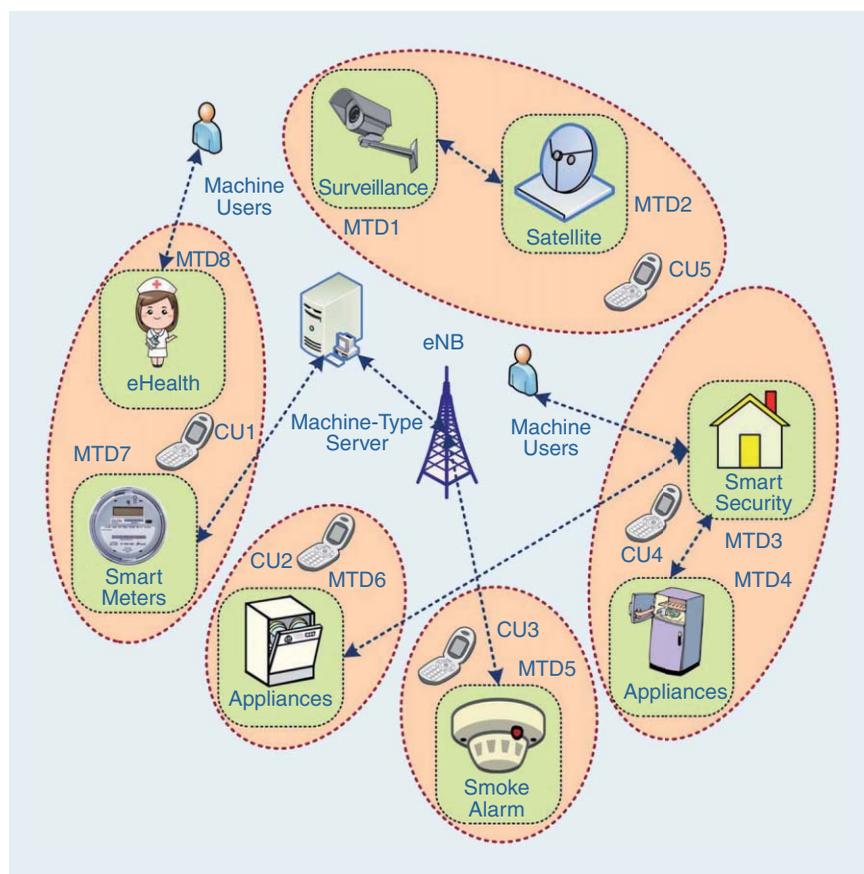


FIGURE 7. An M2M wireless network configuration [59].

allocated to each CU. It is assumed that each subchannel is allocated to only one CU, implying that there will be no interference observed from the CUs at the eNB.

The subchannels allocated to the CUs are fixed for each transmission frame. MTDs seek to obtain access to subchannels occupied by CUs to transmit their data. In exchange for the subchannel access, MTDs will provide monetary compensation to the CUs. This is fictitious money or some kind of credit/reward to coordinate the system. During each transmission frame, each CU has incentive to share its extra time slots with the MTDs, if it is profitable. Each CU has a minimum acceptable price.

If an MTD accesses a CU's subchannel during a fraction of time, of T , then the MTD will pay a certain amount of money to the CU in exchange for accessing its subchannel. It is also shown that the proposed algorithm has a performance comparable to the centralized method with a significantly lower complexity. Figure 8 plots the average percentage of MTDs with their delays not met versus the number of the MTDs in the network. As was expected by increasing the number of MTDs, more MTDs are unable to meet their required delays. This also shows that the proposed algorithm has a comparable performance with a centralized algorithm and performs much better than the random access algorithm.

Externality matching-based user-cell stable association in small-cell networks

Context-aware user-cell association

As discussed previously, recently, the concept of heterogeneous networks has attracted a lot of attention as a suitable and reliable approach to meet the increasing data traffic demands in wireless networks. Although the deployment of small cells is a very interesting area, and the utilization of small cells comes with considerable potential benefits, there are several technical challenges such as interference management and network management that should be resolved.

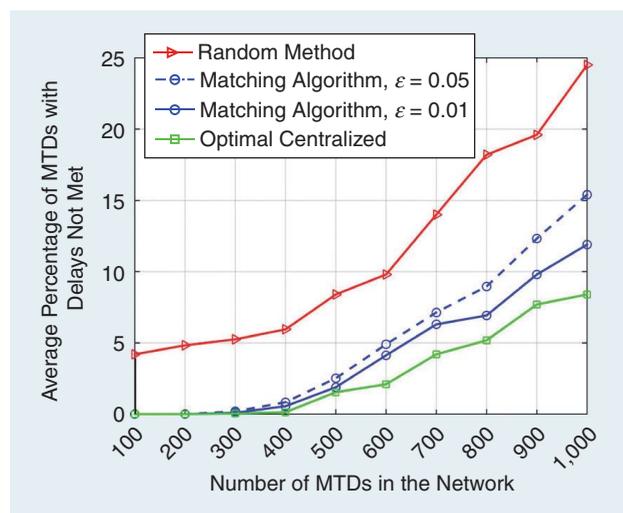


FIGURE 8. The average percentage of MTDs with delays not met versus the number of MTDs in the network.

Among all the difficulties, one main technical challenge in small-cell networks is that of user association. Small-cell networks consist of different types of APs (femtocells, picocells, and microcells) with different requirements. Hence, due to the heterogeneous nature of small-cell networks, traditional macrocell-defined user-cell association strategies to small-cell networks cannot be used, and new association schemes that are tailored to the specific nature of wireless small-cell networks are required. In [60], other user information is considered in the design, such as the user's mobility pattern and the urgency of traffic each user imposes to the network. Here, such additional information is referred to as the *user's context information*.

In [60], it is shown that, by using such suitable context information, the network can make better decisions about which user should be serviced by which SCBS while also satisfying the QoS requirements of the users. In [60], a matching-based distributed association scheme in the downlink of small-cell networks is proposed that uses new context information related to the users' trajectory profiles, their QoS demands, and the current load of each cell.

The effects of this algorithm on the association strategy are modeled via novel utility functions. They formulated the association problem as a many-to-one matching game with externalities, in which the preferences of the players, that is, users and small-cell BSs, are interdependent. Such interdependency, which stems from the mutual interference, significantly changes the game properties as opposed to classical matching games. To solve the proposed matching game, they proposed a novel self-organizing algorithm that can dynamically update the preference lists in the presence of externalities and is able to reach a stable matching between the users and their serving small-cell BSs.

Another promising approach of context-aware information is that the small-cell BS networks can anticipate the resource allocation and devise the UE–small-cell BS association that delivers the largest individual quality of experience (QoE). Anticipating network operations based on user context information is an emerging topic in wireless cellular communications.

In [61], the concepts of QoE prediction and user context for solving the UE–small-cell BS association problem are combined in the downlink of small-cell networks. They discussed how the typical set of active applications and the qualitative feedback on past data services can be translated, at the network level, into decisions on which UE should be served and by which small-cell BS. They also studied this problem by modeling the QoE feedbacks as mean opinion scores, which account for the characteristics of both wireless transmissions and the multimedia data services in use at the UE's side. In particular, collaborative filtering uses known user information and correlations across similar services so as to make recommendations or predictions on which multimedia applications a new user is most likely to request and with which expected QoE. In summary, this work has demonstrated how to exploit user context information, available at each small-cell BS, for making better informed decisions, thereby improving the overall network performance.

Conclusions and future directions

In this tutorial article, a comprehensive overview of matching theory, its branches, and its applications in wireless communications and signal processing was presented. A novel classification of matching models from their practical point of view was presented first, and then the properties and structure of each model were explained such that a network designer is capable to select an appropriate matching model for a specific network. In each matching model, the key elements and the configurations were introduced, and the associated analytical tools were explained. A general form of a distributed matching algorithm that can be customized for all the matching models was discussed and explained. Finally, the wireless communications applications for each matching model were discussed and a broad area of the topics in wireless communications and signal processing with various scenarios were covered. The applications that are presented in this tutorial were selected from a broad range of areas and related works that discussed a variety of research problems. In this tutorial, the goal was to provide a comprehensive overview of matching theory suitable to the demands of wireless communications and network engineers to address major technical opportunities and challenges in today's and future wireless networks.

Acknowledgments

This work was supported in part by ARC grants DP150104019, FT120100487, the faculty research cluster program and the faculty early career researcher scheme from the faculty of engineering and information technologies, the University of Sydney, U.S. NSF CPS-1646607, ECCS-1547201, CCF-1456921, CNS-1443917, ECCS-1405121, and NSFC 61428101.

Authors

Siavash Bayat (bayat@sharif.edu) received his B.S. and M.S. degrees in electrical engineering in 2002 and 2005, respectively, and his Ph.D. degree in electrical engineering from the University of Sydney, Australia, in 2013. He was then a postdoctoral research fellow at the University of Sydney, prior to joining the Sharif University of Technology in 2014, where he is currently an assistant professor. From 2007 to 2009, he was affiliated with Sharif University of Technology, where he held a position of faculty member with responsibility for research in wireless communication networks system design. His research interests include wireless resource management, wireless communications, signal processing, heterogeneous networks, cognitive radio, game theory, and physical layer security. He is a Member of the IEEE.

Yonghui Li (yonghui.li@sydney.edu.au) received his B.S. degree from Jilin University, China, and his Ph.D. degree in 2002 from Beijing University of Aeronautics and Astronautics. From 1999 to 2003, he was affiliated with Linkair Communication Inc. Since 2003, he has been with the Centre of Excellence in Telecommunications, the University of Sydney, Australia. He is now a professor in the School of Electrical and Information Engineering at the University of Sydney. His current research interests are in the area of wireless communications, with a

particular focus on multiple-input, multiple-output millimeter wave communications, machine-to-machine communications, coding techniques, and cooperative communications. He holds a number of patents granted and pending in these fields. He is an editor of *IEEE Transactions on Communications* and *IEEE Transactions on Vehicular Technology* and an executive editor of *European Transactions on Telecommunications*. He received best paper awards from the 2014 IEEE International Conference on Communications and 2014 IEEE Wireless Days Conferences. He is a Senior Member of the IEEE.

Lingyang Song (lingyang.song@pku.edu.cn) received his B.S. degree from Jilin University, China, and his Ph.D. degree from the University of York, United Kingdom, in 2007, where he received the K.M. Stott Prize for excellent research. He worked as a research fellow at the University of Oslo, Norway, until rejoining Philips Research UK in March 2008. In May 2009, he joined the School of Electronics Engineering and Computer Science, Peking University, China, as a full professor. His main research interests include multiple-input, multiple-output, cognitive and cooperative communications, security, and big data. He received the IEEE Leonard G. Abraham Prize in 2016 and the IEEE Asia Pacific Young Researcher Award in 2012. He is currently on the editorial board of *IEEE Transactions on Wireless Communications*. He has been an IEEE Distinguished Lecturer since 2015. He is a Senior Member of the IEEE.

Zhu Han (zhan2@uh.edu) received his B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1997 and his M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. Currently, he is a professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He is an IEEE Communications Society Distinguished Lecturer. He is a Fellow of the IEEE.

References

- [1] L. A. Dasilva, H. Bogucka, and A. Mackenzie, "Game theory in wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 110–111, June 2011.
- [2] M. Debbah, "Mobile flexible networks: The challenges ahead," in *Proc. Int. Conf. Advanced Technologies for Communications*, Hanoi, Vietnam, 2008, pp. 3–7.
- [3] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [4] A. E. Roth, "Deferred acceptance algorithms: History, theory, practice, and open questions," *Internat. J. Game Theory*, vol. 36, no. 6, pp. 537–569, July 2008.
- [5] A. Roth and M. Sotomayor, *Two Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1989.
- [6] Y. Leshem and E. Zehavi, "Stable matching for channel access control in cognitive radio systems," in *Proc. 2nd Int. Workshop Cognitive Information Processing (CIP)*, Elba Island, Italy, 2010, pp. 470–475.
- [7] S. Bayat, R. H. Y. Louie, B. Vucetic, and Y. Li, "Dynamic decentralized algorithms for cognitive radio relay networks with multiple primary and secondary users utilizing matching theory," *Trans. Emerg. Telecomm. Techs.*, vol. 24, no. 5, pp. 486–502, Aug. 2013.
- [8] S. Bayat, R. Louie, Z. Han, B. Vucetic, and Y. Li, "Distributed user association and femtocell allocation in heterogeneous wireless networks," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 3027–3043, Aug. 2014.

- [9] S. Bayat, R. Louie, Z. Han, B. Vucetic, and Y. Li, "Physical-layer security in distributed wireless networks using matching theory," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 5, pp. 717–732, May 2013.
- [10] E. Jorswieck, "Stable matchings for resource allocation in wireless networks," in *Proc. 17th Int. Conf. Digital Signal Processing (DSP)*, Corfu, Greece, 2011, pp. 1–8.
- [11] E. Y. Lu and S. Q. Zheng, "A parallel iterative improvement stable matching algorithm," in *Proc. High Performance Computing*, Hyderabad, India, 2003, pp. 55–65.
- [12] A. Schweizer, J. Siliquini, and T. Devadason, "On the performance of advanced stable matching algorithms in combined input output queued network switches," in *Proc. Asia-Pacific Conf. Communications (APCC)*, Bangkok, Thailand, 2007, pp. 223–226.
- [13] K. Iwama and S. Miyazaki, "A survey of the stable marriage problem and its variants," in *Proc. Int. Conf. Informatics Education and Research for Knowledge-Circulating Society (ICKS)*, Kyoto, Japan, 2008, pp. 131–136.
- [14] A. Kipnis and B. Patt-Shamir, "A note on distributed stable matching," in *Proc. 29th IEEE Int. Conf. Distributed Computing Systems (ICDCS)*, Montreal, Canada, 2009, pp. 466–473.
- [15] H. Xu and B. Li, "Seen as stable marriages," in *Proc. IEEE Conf. Computer Communications (INFOCOM)*, Shanghai, China, 2011, pp. 586–590.
- [16] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, "Matching theory for future wireless networks: Fundamentals and applications," *IEEE Commun. Mag.*, vol. 53, no. 15, pp. 52–59, May 2015.
- [17] P. Liu, R. Berry, and M. Honig, "A fluid analysis of a utility-based wireless scheduling policy," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 2872–2889, July 2006.
- [18] Z. Jiang, Y. Ge, and Y. Li, "Max-utility wireless resource management for best-effort traffic," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 100–111, Jan. 2005.
- [19] K. Eriksson, J. Sjöstrand, and P. Strimling, "Three-dimensional stable matching with cyclic preferences," *Mathematical Social Science*, vol. 52, pp. 77–87, May 2006.
- [20] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *Amer. Math. Monthly*, vol. 69, no. 1, pp. 9–14, Jan. 1962.
- [21] D. Gusfield and R. W. Irving, *The Stable Marriage Problem: Structure and Algorithms*. Cambridge, MA: MIT Press, 1989.
- [22] A. Roth, "Stability and polarization of interests in job matching," *Econometrica*, vol. 52, pp. 47–57, 1984.
- [23] V. P. Crawford and E. M. Knoer, "Job matching with heterogeneous firms and workers," *Econometrica*, vol. 49, no. 2, pp. 437–450, Mar. 1981.
- [24] V. Krishna, *Auction Theory*, 2nd ed. Amsterdam, The Netherlands: Elsevier, 2010.
- [25] G. Demange and D. Gale, "The strategy structure of two-sided matching markets," *Econometrica*, vol. 53, no. 4, pp. 873–888, 1985.
- [26] D. Gale, "Equilibrium in a discrete exchange economy with money," *Internat. J. Game Theory*, vol. 13, no. 1, pp. 61–64, 1984.
- [27] R. W. Irving and D. F. Manlove, "The stable roommates problem with ties," *J. Algorith.*, vol. 43, no. 1, pp. 85–105, Oct. 2002.
- [28] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [29] B. Wang, Y. Wu, Z. J. K. Liu, and T. Clancy, "Game theoretical mechanism design methods," *IEEE Trans. Signal Process.*, vol. 25, no. 6, pp. 74–84, Nov. 2008.
- [30] S. Bayat, R. H. Y. Louie, Y. Li, and B. Vucetic, "Cognitive radio relay networks with multiple primary and secondary users: Distributed stable matching algorithms for spectrum access," in *Proc. 2011 IEEE Int. Conf. Communications (ICC)*, Kyoto, Japan, pp. 1–6.
- [31] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.
- [32] L. Giupponi and C. Ibars, "Distributed cooperation among cognitive radios with complete and incomplete information," *EURASIP J. Applied Signal Processing*, vol. 2009, pp. 1–13, June 2009.
- [33] Y. Han, A. Pandharipande, and S. H. Ting, "Cooperative decode-and-forward relaying for secondary spectrum access," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4945–4950, Oct. 2009.
- [34] L. Huang, G. Zhu, X. Du, and K. Bian, "Stable multiuser channel allocations in opportunistic spectrum access," in *Proc. 2013 IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, 2013, pp. 1715–1720.
- [35] N. Namvar and F. Afghah, "Spectrum sharing in cooperative cognitive radio networks: A matching game framework," in *Proc. 49th Annu. Conf. Information Sciences and Systems (CISS)*, Baltimore, MD, 2015, pp. 1–5.
- [36] R. Mochaourab, B. Holfeld, and T. Wirth, "Distributed channel assignment in cognitive radio networks: Stable matching and walrasian equilibrium," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3924–3936, July 2015.
- [37] R. El-Bardan, W. Saad, S. Brahma, and P. K. Varshney, "Matching theory for cognitive spectrum allocation in wireless networks," in *Proc. Annu. Conf. Information Science and Systems (CISS)*, Princeton, NJ, 2016, pp. 466–471.
- [38] Z. Fu, Q. Guan, F. Ji, S. Jiang, and Y. Lu, "Stable user pairing in virtual MIMO systems: A game theoretical perspective," in *Proc. IEEE/CIC Int. Conf. Communications in China (ICCC)*, Xi'an, China, 2013, pp. 430–435.
- [39] A. El-Hajj, Z. Dawy, and W. Saad, "A stable matching game for joint uplink/downlink resource allocation in OFDMA wireless networks," in *Proc. IEEE Int. Conf. Communications (ICC)*, Ottawa, Canada, 2012, pp. 5354–5359.
- [40] B. Kauffmann, F. Baccelli, A. Chaintreau, V. Mhatre, K. Papagiannaki, and C. Diot, "Measurement-based self organization of interfering 802.11 wireless access networks," in *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, Anchorage, AK, 2007, pp. 1451–1459.
- [41] J. Huang and A. Swindlehurst, "Cooperative jamming for secure communications in MIMO relay networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4871–4884, Oct. 2011.
- [42] E. A. Jorswieck and R. Mochaourab, "Secrecy rate region of MISO interference channel: Pareto boundary and non-cooperative games," in *Proc. Int. ITG Workshop on Smart Antennas (WSA)*, Berlin, Germany, 2009, pp. 1–8.
- [43] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, July 2010.
- [44] J. Huang and A. Swindlehurst, "Robust secure transmission in MISO channels based on worst-case optimization," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1696–1707, Apr. 2012.
- [45] Z. Han, N. Marina, M. Debbah, and A. Hjørungnes, "Improved wireless secrecy rate using distributed auction theory," in *Proc. Int. Conf. Mobile Ad-Hoc and Sensor Networks*, Wu Yi Mountain, China, 2009, pp. 442–447.
- [46] R. Zhang, L. Song, Z. Han, and B. Jiao, "Improve physical layer security in cooperative wireless network using distributed auction games," in *Proc. IEEE Conf. Computer Communications (INFOCOM)*, Shanghai, China, 2011, pp. 18–23.
- [47] F. Pantisano, M. Bennis, W. Saad, R. Verdone, and M. Latva-Aho, "Coalition formation games for femtocell interference management: A recursive core approach," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Cancun, Mexico, 2011, pp. 1161–1166.
- [48] J. Hoydis and M. Debbah, "Green, cost-effective, small cell networks," *IEEE COMSOC MMT E-Letter*, vol. 5, no. 5, pp. 23–26, Sept. 2010.
- [49] D. Lopez-Perez, A. Valcarce, G. de la Roche, and J. Zhang, "OFDMA femtocells: A roadmap on interference avoidance," *IEEE Commun. Mag.*, vol. 47, no. 9, pp. 41–48, Sept. 2009.
- [50] V. Chandrasekhar and J. Andrews, "Uplink capacity and interference avoidance for two-tier femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3498–3509, July 2009.
- [51] S. Carlaw, A. Giustina, R. R. Bhat, V. S. Rao, R. Sieberg, and S. R. Saunders, *Femtocells: Opportunities and Challenges for Business and Technology*. Hoboken, NJ: Wiley, 2009.
- [52] Z. Han and K. J. Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [53] S. Bayat, R. Y. H. Louie, Z. Han, Y. Li, and B. Vucetic, "Multiple operator and multiple femtocell networks: Distributed stable matching," in *Proc. 2012 IEEE Int. Conf. Communications (ICC)*, Ottawa, pp. 5140–5145.
- [54] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, June 2013.
- [55] S. Lien, T. Liau, C. Kao, and K. Chen, "Cooperative access class barring for machine-to-machine communications," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 27–32, Jan. 2012.
- [56] 3GPP, "System improvement for machine-type communications," 3GPP, Sophia Antipolis, France, TR 23.888 V1.6.1, Sept. 2012.
- [57] S. Lien and K. Chen, "Massive access management for QoS guarantees in 3GPP machine-to-machine communications," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 311–313, July 2011.
- [58] S. Shin and D. Triwacskono, "Radio resource control scheme for machine-to-machine communication in LTE infrastructure," in *Proc. 2012 Int. Conf. ICT Convergence (ICTC)*, Jeju Island, South Korea, pp. 1–6.
- [59] S. Bayat, Y. Li, Z. Han, M. Dohler, and B. Vucetic, "Distributed massive wireless access for cellular machine-to-machine communication," in *Proc. IEEE Int. Conf. Communications (ICC)*, Sydney, Australia, 2014, pp. 2767–2772.
- [60] N. Namvar, W. Saad, B. Maham, and S. Valentin, "A context-aware matching game for user association in wireless small cell networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, 2014, pp. 439–443.
- [61] F. Pantisano, M. Bennis, W. Saad, S. Valentin, M. Debbah, and A. Zappone, "Proactive user association in wireless small cell networks via collaborative filtering," in *Proc. Asilomar Con. Signals, Systems and Computers*, Asilomar, CA, 2013, pp. 1601–1605.

Ed Richter and Arye Nehorai

Enriching the Undergraduate Program with Research Projects

EDITORS' NOTE

The article in this "SP Education" column is the second in a special series on hands-on and design projects for signal processing education. The authors from Washington University in St. Louis, Missouri, United States, share their experiences and best practices of introducing undergraduate research projects as a key element in the undergraduate education of their electrical and system engineering program. Many of the projects are related to signal processing. As we will see, the program has energized the students' understanding and interests toward signal processing and, more generally, electrical engineering, and contributed to an impressive record of attracting students to pursue the field. We hope you will find the article inspiring, and stay tuned for more articles in the series in upcoming issues.

—Hana Godrich, lead guest editor
and Kenneth Lam, area editor, columns and forums

their involvement in undergraduate research projects.

Over the years, we have streamlined our approach and developed the program described here that works well for students interested in gaining independent project experience at the undergraduate level. This approach includes both undergraduate participation in faculty research as well as projects with Ph.D. students and our Professor of Practice, Ed Richter, as mentors. Overall, we have had more than 200 students from the ESE Department and from other departments participate in our undergraduate research projects program.

ESE297: Introduction to Undergraduate Research Projects

Freshmen and sophomores who are interested in undergraduate research projects are encouraged to take our Introduction to Undergraduate Research Projects (ESE297) course [1]. This three-credit course provides exposure to signal processing topics and introduces the design process early in the curriculum. By introducing design earlier, the students are not only preparing for undergraduate research projects but also for our Capstone Design course. The students learn valuable teamwork skills by working in groups of two. They are guided through the design process for two signal processing projects. These projects were selected after a few early undergraduate researchers and their mentors worked on these projects for a few semesters.

While some engineering departments struggle with low enrollment today, the Preston M. Green Electrical and Systems Engineering (ESE) Department at Washington University in St. Louis, Missouri, more than tripled over the last ten years. This was done under the guidance of Dr. Arye Nehorai, who was appointed its chair in 2006. This growth, including 34% female enrollment, is mostly due to students joining our department during or after their freshman year. The growth was fueled by improvements to the undergraduate educational experience including creating an undergraduate research program, improving the instruction quality, revising old courses, creating new courses, modernizing

our laboratories, creating study abroad programs, and communicating the new opportunities in the ESE Department to the engineering student body. This article focuses on the undergraduate research program.

Introduction

The undergraduate research program was created in response to student interest and current trends in education and gave us an opportunity to add more exciting design elements to our curriculum. Our students gain valuable experience in research, design, algorithm development, project implementation, and verification. These are skills that students can share during job interviews for both academia and industry. Also, some of our undergraduate researchers have decided to pursue careers in academia as a direct result of

Digital Object Identifier 10.1109/MSP.2016.2601652
Date of publication: 4 November 2016

The first project is acoustical source localization using a four-element microphone array [2]. The students are guided through an algorithm for triangulating a source location using a four-element microphone array. Next, they develop the signal processing implementation using LabVIEW as the programming language and a simulator that we provide them. Once the triangulation works in simulation, they test their implementation on the actual acoustic array. In the next phase of the project, the students implement the triangulation using a central computer and two modified National Instruments Dani Robots, each equipped with Wi-Fi and a microphone pair as shown in Figure 1(a). In their final presentation, we move a chirping speaker around and their mobile robots automatically track it.

The second project in this course is a multidisciplinary electroencephalogram (EEG) signal processing application to

create a brain-computer interface (BCI) to control robot motion in one dimension. We use a commercial 14-channel EEG electrode array from Emotiv, portions of the BCI2000 framework, and MATLAB as the programming language. The students are guided through the algorithm development to measure the average signal power in 2-Hz bandwidths for the 14 electrodes and to build a classifier to detect a power change in the “best” channel/frequency pairs.

We start with the clinical data provided with the BCI2000 framework collected during an experiment where the subject was asked to raise his legs or to relax. The clinical data file contains the raw data from the EEG sensors as well as the timing information of the action and rest trials. The BCI2000 framework also provides the MATLAB functions to extract the raw data and timing information from the file. The students then implement the signal processing and

classification in MATLAB working on the clinical data. After they validate their implementation against the BCI2000 results for the same data, they collect their own data with the Emotiv headset while being prompted to clench their fists or rest. Using their collected data as the training set, the students design their own classifier using their “best” channel/frequency pairs. When they are satisfied with the receiver operating characteristic (ROC) curve of their system, they design/implement a real-time LabVIEW/MATLAB/Emotiv software development kit (SDK) application that processes the EEG data in real time and moves the mobile robot every time they clench their fists as shown in Figure 1(b) [3].

ESE497: Undergraduate Research Projects

The top students from ESE297 are encouraged to sign up for the course Undergraduate Research Projects (ESE497) [4]. However, ESE297 is not necessarily a prerequisite for ESE497. We also encourage the top third- and fourth-year students from across the engineering school to get involved in our undergraduate research projects. The students are required to work on their projects for at least ten hours per week for two credits, create a blog/webpage of their project, and participate in several poster sessions. The required presentations are a good experience for the students, where feedback is provided from their peers as well as from the faculty. Where applicable, we encourage our students to publish their projects in a peer-reviewed journal.

Students are given a choice of 50 projects [5] listed on our website that cover topics such as signal and image processing, machine learning, optimization, power and energy, controls and operations research. We interview each student to find a good match between his or her interest and the available projects. One of the themes of our projects is multidisciplinary robotic sensing [6], merging signal and image processing, robotics, and biomedical and environmental sensing. We are also flexible when an entrepreneurial-minded student comes to us with a good idea. Sometimes the student's ideas/projects are part of our design



(a)



(b)

FIGURE 1. Platforms used in ESE297: (a) the mobile microphone array and (b) a student controls a robot with EEG signals from the Emotiv headset.

competition, like our Engineering School Discovery Competition. Additionally, many of our younger faculty post projects directly related to their research interests to attract potential Ph.D. students. When the proposed projects are part of a faculty member's active research, that faculty member is the mentor and the students work on their projects in the faculty member's lab.

More commonly, we have more student interest than our faculty can support. For some of these cases, we carefully select Ph.D. students working in a related area to mentor these students. The undergraduates benefit from this because they get more one-on-one opportunities with a top Ph.D. student than they might otherwise get with a busy full-time faculty member. It is mutually beneficial for the Ph.D. student as they get an opportunity to work with our bright undergraduates. In many cases, the projects are multi-semester, and the more senior undergraduate students are also involved in the mentoring. Additionally, we have a Professor of Practice, Ed Richter, who mentors projects and supports other mentors in their efforts.

We have developed many platforms that students can use for their projects and we provide "Getting Started Guides" [7] for these platforms on our undergraduate research webpage. This infrastructure allows the students to focus on the signal processing algorithm development instead of spending the semester figuring out the low-level communication details to some new piece of hardware. Several of the platforms available to our students are listed below:

- 1) ESE297 students can continue to work with the mobile robots used in ESE297 to develop autonomous navigation algorithms.
- 2) ESE297 students can continue to work with the EEG headset to develop more advanced classification applications for BCI projects [8].
- 3) In 2012, we purchased an industrial robotic arm (Figure 2) [9] made by Fanuc, and we developed a communication scheme using the National Instruments Ethernet/IP tool kit that allows the students to position the robotic arm using LabVIEW or

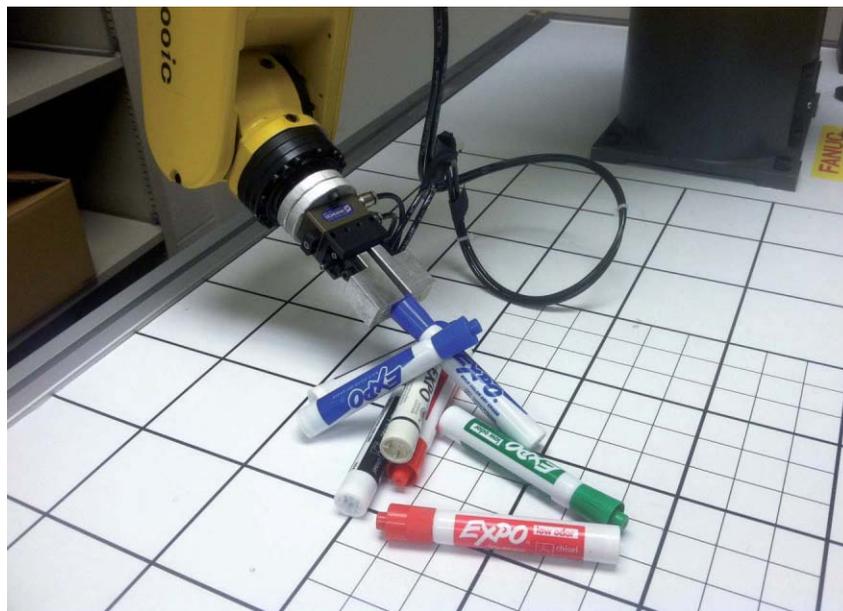


FIGURE 2. An industrial robotic arm for student projects.

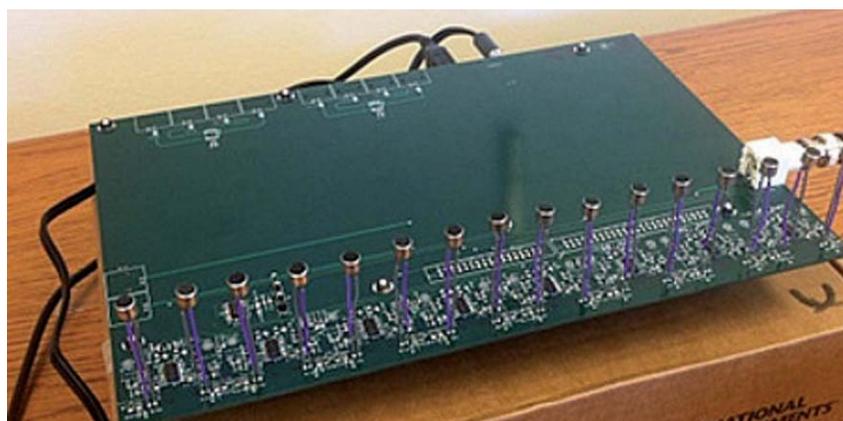


FIGURE 3. Sixteen channels of a 64-channel microphone array.

- 4) In 2010, we purchased several National Instruments USRP 2920 and USRP 2921 for wireless communication and array signal processing applications. Students have used this platform for localization and Doppler measure-

ments (video [10] by Prof. Ed Richter and Dr. Martin Hurtado).

- 5) In 2014, we purchased two inertial movement unit sensors [11] as part of our collaboration with researchers at KTH in Sweden. Students have used this platform to analyze tremors in patients who have Parkinson's disease.
- 6) In 2012, we developed a 64-channel microphone array as shown in Figure 3. Students can work with the array to develop audio beamforming applications (beamformed audio file created by Ricky Chen and Xiangyang Mou [12]).
- 7) Our labs are equipped with data acquisition systems that we use in our

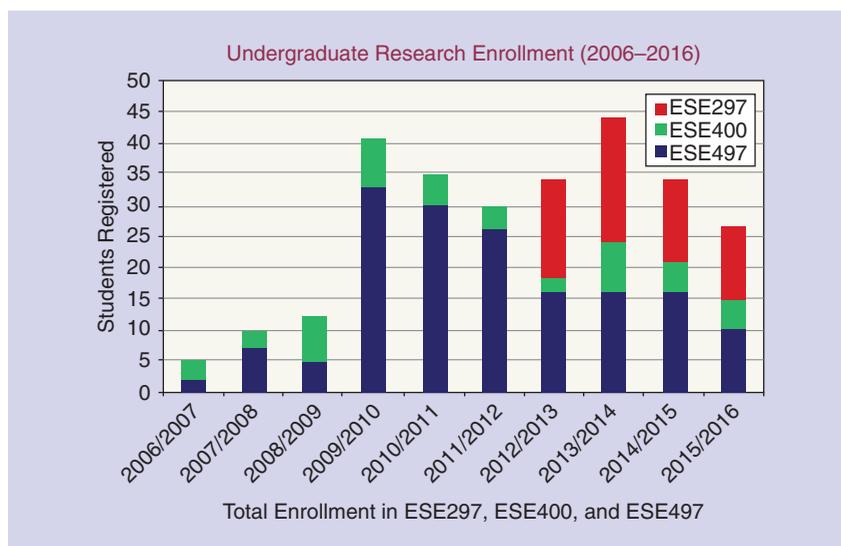


FIGURE 4. Total enrollment in ESE297, ESE400, and ESE497.

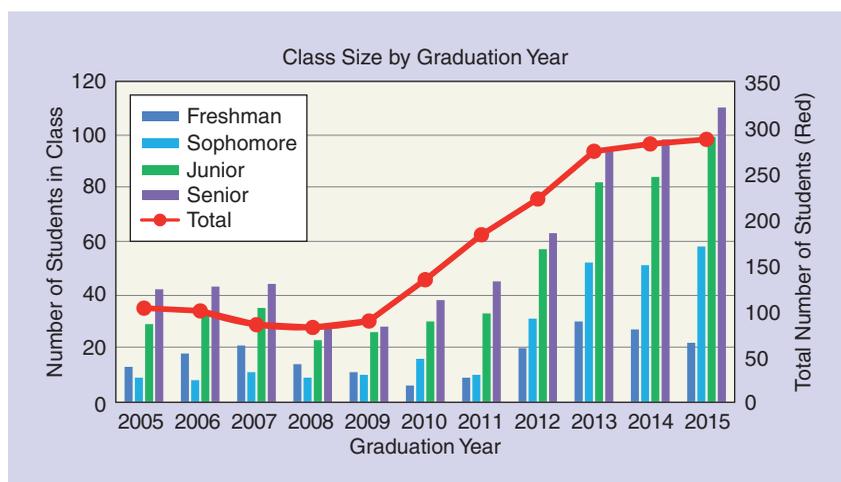


FIGURE 5. Class size by graduation year (total students on right y-axis).

courses but that are also available for undergraduate research projects that can use the computer as the digital signal processor. Real-time applications with data acquisition can be easily written in both LabVIEW and/or MATLAB.

Undergraduate research examples

Some of the more successful undergraduate research projects to date from the ESE497 research course are as follows:

- Yuni Teh (supervisors: Arye Nehorai and Ed Richter), Human-Machine Interface: Myoelectric Control Scheme to Restore Upper Extremity Motor Function, Spring 2015 [13]

- Ren Liang Liang, Tsinghua University (supervisors: Arye Nehorai, Martin Hurtado, and Ed Richter), Radio Tomographic Imaging-Based Fall Detection, Summer 2015 [14]
- Yifan Wang and Stephen Gower (supervisor: Arye Nehorai), Robotic Avoidance Using Kinect and UST, Summer 2015 [15]
- Daniel Wasseman (supervisor: Zachary Feinstein), Grid Search Algorithm for Set Optimization, Spring 2015 [16]
- David Sehloff and Celso Torres (supervisors: Arye Nehorai and Ed Richter), Predicting the Solar

Resource and Power Load, Spring 2015 [17].

ESE Department growth

We created our undergraduate research program in 2006 as part of the effort to attract more students to electrical and/or systems engineering. The result of these efforts has shown a big jump in undergraduate research enrollment in 2009/2010 shown in Figure 4. This is one of the reasons our department has tripled in size since 2008 as shown in the bold red line in Figure 5.

Other factors contributing to the growth of the ESE Department include improving the instruction quality, modernizing our laboratories, adding introductory courses in electrical and systems engineering, and offering study-abroad opportunities. Additionally, we have improved our outreach to attract the undecided students and inform students from other departments about our unique programs. These presentations and web documents highlight our flexible curricula and give examples to show how to add a second major in electrical or systems engineering. One more way we are successful in increasing the enrollment is by interacting with students and getting their feedback. In particular, Dr. Nehorai contacts every student who decides to move to our department to understand his or her motivation for changing majors, and he has personally mentored more than 100 students in his lab over the last ten years. The success of these efforts is shown in Figure 5. Our entering class size is small (shown by the dark blue bar) but the class size has increased from freshman year (dark blue) to senior year (purple) every year since 2010 as students join our department after arriving at Washington University.

Conclusions

Student interest in undergraduate research projects continues to grow. We have tried to meet that demand by developing a “training” course (ESE297) and by offering many diverse topics and platforms for them to use

for their own projects. In general, the students find ESE297 and undergraduate research projects (ESE497) extremely rewarding, albeit challenging. The students really appreciate the opportunity to apply the theory they have been studying. The one-on-one mentoring with tenured, tenure-track faculty, Professor of Practice, or Ph.D. students is a valuable opportunity to our students. The feedback they get from their peers and the faculty about their presentations forces them to think about their work in a new light. The students can often showcase these projects in job interviews and graduate school applications. Our undergraduate research program is just one way that we have expanded the enrollment in the ESE Department over the last ten years.

Authors

Ed Richter (ed@ese.wustl.edu) is a Professor of Practice in the Preston M. Green Department of Electrical and Systems Engineering at Washington University in St. Louis. He brings 30 years of real-world engineering experience to the classroom and to undergraduate research projects.

Arye Nehorai (nehorai@ese.wustl.edu) is the Eugene and Martha Lohman Professor of Electrical Engineering in

the Preston M. Green Department of Electrical and Systems Engineering at Washington University in St. Louis. He served as chair of this department from 2006 to 2016. Under his leadership, the department's undergraduate enrollment has more than tripled in four years and the master's enrollment grew sevenfold in the same time period. He is a Fellow of the IEEE, Royal Statistical Society, and the American Association for the Advancement of Science.

For More Information

All references in this article can also be found by visiting <https://sites.wustl.edu/spmrefs2016>.

References

- [1] ESE297—Introduction to undergraduate research [Online]. Available: http://classes.engineering.wustl.edu/ese497/index.php/ESE297_-_Intro_to_Undergraduate_Research
- [2] Mobile microphone array video [Online]. Available: http://classes.engineering.wustl.edu/ese497/index.php/Project2:_Triangulation_with_sbRIO_robots
- [3] BCI robot control video [Online]. Available: http://classes.engineering.wustl.edu/ese497/index.php/BCI_Projects
- [4] ESE497: Undergraduate research projects [Online]. Available: http://classes.engineering.wustl.edu/ese497/index.php/Main_Page
- [5] Wash. Univ. St. Louis available undergraduate research projects [Online]. Available: http://ese.wustl.edu/undergraduateprograms/Pages/Available_UndergraduateResearchProjects.aspx

[6] Robotic sensing project [Online]. Available: http://classes.engineering.wustl.edu/ese497/images/3/3c/Robotic_Sensing_V4.pdf

[7] Getting started guides [Online]. Available: http://classes.engineering.wustl.edu/ese497/index.php/Main_Page#Getting_Started_Guides

[8] BCI hand control download

[9] Fanuc Robot control [Online]. Available: http://classes.engineering.wustl.edu/ese497/index.php/Matlab_Fanuc_Robot_Control

[10] Doppler shift velocity measurement with NI USRP [Online]. Available: http://classes.engineering.wustl.edu/ese497/index.php/Getting_Started_with_the_USRP

[11] Openshoe website [Online]. Available: <http://www.openshoe.org>

[12] Beamformed audio file [Online]. Available: <http://ese.wustl.edu/ContentFiles/Research/UndergraduateResearch/CompletedProjects/WebPages/su15/MouXiangyang/TwoSource.wav>

[13] Y. Teh. (2015). Human-machine interface: Myoelectric control scheme to restore upper extremity motor function, ESE 497 [Online]. Available: <https://sites.wustl.edu/spmrefs2016/yuniteh>

[14] L. L. Ren. (2015). Radio tomographic imaging-based fall detection, Tsinghua Univ., ESE 497 [Online]. Available: <https://sites.wustl.edu/spmrefs2016/renliangliang>

[15] Y. Wang and S. Gower (2015). Robotic avoidance using Kinect and UST, ESE 497 [Online]. Available: <https://sites.wustl.edu/spmrefs2016/wanggower>

[16] D. Wassemán. (2015). Grid search algorithm for set optimization, ESE 497 [Online]. Available: <https://sites.wustl.edu/spmrefs2016/danielwassemán>

[17] D. Sehloff and C. Torres. (2015). Predicting the solar resource and power load, ESE 497 [Online]. Available: <https://sites.wustl.edu/spmrefs2016/sehloffrorres>

SP

Sign Up or Renew Your 2017 SPS Memberships

A Membership in the IEEE Signal Processing Society (SPS), the IEEE's first society, can help you lay the foundation for many years of success ahead:

- **CONNECT** with more than 19,000 signal processing professionals through SPS conferences, and local events hosted by more than 170 SPS Chapters worldwide.
- **SAVE** with member discounts on conferences and publications, and access to travel grants, SigPort repository, and SPS Resource Center.
- **ADVANCE** with world-class educational resources, awards and recognitions, and society-wide volunteer opportunities in publications, conferences, membership, and more.



APPLICATIONS CORNER

Chen Chen, Yi Han, Yan Chen, and K.J. Ray Liu

Indoor Global Positioning System with Centimeter Accuracy Using Wi-Fi

The global positioning system (GPS) is a space-based navigation system that can provide location and time information whenever there is an unobstructed line of sight (LOS) to four or more GPS satellites [1]. Such a system provides critical capabilities to military, civil, and commercial applications around the world. On the other hand, considering the fact that people today spend more than 80% of their time in indoor environments, accurate indoor localization is highly desirable and has a great potential impact on many applications. Unfortunately, the use of GPS satellites to enable indoor localization is a nonstarter due to a variety of reasons including poor signal strength, multipath effect, and limited on-device computation and communication power [2]. Therefore, over the past two decades, the research community has

been urgently seeking new technologies that can enable high-accuracy indoor localization. However, the results are still mostly unsatisfactory. Microsoft hosted Indoor Localization Competitions in recent years and concluded that “the indoor location problem is not solved” [3].

Many indoor positioning systems (IPSS) have been developed by leveraging radio wave, magnetic field, acoustic signal, or other sensory information collected by mobile devices [4]. Most of these systems are based on the ranging technique. Ranging is a process to determine the distance from one location to another by utilizing the collected information such as the received signal strength indicator (RSSI) and/or time of arrival (TOA). Typically, these systems require multiple anchors at known locations and dedicated devices to collect fine-grained information for accurate ranging.

However, when there exist obstacles between the localized device and the anchors, the localization performance degrades significantly. In other words, the performance of ranging-based systems cannot maintain under non-line-of-sight (NLOS) scenarios, which is very common for an indoor environment. Such degradation is due to that the physical ranging rules that translate the collected information into the distances are impaired by the blockage and multipath components naturally existing indoor. Developing a general physical ranging rule that suits NLOS conditions is practically difficult, if not impossible, due to the complicated indoor environment, which motivates the development of the fingerprint-based IPSSs. A summary of the existing state-of-the-art capabilities from the Microsoft-hosted Indoor Localization Competitions is given in Figure 1, in

Digital Object Identifier 10.1109/MSP.2016.2600734
Date of publication: 4 November 2016

Technology	Existing Hardware?	Minimum Anchors	Low Cost?	Res. (m) LOS	Res. (m) NLOS	Commercial Examples
RSSI	✓	3	—	1–3	5–10	Active RFID, iBeacon, SPIRIT Navigation, Modulated LEDs
TOA TDOA	✗	3	✗	0.2–0.4	1–5	UWB, Decawave, Time Domain, Zebra, Nanotron
AOA	✓	2	—	0.4	1–5	None
Time Reversal	✓	1	✓	0.02	0.02	Origin Wireless

FIGURE 1. State-of-the-art IPSSs.

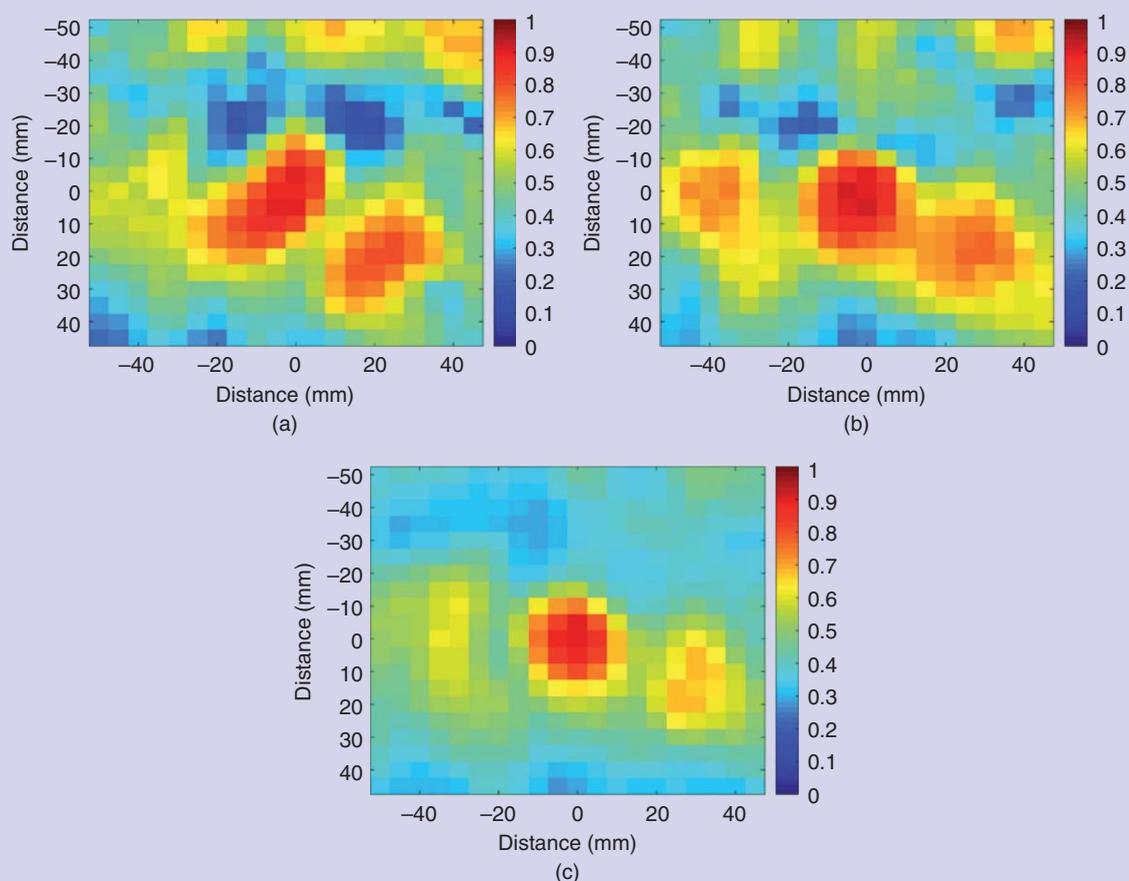


FIGURE 2. Ambiguity among nearby locations under (a) 40-MHz bandwidth, (b) 120-MHz bandwidth, and (c) 360-MHz bandwidth.

which one can see that under the LOS condition, with more than one anchor, submeter accuracy can be achieved. However, under the NLOS condition, only the meter range can be obtained by most methods, except the recently proposed time-reversal approach that can obtain 1–2 cm accuracy for both LOS and NLOS conditions [5].

In an indoor environment, there naturally exists some location-specific information, known as the fingerprints. Examples include the magnetic field, RSSI, and the channel state information (CSI). All of these fingerprints can be exploited for indoor localization. Specifically, in the fingerprint-based IPS, the location-specific fingerprints are collected and stored in a database in the mapping phase. Then, in the localization phase, the location of the device is determined by comparing the device fingerprint with those in the database.

In [5], it was shown that the physical phenomenon of the time-reversal focusing effect can provide a high-resolution fingerprint for indoor localization. The authors used a dedicated device to obtain the channel impulse response under the 5 GHz industrial, scientific, and medical radio (ISM) band with a bandwidth of 125 MHz as the fingerprint and utilized the time-reversal resonating strength (TRRS) as the similarity measure, which gives an accuracy of 1–2 cm.

The question now is: Can one use the ubiquitous Wi-Fi devices to achieve the same? The answer is yes as evidenced by the recent works in [6]–[8]. The work in [6] and [7] leveraged frequency hopping, while the work in [8] used multiantenna spatial diversity to increase the effective bandwidth. As a result, the localization resolution can be significantly improved to 1–2 cm.

This article will show the basic principles of how one can achieve indoor localization resolution down to the centimeter accuracy level using standard Wi-Fi devices. A unified view by combining both the frequency and spatial diversities is also presented.

How does bandwidth affect the localization performance?

The main reason that most of the fingerprint-based methods utilizing CSI in Wi-Fi systems cannot achieve centimeter localization accuracy is due to the bandwidth limitation. More specifically, the maximum bandwidth in mainstream Wi-Fi devices is only either 20 or 40 MHz, which introduces severe ambiguity into the fingerprints of different locations and thus leads to the poor accuracy for indoor localization.

To clearly illustrate the impact of bandwidth on localization performance, we have conducted experiments to collect CSIs under different bandwidths in a typical indoor environment. Two channel sounders are placed in an NLOS setting, where one of them is placed on a customized experiment structure with 5-mm resolution.

To characterize the similarity between CSIs collected at the same or different locations, the TRRS is calculated, given by

$$\gamma[\mathbf{H}, \mathbf{H}'] = \left(\frac{\eta}{\sqrt{\Lambda} \sqrt{\Lambda'}} \right)^2, \quad (1)$$

with

$$\eta = \max_{\phi} \left| \sum_{k=1}^K H_k H_k'^* e^{-jk\phi} \right|, \\ \Lambda = \sum_{k=1}^K |H_k|^2, \Lambda' = \sum_{k=1}^K |H_k'|^2, \quad (2)$$

where \mathbf{H} and \mathbf{H}' represent two fingerprints, K is the total number of usable subcarriers, H_k and H_k' are the CSIs on subcarrier k , η is the modified cross-correlation between \mathbf{H} and \mathbf{H}' with synchronization error compensated, and Λ, Λ' are the channel energies of \mathbf{H} and \mathbf{H}' , respectively. Realizing that the Wi-Fi receiver may not be fully synchronous with the Wi-Fi transmitter due to mismatches in their radio-frequency front-end components [9], an additional phase rotation of $e^{-jk\phi}$ is employed to counteract the phase distortions incurred by the synchronization errors in the calculation of η , where ϕ can be estimated and compensated using Algorithm 1 shown later in the section “Calculating Time-Reversal Resonating Strength by Diversity Exploitation.” Equation (1) implies that TRRS ranges from 0 to 1. More specifically, a larger TRRS indicates a higher similarity between two fingerprints and thus the two associated locations.

The corresponding TRRS between the target location and nearby locations are illustrated in Figure 2 under different bandwidth settings. It is shown in Figure 2(a) that with 40 MHz bandwidth, a large region of nearby locations is ambiguous with the target location in terms of the TRRS. Enlarging the bandwidth shrinks the area of ambiguous regions. As

demonstrated in Figure 2(c), when the bandwidth increases to 360-MHz, the ambiguous region is reduced to a ball of 1 cm radius, which implies centimeter accuracy in localization.

The experiment results motivate us to formulate a large effective bandwidth by exploiting diversities on Wi-Fi devices to facilitate centimeter accuracy indoor localization.

Increasing effective bandwidth via diversity exploitation

Two different diversities exist in the current Wi-Fi system, i.e., frequency diversity and spatial diversity. According to IEEE 802.11n, 35 Wi-Fi channels are dedicated to Wi-Fi transmission in 2.4- and 5-GHz-frequency bands with a maximum bandwidth of 40 MHz. The multitude of Wi-Fi channels leads to frequency diversity in that they provide opportunities for Wi-Fi devices to perform frequency hopping when experiencing deep fading or severe interference. On the other hand, spatial diversity can be exploited on multiple-input, multiple-output (MIMO) Wi-Fi devices, which is a mature technique that greatly boosts the spectral efficiency. MIMO has not only become an essential component of IEEE 802.11n/ac but also been ubiquitously deployed on numerous commercial Wi-Fi devices. For Wi-Fi systems, both types of diversity can be harvested to provide fingerprints with much finer granularity and thus lead to less ambiguity in comparison with the fingerprint measured with a bandwidth of only 40 MHz.

Figure 3 shows the general principle of creating a large effective bandwidth by exploiting the frequency and spatial diversities either independently or jointly. Since Wi-Fi devices can work on multiple Wi-Fi channels, one can exploit the frequency diversity by performing frequency hopping to obtain CSIs on different Wi-Fi channels. As demonstrated in Figure 3(a), CSIs on four different Wi-Fi channels are concatenated together to formulate a fingerprint of a large effective bandwidth. Despite the fact that the frequency diversity can be exploited on a single-antenna Wi-Fi device, it is

time-consuming to perform frequency hopping. For time efficiency, spatial diversity can be exploited on multi-antenna Wi-Fi devices. For a Wi-Fi receiver with four antennas, e.g., in Figure 3(b), CSIs on the four receiving antennas can be combined together to formulate the fingerprint with a large effective bandwidth. Figure 3(c) shows an example of utilizing both the frequency and spatial diversities, where CSIs on two Wi-Fi channels and from two receiving antennas are combined into the fingerprint.

For a Wi-Fi system, the spatial diversity is determined by the number of antenna links, while the frequency diversity is dependent on the number of available Wi-Fi channels. Denote the maximum spatial diversity by S , the maximum frequency diversity by F , and the bandwidth for each Wi-Fi channel by W , the effective bandwidth is calculated as $S \times F \times W$.

Achieving centimeter accuracy via TRRS

As discussed in the section “Increasing Effective Bandwidth via Diversity Exploitation,” a fine-grained fingerprint associated with a large effective bandwidth can be generated through diversity exploitation on Wi-Fi devices. In this section, we first introduce the calculation of TRRS when both of the frequency and spatial diversities are available. Then, we present the algorithm for indoor localization.

Calculating TRRS by diversity exploitation

As discussed in the sections “How Does Bandwidth Affect the Localization Performance?” and “Increasing Effective Bandwidth via Diversity Exploitation,” to achieve centimeter localization accuracy, a large effective bandwidth beyond 40 MHz is required, which can be obtained by diversity exploitation. For Wi-Fi devices with a spatial diversity of S and a frequency diversity of F , the CSI measurements can be written as $\bar{\mathbf{H}} = \{\mathbf{H}_{s,f}\}_{s=1,2,\dots,S}^{f=1,2,\dots,F}$, where $H_{s,f}$ stands for the CSI measured with the s th antenna link on the f th Wi-Fi channel, denoted as the virtual link (s, f) .

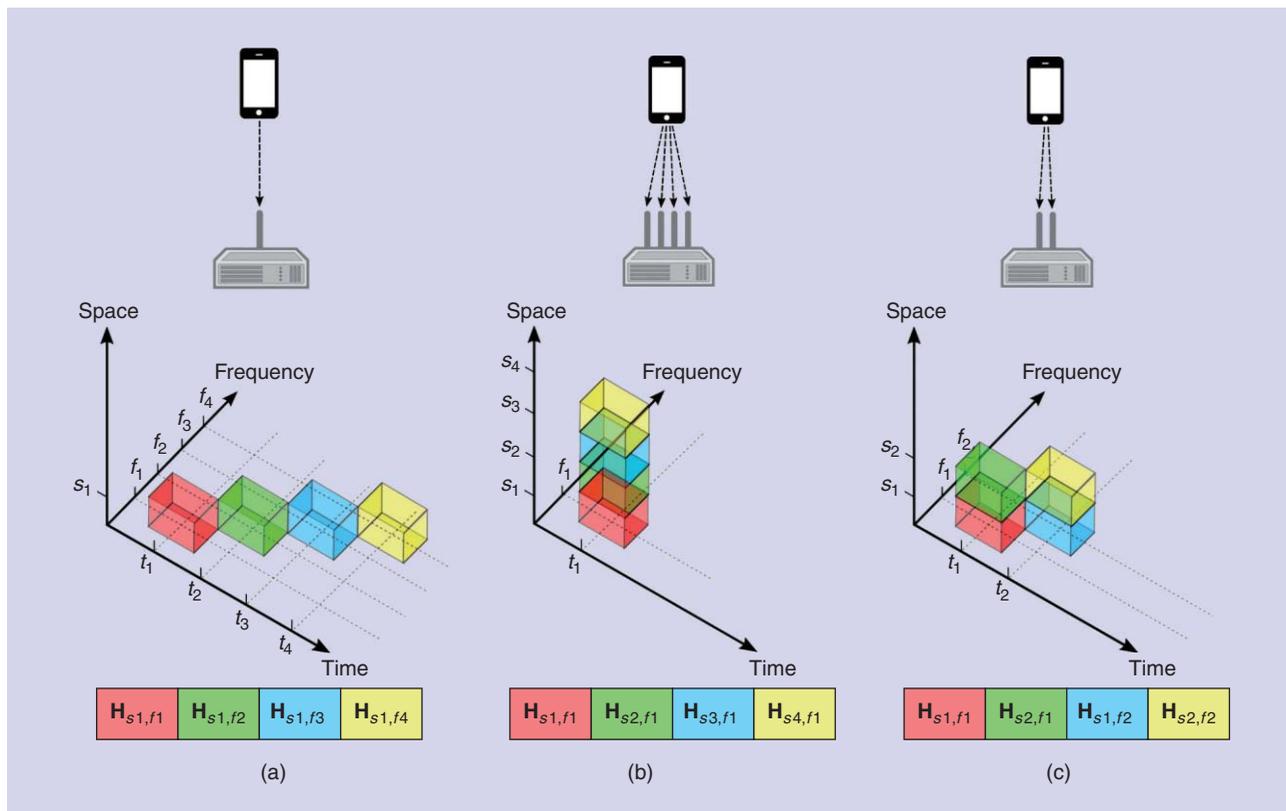


FIGURE 3. Leveraging frequency and spatial diversities in Wi-Fi to achieve large effective bandwidth: (a) using only the frequency diversity, (b) using only the spatial diversity, and (c) using both of the frequency and spatial diversities.

$\bar{\mathbf{H}} = \{\mathbf{H}_{s,f}\}_{s=1,2,\dots,S}^{f=1,2,\dots,F}$ can provide fine-grained fingerprint with an effective bandwidth of $S \times F \times W$. Consequently, TRRS in (1) can be extended to the fine-grained fingerprint $\bar{\mathbf{H}}$ and $\bar{\mathbf{H}}'$, with η and Λ, Λ' modified as

$$\eta = \sum_{s=1}^S \sum_{f=1}^F \eta_{s,f},$$

$$\Lambda = \sum_{s=1}^S \sum_{f=1}^F \Lambda_{s,f}, \quad \Lambda' = \sum_{s=1}^S \sum_{f=1}^F \Lambda'_{s,f}, \quad (3)$$

where

$$\eta_{s,f} = \max_{\phi} \left| \sum_{k=1}^K H_{s,f,k} H_{s,f,k}^* e^{-jk\phi} \right| \quad (4)$$

represents the modified cross-correlation on the virtual link (s,f) , and $\Lambda_{s,f} = \sum_{k=1}^K |H_{s,f,k}|^2$, $\Lambda'_{s,f} = \sum_{k=1}^K |H'_{s,f,k}|^2$ are the channel energies of $\mathbf{H}_{s,f}$ and $\mathbf{H}'_{s,f}$ on the virtual link (s,f) , respectively.

Algorithm 1 elaborates on the calculation of $\gamma[\bar{\mathbf{H}}, \bar{\mathbf{H}}']$. As shown in Algorithm 1, steps 4–9 are used to calculate

the channel energies on the virtual link (s,f) , while steps 10–14 are targeted to compute the modified cross-correlation of two CSIs on the virtual link (s,f) . The channel energies and modified cross-correlation on each virtual link are accumulated as shown in step 9 and step 15, respectively. Finally, the TRRS is obtained by step 18. The computation of $\eta_{s,f}$ is approximated by $\tilde{\eta}_{s,f} = \max_n \left| \sum_{k=1}^K H_{s,f,k} H_{s,f,k}^* e^{-j\frac{2\pi nk(n-1)}{N}} \right|$ that takes the same format of a discrete Fourier transform of size N and thus can be computed efficiently by fast Fourier transform. Using a large N in the computations leads to a more accurate approximation of $\eta_{s,f}$.

Localization using TRRS

There are two phases in the proposed IPS: a mapping phase and a localization

phase. During the mapping phase, the CSIs are collected from L locations-of-interest using Wi-Fi devices with S antenna links and across F Wi-Fi channels, denoted by $\{\bar{\mathbf{H}}_\ell\}_{\ell=1,2,\dots,L}$. In the localization phase, $\bar{\mathbf{H}}'$ is obtained at a testing location, which may either be one of the L locations-of-interest or an unmapped location in the mapping phase. Then, the pairwise TRRS $\gamma[\bar{\mathbf{H}}_\ell, \bar{\mathbf{H}}']$, is calculated for all locations-of-interest. Finally, the location is determined based on $\gamma[\bar{\mathbf{H}}_\ell, \bar{\mathbf{H}}']$, i.e., (5), shown in the box at the bottom of the page, where Γ is a threshold introduced to balance off the true positive rate and false positive rate in location determination. When $\gamma[\bar{\mathbf{H}}_\ell, \bar{\mathbf{H}}']$ falls below Γ , the IPS cannot obtain a credible location estimation and returns 0 to imply an unmapped location.

$$\hat{\ell} = \begin{cases} \operatorname{argmax}_{\ell=1,2,\dots,L} \gamma[\bar{\mathbf{H}}_\ell, \bar{\mathbf{H}}'], & \text{If } \max_{\ell=1,2,\dots,L} \gamma[\bar{\mathbf{H}}_\ell, \bar{\mathbf{H}}'] \geq \Gamma \\ 0, & \text{Otherwise} \end{cases} \quad (5)$$



FIGURE 4. The Wi-Fi device used in the proposed IPS.

Experiment results

Extensive experiments are conducted to validate the theoretical analysis and evaluate the performance of the proposed IPS. The proposed system contains two Wi-Fi devices, each equipped with three omnidirectional antennas. One Wi-Fi device, called Origin, estimates CSI from the other Wi-Fi

device, named the Bot. With the proposed algorithm in the section “Achieving Centimeter Accuracy Via TRRS,” the Origin estimates the location of the Bot. Figure 4 shows one Wi-Fi device used in the proposed IPS.

The experiments are conducted in a typical office of a multistory building. The indoor space is filled with a large number of reflectors, e.g., chairs, desks, shelves, sofas, walls, and ceilings. The CSIs of 50 candidate locations are measured, with 20 measurements for each location.

To evaluate the performance, the CSIs at each location are partitioned into a training set and a testing set, with ten CSIs for each. The TRRS matrix is calculated using the CSIs collected at the 50 candidate locations. Each element of the matrix represents the TRRS between the CSIs at the training location and the testing location. In other words, the

diagonal elements of the matrix indicate the similarity between CSIs at the same location, while the off-diagonal elements stand for the similarity between CSIs of different locations.

Figure 5 illustrates the TRRS matrices under effective bandwidths of 10, 40, 120, and 360 MHz. First of all, it is easily seen from Figure 4 that the diagonal elements of the matrices are close to one, signifying high similarities among CSIs of the same locations. Regarding the off-diagonal elements, they become smaller with an increasing effective bandwidth. When the effective bandwidth is small, e.g., 10 MHz, some off-diagonal elements are even larger than the diagonal elements, giving rise to localization errors. In other words, it is very likely to localize the Bot to incorrect positions when the effective bandwidth is small. When the effective bandwidth is increased, the gap between diagonal and off-diagonal elements enlarges, which provides a clear watershed between the correct and incorrect locations and leads to an enhanced system performance in return.

To provide a statistical point of view, Figure 6 shows the cumulative density functions of the diagonal and off-diagonal elements in TRRS matrices under a variety of effective bandwidths. As we can see, the gap between the diagonal and off-diagonal elements increases with the effective bandwidth, indicating a better distinction between different locations. Whenever there is a gap between the diagonal and off-diagonal elements, a perfect localization can be achieved with an appropriate threshold, i.e., 100% true positive rate and 0% false positive rate.

In a practical indoor environment, there usually exists environment dynamics that might degrade the localization performance. To evaluate the proposed IPS in a dynamic indoor environment, the testing CSIs are recollected in the presence of human activities and large object movement. In particular, to emulate dynamics from human activities, one participant was asked to walk continuously in the vicinity of the Bot. Then, the participant was asked to open and close a door that blocks the direct link between the Origin and Bot so as

Algorithm 1. Calculating TRRS by exploiting diversities.

Input: $\bar{\mathbf{H}} = \{\mathbf{H}_{s,f}\}_{s=1,2,\dots,S}^{f=1,2,\dots,F}$, $\bar{\mathbf{H}}' = \{\mathbf{H}'_{s,f}\}_{s=1,2,\dots,S}^{f=1,2,\dots,F}$

Output: $\gamma[\bar{\mathbf{H}}, \bar{\mathbf{H}}']$

```

1:  $\Lambda = 0, \Lambda' = 0, \eta = 0$ 
2: for  $s = 1, 2, \dots, S$  do
3:   for  $f = 1, 2, \dots, F$  do
4:      $\Lambda_{s,f} = 0, \Lambda'_{s,f} = 0$ 
5:     for  $k = 1, 2, \dots, K$  do
6:        $\Lambda_{s,f} \leftarrow \Lambda_{s,f} + |H_{s,f,k}|^2$ 
7:        $\Lambda'_{s,f} \leftarrow \Lambda'_{s,f} + |H'_{s,f,k}|^2$ 
8:     end for
9:      $\Lambda \leftarrow \Lambda + \Lambda_{s,f}, \Lambda' \leftarrow \Lambda' + \Lambda'_{s,f}$ 
10:    for  $n = 1, 2, \dots, N$  do
11:       $z[n] \leftarrow \sum_{k=1}^N H_{s,f,k} H_{s,f,k}^* e^{-j\frac{2\pi n(k-1)}{N}}$ 
12:    end for
13:     $n^* = \operatorname{argmax}_{n=1,2,\dots,N} |z[n]|$ 
14:     $\tilde{\eta}_{s,f} = z[n^*]$ 
15:     $\eta \leftarrow \eta + \tilde{\eta}_{s,f}$ 
16:  end for
17: end for
18:  $\gamma[\bar{\mathbf{H}}, \bar{\mathbf{H}}'] \leftarrow \left(\frac{\eta}{\sqrt{\Lambda}\sqrt{\Lambda'}}\right)^2$ 

```

} Calculating channel energies on virtual link (s, f)

} Calculating modified cross-correlation on virtual link (s, f)

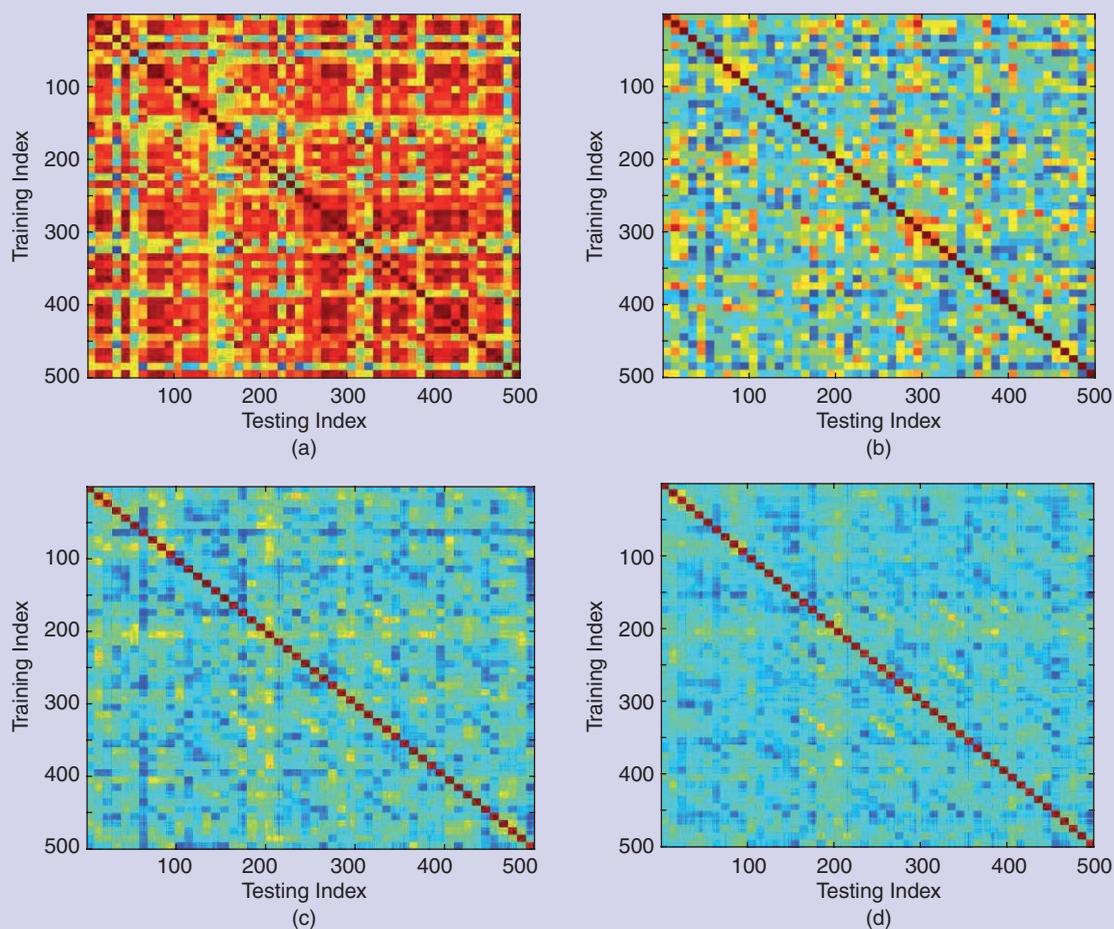


FIGURE 5. The TRRS matrix under an effective bandwidth of (a) 10 MHz, (b) 40 MHz, (c) 120 MHz, and (d) 360 MHz.

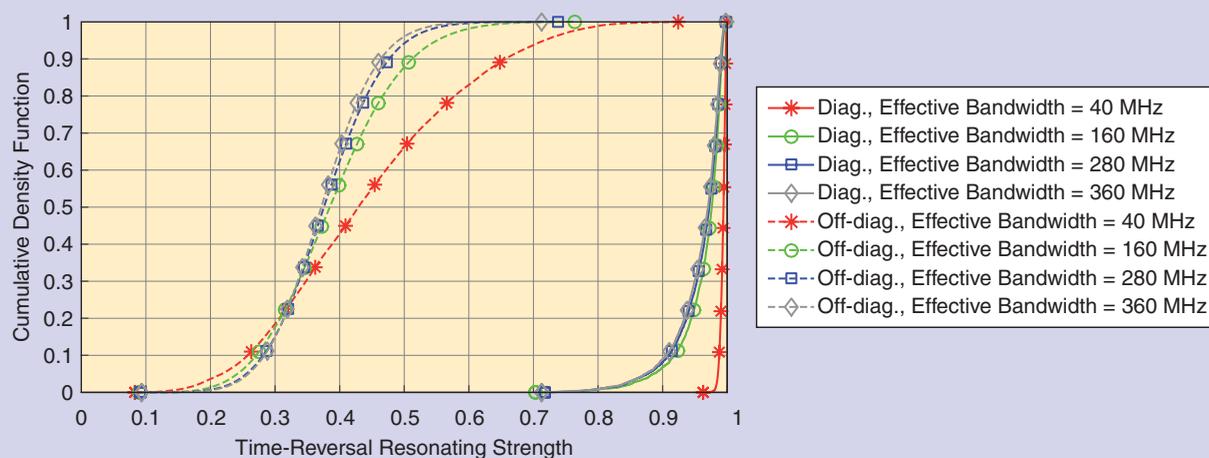


FIGURE 6. Cumulative density functions of the TRRS of the diagonal and off-diagonal elements.

to emulate the dynamics from large object movement.

Figure 7(a) demonstrates the receiver operating characteristic curve with

human activities. For a fixed false positive rate 0.15%, the true positive rate increases from 94.17% with 40-MHz effective bandwidth to 99.11% with

120-MHz effective bandwidth. Further enlarging the effective bandwidth to 240 MHz and 360 MHz boosts the true positive rate to 99.61% and 99.89%,

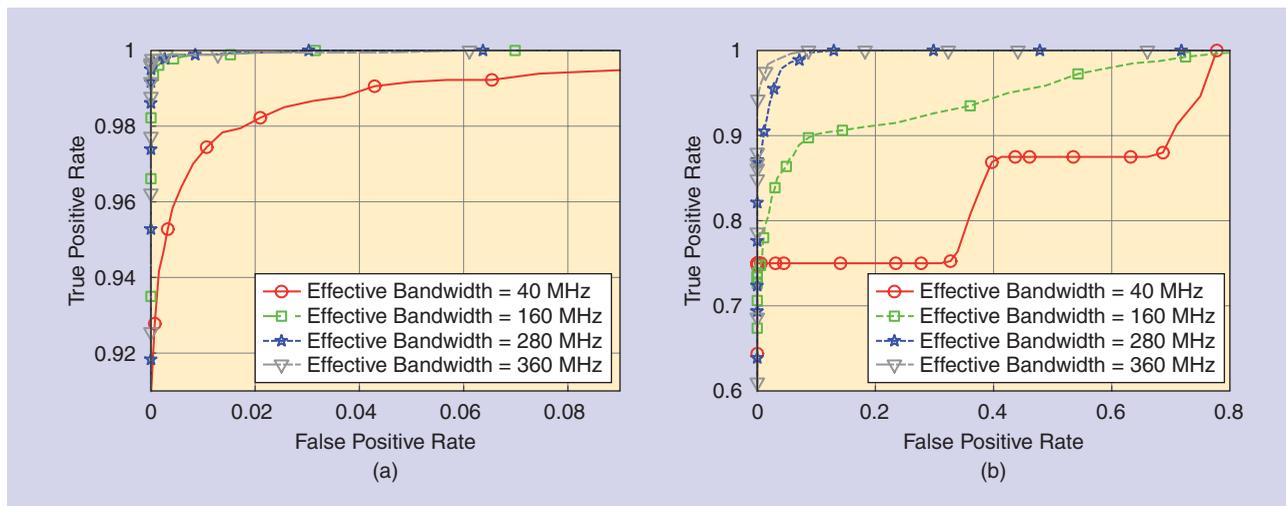


FIGURE 7. The receiver operating characteristic curve in the presence of environment dynamics: (a) dynamics from human movement and (b) dynamics from door opening and closing.

respectively. On the other hand, Figure 7(b) depicts the receiver operating characteristic curve with large object movement. For a fixed false positive rate 0.15%, the true positive rate increases from 75% with 40 MHz effective bandwidth to 76.38, 87.12, and 95% with 120, 240, and 360 MHz effective bandwidths, respectively. This can be justified by that, with a large effective bandwidth, the environment dynamics only affect very limited information in the fingerprint while leaving the majority intact. In other words, a large effective bandwidth enhances the robustness of the proposed IPS against environment dynamics.

During the experiments, we observe multiple Wi-Fi access points coexisting with the proposed IPS. However, their impact on the proposed IPS is minimal. Due to the carrier sense multiple access with collision avoidance (CSMA/CA) mechanism, these Wi-Fi access points would not transmit signals whenever they sense that the IPS is operating to obtain CSIs, which guarantees that the captured CSIs are free from interference.

Conclusions

In this column, we present a time-reversal method for indoor localization that achieves centimeter accuracy with a single pair of off-the-shelf Wi-Fi devices. The high accuracy for localization is maintained under strong NLOS

scenarios. With the exploitation of the inherent frequency and spatial diversities in Wi-Fi systems, it is capable of creating a large effective bandwidth to enable centimeter accuracy. Extensive experiment results in a typical office environment show that the centimeter accuracy as well as robustness against dynamics can be simultaneously achieved with a large effective bandwidth. The global GPS can achieve 3–15 m of accuracy by mapping the world into latitude and longitude coordinates. The presented “indoor GPS” can achieve 1–2 cm accuracy when an indoor environment is fingerprinted and mapped.

Authors

Chen Chen (cc8834@umd.edu) is a graduate student of the Department of Electrical and Computer Engineering, University of Maryland, College Park. He is also affiliated with Origin Wireless.

Yi Han (yi.han@originwireless.net) received his Ph.D. degree from the Department of Electrical and Computer Engineering, University of Maryland, College Park. Currently, he is the wireless architect of Origin Wireless.

Yan Chen (eecyan@uestc.edu.cn) is a professor with the School of Electronic Engineering, University of Electronic Science and Technology of China. He was a principal technologist of Origin Wireless.

K.J. Ray Liu (kjrlu@umd.edu) is the founder of Origin Wireless. He is also the Christine Kim Eminent Professor of Information Technology at the University of Maryland, College Park.

References

- [1] J. G. McNeff, “The global positioning system,” *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 645–652, Mar. 2002.
- [2] S. Nirjon, J. Liu, G. DeJean, B. Priyantha, Y. Jin, and T. Hart, “Coin-GPS: Indoor localization from direct GPS receiving,” in *Proc. ACM 12th Annu. Int. Conf. Mobile Systems, Applications, and Services*, 2014, pp. 301–314.
- [3] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, S. Sen, and V. Handziski, “Microsoft Indoor Localization Competition: IPSN 2016,” *GetMobile: Mobile Comp. and Comm.* [Online]. Available: <http://doi.acm.org/10.1145/2721914.2721923>
- [4] K. Curran, E. Furey, T. Lunney, J. Santos, D. Woods, and A. McCaughey, “An evaluation of indoor location determination technologies,” *J. Location Based Services*, vol. 5, no. 2, pp. 61–78, 2011.
- [5] Z.-H. Wu, Y. Han, Y. Chen, and K. J. R. Liu, “A time-reversal paradigm for indoor positioning system,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1331–1339, 2015.
- [6] C. Chen, Y. Chen, H. Q. Lai, Y. Han, and K. J. R. Liu, “High accuracy indoor localization: A WiFi-based approach,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2016, pp. 6245–6249.
- [7] C. Chen, Y. Chen, Y. Han, H. Lai, and K. J. R. Liu, “Achieving centimeter accuracy indoor localization on Wi-Fi platforms: A frequency hopping approach,” *IEEE Trans. Internet Things*, to be published.
- [8] C. Chen, Y. Chen, Y. Han, H. Lai, F. Zhang, and K. J. R. Liu, “Achieving centimeter accuracy indoor localization on Wi-Fi platforms: A multi-antenna approach,” *IEEE Trans. Internet Things*, to be published.
- [9] M. Speth, S. Fechtel, G. Fock, and H. Meyr, “Optimum receiver design for wireless broad-band systems using OFDM—Part I,” *IEEE Trans. Commun.*, vol. 47, pp. 1668–1677, Nov. 1999.

James Folberth and Stephen Becker

Efficient Adjoint Computation for Wavelet and Convolution Operators

First-order optimization algorithms, often preferred for large problems, require the gradient of the differentiable terms in the objective function. These gradients often involve linear operators and their adjoints, which must be applied rapidly. We consider two example problems and derive methods for quickly evaluating the required adjoint operator. The first example is an image deblurring problem where we must compute efficiently the adjoint of multistage wavelet reconstruction. Our formulation of the adjoint works for a variety of boundary conditions, which allows the formulation to generalize to a larger class of problems. The second example is a blind channel estimation problem taken from the optimization literature where we must compute the adjoint of the convolution of two signals. In each example, we show how the adjoint operator can be applied efficiently while leveraging existing software.

Prerequisites

The reader should be familiar with linear algebra, wavelets, and basic Fourier analysis. Knowledge of first-order iterative optimization algorithms is beneficial for motivating the need for a fast adjoint computation but should not be necessary to understand the adjoint computation itself.

Digital Object Identifier 10.1109/MSP.2016.2594277
Date of publication: 4 November 2016

Motivation

Many estimation problems are modeled using the composition of a differentiable loss ℓ with an affine map $x \mapsto \mathcal{A}x + b$, the simplest such examples being generalized linear models such as least-square data fitting. The gradient of $\ell \circ \mathcal{A}$ is given by the chain rule as $\mathcal{A}^* \nabla \ell(\mathcal{A}x + b)$. For small problems, \mathcal{A} may be explicitly represented as a matrix and \mathcal{A}^* calculated as the conjugate transpose of the matrix.

In contrast, larger problems are often carefully formulated to involve only linear operators \mathcal{A} , which have a fast transform [e.g., fast Fourier transform (FFT), wavelet transform, convolution] and are therefore never explicitly stored as matrices. While many software packages provide fast routines to compute \mathcal{A}^{-1} or the pseudoinverse \mathcal{A}^\dagger , few packages include routines for \mathcal{A}^* .

The purpose of this lecture note is to describe a few cases where one can compute the action of \mathcal{A}^* with the same complexity as applying \mathcal{A} . Boundary conditions play a confounding role in the calculation, and we present a simple framework to correctly take them into account.

Example 1: An image deblurring problem

Digital images can be blurred through a variety of means. For example, the optical system can be out of focus and/or atmospheric turbulence can cause

blurring of astronomical images. The goal of image deblurring is to recover the original, sharp image. One class of techniques poses the deblurring problem as an optimization problem, where the optimization variables are the wavelet coefficients of the recovered image. We assume that the blurring operator \mathcal{R} is known. For instance, \mathcal{R} could represent a Gaussian point spread function (PSF) under symmetric (reflexive) boundary conditions. The blurring operator is usually singular or ill conditioned [1].

Let \mathcal{W} represent a multilevel wavelet reconstruction operator with suitable boundary conditions, b be the observed, blurry image, and $\mathcal{A} = \mathcal{R}\mathcal{W}$ be the linear operator that synthesizes an image from wavelet coefficients x and then blurs the synthesized image under the blurring operator \mathcal{R} . To both regularize the problem and take into account that many real-world images have a sparse wavelet representation, we solve the following standard model, which seeks to recover sparse wavelet coefficients that accurately reconstruct the blurred image:

$$\min_x \frac{1}{2} \|\mathcal{R}\mathcal{W}x - b\|_2^2 + \lambda \|x\|_1, \quad \lambda > 0. \quad (1)$$

This commonly used formulation [2] is sometimes called the *synthesis formulation* since we reconstruct an image $\mathcal{W}x$ from the wavelet coefficients in the optimization variable x . We seek to find an

$$(\mathcal{E}^\dagger)^* = \begin{bmatrix} 0 & 0 & \frac{1}{2}I_{L' \times L'} \\ \frac{1}{2}I_{L' \times L'} & 0 & 0 \\ 0 & I_{N-2L' \times N-2L'} & 0 \\ 0 & 0 & \frac{1}{2}I_{L' \times L'} \\ \frac{1}{2}I_{L' \times L'} & 0 & 0 \end{bmatrix}$$

Again, in the unusual case that $N \leq 2L'$, the form of the pseudoinverse changes slightly. By viewing other signal extension operators as matrices acting on signal vectors, one can readily find the appropriate $(\mathcal{E}^\dagger)^*$.

The adjoint for orthogonal wavelets

For orthogonal wavelets (e.g., Haar and more generally Daubechies wavelets), the adjoint \mathcal{W}_{zpd}^* is exactly the analysis operator $\mathcal{W}_{zpd}^\dagger$. For orthogonal wavelets with general boundary conditions, we can use the splitting (2) and implement $(\mathcal{E}^\dagger)^*$.

For an example image, consider a standard resolution test chart [6], shown in Figure 1. We prescribe symmetric boundary conditions on the image, which will affect both \mathcal{W} , \mathcal{R} , and their adjoints. Symmetric boundary conditions (compared to periodic or zero-padded boundary conditions) produce significantly fewer edge effects, and so they are a natural choice.

Fortuitously, for orthogonal wavelets, the adjoint of \mathcal{W} with symmetric boundary conditions is “very close to” the analysis operator \mathcal{W}^\dagger . This fact is often used (e.g., [2]), and as we can see from the splitting (2), the error is introduced through $\mathcal{E} \approx (\mathcal{E}^\dagger)^*$. Again, by merely implementing $(\mathcal{E}^\dagger)^*$, we can use the true adjoint \mathcal{W}^* .

Figure 2 shows the deblurred image after 2,500 iterations of a fast iterative shrinkage-thresholding algorithm (FISTA), a fast proximal gradient method, developed by Beck and Teboulle [2]. The wavelet reconstruction operator \mathcal{W} is taken to be a three-stage Haar discrete wavelet transform with symmetric boundary conditions. We set $\lambda = 2 \times 10^{-5}$. We use both the true adjoint \mathcal{W}^* and the approximation \mathcal{W}^\dagger . Using the

true adjoint, the image reconstruction relative error (versus the unblurred image) is 8.91×10^{-4} and 31.8% of the coefficients are nonzero. Using the pseudoinverse approximation, the relative error is 8.91×10^{-4} and 32.1% of the coefficients are nonzero. In this case the pseudoinverse approximation works extremely well.

The adjoint for biorthogonal wavelets

For biorthogonal wavelets, we no longer have $\mathcal{W}_{zpd}^* = \mathcal{W}_{zpd}^\dagger$. Since we have relaxed ourselves to biorthogonal wavelets (which are nice for symmetric boundary conditions [7]), it is perhaps too much to ask that the adjoint of the primal wavelet reconstruction operator involves only the primal wavelets. Let us recall briefly the pertinent aspects of frames of \mathbb{R}^d . These facts are described in more detail and generality in [4].

Let $y \in \mathbb{R}^N$ be an arbitrary signal vector and $\{\phi_i\}_{i=1}^p, p \geq N$ be a set of vectors in \mathbb{R}^N . Define the analysis operator Φ as the $p \times N$ matrix

$$\Phi = \begin{bmatrix} \phi_1^T \\ \vdots \\ \phi_p^T \end{bmatrix}$$

If Φ has full rank, we say that $\{\phi_i\}$ is a frame and we call Φ the frame analysis operator. Henceforth we assume $\{\phi_i\}$ is a frame.

The product Φu computes the expansion coefficients of the signal $u \in \mathbb{R}^N$ in the frame $\{\phi_i\}$. The frame synthesis operator Φ^* constructs a vector in \mathbb{R}^N given some expansion coefficients. Since $\{\phi_i\}$ is assumed to be a frame, $\Phi^* \Phi$ is invertible and we may define the Moore–Penrose pseudoinverse Φ^\dagger , which implements reconstruction in the frame as $\Phi^\dagger = (\Phi^* \Phi)^{-1} \Phi$.

A dual frame can be associated with a (primal) frame by defining the dual frame vectors $\tilde{\phi}_i = (\Phi^* \Phi)^{-1} \phi_i$ for $i = 1, \dots, p$. We may define analogously the dual frame analysis operator $\tilde{\Phi}$ and dual frame synthesis operator $\tilde{\Phi}^*$.

A fundamental relationship between primal and dual frame operators is $\tilde{\Phi}^* = \Phi^\dagger$ (see [4, Theorem 5.5]). We can directly relate the frame operators

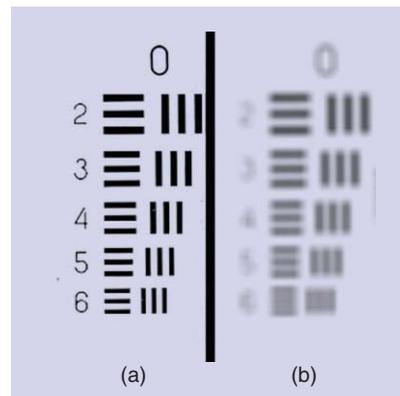


FIGURE 1. (a) The left half of the original, unblurred image and (b) the left half of the blurred image. To create the blurred image b , the blurring operator \mathcal{R} is applied to the original image and small amount of Gaussian noise is added.

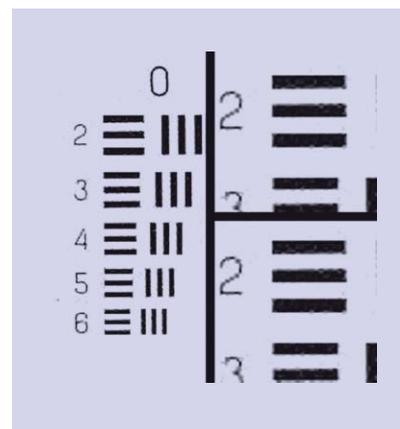


FIGURE 2. The deblurred image after 2,500 iterations of FISTA with \mathcal{W} a three-stage Haar transform with symmetric boundary conditions. The left half of the figure shows the deblurred image using the true adjoint \mathcal{W}^* ; the top right shows a zoomed-in view using \mathcal{W}^* ; the bottom right shows the same zoomed-in view using \mathcal{W}^\dagger and 2,500 iterations of FISTA.

and the wavelet operators under zero-padded boundary conditions. We can write zero-padded wavelet reconstruction as frame reconstruction: $\mathcal{W}_{zpd} = \Phi^\dagger$. Then using the above frame relations, we have

$$\mathcal{W}_{zpd}^* = (\Phi^\dagger)^* = (\Phi^*)^\dagger = \tilde{\Phi}. \quad (3)$$

We present other relationships between the primal and dual frames in Table 1. Note that $\Phi^\dagger \Phi = I$ but $\Phi \Phi^\dagger \neq I$ in general.

In (3), we have \mathcal{W}_{zpd}^* in terms of dual frame analysis. For biorthogonal wavelets, dual frame analysis corresponds

Table 1. A few primal and dual frame relationships.

	Primal Frame	Dual Frame
Analysis	$\Phi = \mathcal{W}_{zpd}^\dagger$	$\tilde{\Phi} = (\Phi^\dagger)^*$
Synthesis	$\Phi^* = (\mathcal{W}_{zpd}^\dagger)^*$	$\tilde{\Phi}^* = \Phi^\dagger$
Reconstruction	$\Phi^\dagger = \mathcal{W}_{zpd}$	$\tilde{\Phi}^\dagger = \Phi^*$

to wavelet analysis with the dual wavelets, which has a fast transform and is typically implemented together with the primal wavelets [3]. For biorthogonal (and orthogonal) wavelets, this states $\mathcal{W}_{zpd}^* = \overline{\mathcal{W}_{zpd}^\dagger}$, the latter of which is a standard operation and has a fast implementation. Combining this with the signal extension operator, we finally find

$$\mathcal{W}^* = \mathcal{W}_{zpd}^* (\mathcal{E}^\dagger)^* = \overline{\mathcal{W}_{zpd}^\dagger} (\mathcal{E}^\dagger)^*. \quad (4)$$

From this we see that with a complete wavelet software library, one only needs to implement $(\mathcal{E}^\dagger)^*$ to use the true adjoint instead of the approximation $\mathcal{W}^* \approx \overline{\mathcal{W}_{zpd}^\dagger}$.

Consider deblurring the resolution chart from Figure 1 with 2,500 iterations of FISTA using a three-stage Cohen–Daubechies–Feauveau (CDF) 9/7 discrete wavelet transform with symmetric boundary conditions. The results of deblurring are visually very similar to Figure 2. Using the true adjoint, the image reconstruction relative error (versus the unblurred image) is 9.45×10^{-4} and 29.23% of the coefficients are nonzero. Using

the pseudoinverse approximation, the relative error is 9.67×10^{-4} and 29.74% of the coefficients are nonzero. Here the pseudoinverse approximation works, but not quite as well as the true adjoint.

Figure 3 shows the structural similarity (SSIM) index [8] for the CDF 9/7 experiment using both the pseudoinverse approximation and true adjoint. For the first few hundred iterations of FISTA, the pseudoinverse approximation works well, but using the true adjoint appears to recover an image more similar to the original, unblurred image.

Example 2: Blind channel estimation

Another application that depends on a fast adjoint in the gradient computation is in some formulations of blind channel estimation. In blind channel estimation, a single source sends an unknown signal over multiple channels with unknown responses. Observers collect the output of each channel and collectively attempt to determine the source signal and the channel impulse response from each channel. Let s be the unknown source signal of length N and h_i the channel impulse response of the i th channel, each of length K . Then the output of the i th channel is

$$x_i = h_i * s,$$

where $*$ denotes a linear convolution. Here we prescribe zero-padded bound-

ary conditions and take x_i to be the “full” convolution of length $K + N - 1$. We hope to recover the source s and channel responses h_i from the output signals x_i . However, note that there is both a magnitude and phase ambiguity in h_i and s , since we can multiply h_i by $\alpha \neq 0$ and divide s by α , leaving x_i unchanged.

For simplicity, assume we have a single channel h with observed output x ; we will extend the problem to more channels in an example to follow. We can pose the blind deconvolution problem for a single channel as

$$\min_{h,s} \frac{1}{2} \|h * s - x\|_2^2 + \lambda_h \|h\|_1 + \lambda_s \|s\|_1, \quad (5)$$

where we include the terms $\lambda_h \|h\|_1$ and $\lambda_s \|s\|_1$ to promote sparsity in the recovered signals. To promote other structure in the recovered signals (assuming it is present in the unknown true signals), we can include other terms, such as $\|h\|_{TV}$, the 1-D total-variation (TV) seminorm.

Since (5) involves $h * s$, it involves products of the problem variables and so it is a nonconvex problem. However, note that the outer product matrix hs^T contains all of the terms used in the computation of the convolution $h * s$. Indeed, we may define a linear operator on $\mathcal{A}(hs^T)$ that computes the convolution: $\mathcal{A}(hs^T) = h * s$ [9]. Of course this is still a nonlinear operation, but the nonlinear operation is confined to the outer product hs^T . In the following gradient computations, we will implicitly be computing the adjoint \mathcal{A}^* .

Let $f(h, s) = 1/2 \|h * s - x\|_2^2$ be the differentiable part of the objective. First order methods require $\nabla_h f$ and $\nabla_s f$. Note that when computing $\nabla_h f$, we hold s fixed, and in doing so we may view the convolution $h * s$ as a linear operation on h . Similarly, when computing $\nabla_s f$, we hold h fixed. Using the results of [10], we may determine a fast method for evaluating these gradients. To illustrate the derivation, consider the case $K = 3$ and $N = 5$. Set $x_{est} = h * s$. We can write the convolution as a matrix-vector product in a couple different ways. The first,

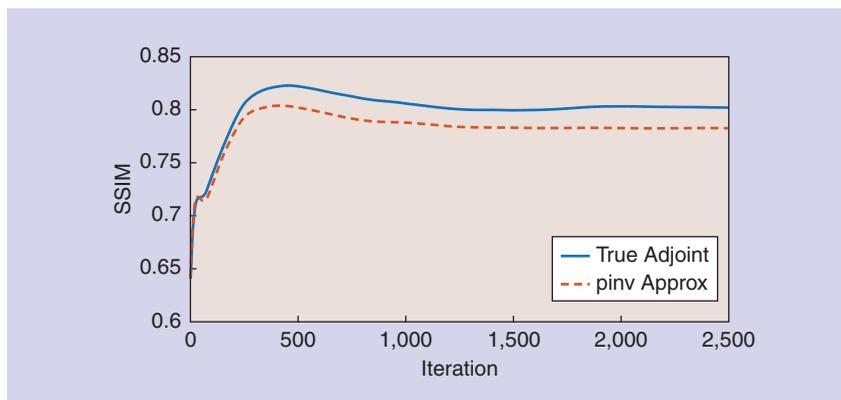


FIGURE 3. The SSIM index over 2,500 iterations of FISTA with \mathcal{W} a three-stage CDF 9/7 transform and using the true adjoint and the pseudoinverse approximation. SSIM compares the deblurred image to the original, unblurred image.

$$x_{\text{est}} = \begin{bmatrix} h[0] \\ h[1] & h[0] \\ h[2] & h[1] & h[0] \\ & h[2] & h[1] & h[0] \\ & & h[2] & h[1] & h[0] \\ & & & h[2] & h[1] \\ & & & & h[2] \end{bmatrix} \begin{bmatrix} s[0] \\ s[1] \\ s[2] \\ s[3] \\ s[4] \end{bmatrix}$$

forms the matrix with entries from h , and the second,

$$x_{\text{est}} = \begin{bmatrix} s[0] \\ s[1] & s[0] \\ s[2] & s[1] & s[0] \\ s[3] & s[2] & s[1] \\ s[4] & s[3] & s[2] \\ & s[4] & s[3] \\ & & s[4] \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ h[2] \end{bmatrix} = Sh,$$

uses the entries of s in the matrix S . Note that here $h * s$ computes the “full” convolution, which includes using the zero-padded boundaries. In MATLAB, we can compute the convolution with the `conv` function:

```
x_est=conv(h, s, 'full');
```

For the gradient $\nabla_h f$, we treat s as a constant, and so it is natural to use the second matrix-vector product where we use the entries of s to form the linear operator. We have

$$\begin{aligned} \nabla_h f &= \nabla_h \frac{1}{2} \|Sh - x\|^2 \\ &= S^*(Sh - x) = S^*r, \end{aligned}$$

where we define the residual $r = Sh - x = x_{\text{est}} - x$. Written out, this gradient is shown in the boxed equation at the bottom of the page. We recognize this as the cross correlation between s and r :

$$\begin{aligned} \nabla_h f[n] &= \sum_{k=0}^4 \bar{s}[k] r[k+n] \\ &= (\bar{s}[-] * r[.])[n], \end{aligned}$$

where $\bar{s}[-]$ is the matched filter of $s[.]$. Note that the convolution here does not include the entries that use the zero-padded boundaries. In MATLAB, we can compute this convolution with

```
Dfh = conv(r, conj(s(end:-1:1)), 'valid');
```

The derivation and computations for $\nabla_s f$ are very similar.

Now that we can efficiently compute the required gradients, we can use existing first-order methods. As an example, we consider a simulated underwater acoustic channel. A single unknown real source signal is broadcast over two noisy acoustic channels with unknown real impulse responses. We extend the blind channel estimation problem to two channels quite naturally:

$$\min_{h_i, s} \sum_{i=1}^2 (\|h_i * s - x_i\|_2^2 + \lambda_h \|h_i\|_1) + \lambda_s \|s\|_1.$$

Further, we may add a differentiable 1-D TV-like term to promote sharp transitions and stretches of nearly constant signal values (e.g., where the channel impulse response is constant for brief periods). The true TV seminorm of h may be computed as $\|\mathcal{D}h\|_1$, where \mathcal{D} is the operator that subtracts consecutive pairs of elements of h : $\{\mathcal{D}h\}_j = h_{j+1} - h_j$. We can, in fact, form \mathcal{D} as a sparse matrix and compute \mathcal{D}^* easily. To “soften” the TV norm to a differentiable term, we may use the Huber function, defined as

$$L_\delta(z) = \begin{cases} \frac{1}{2}z^2 & |z| \leq \delta \\ \delta(|z| - \frac{1}{2}\delta) & |z| > \delta. \end{cases}$$

The Huber function is commonly used in regression with outliers, as it is linear for $|z| > \delta$, and so it is less sensitive to outliers. Unlike $\|\cdot\|_1$, the Huber function is differentiable for all z . We can approximate the TV norm with

$\|h\|_{\text{TV}} := \|\mathcal{D}h\|_1 \approx L_\delta(\mathcal{D}h)/\delta$. The augmented problem, with the differentiable TV-like term is

$$\begin{aligned} f_\delta(h_i, s) &= \|h_i * s - x_i\|_2^2 \\ &\quad + \lambda_{h, \text{TV}} L_\delta(\mathcal{D}h_i)/\delta, \\ \min_{h_i, s} \sum_{i=1}^2 (f_\delta(h_i, s) + \lambda_h \|h_i\|_1) &+ \lambda_s \|s\|_1. \end{aligned} \tag{6}$$

It is straightforward to include the “soft” TV term in the gradients, and we can use standard first-order algorithms to solve the augmented problem. Simulated data containing an impulsive source, two underwater acoustic channel impulse responses, and the two outputs of the underwater channels were provided by [11]. The simulation also added zero mean Gaussian noise with standard deviation 0.005. We used the two channel outputs to attempt to recover the impulsive source and channel impulse responses.

We take $\lambda_h = 0.1$, $\lambda_s = 0.01$, $\lambda_{h, \text{TV}} = 0.01$, and $\delta = 0.1$ for the Huber functions $L_\delta(\cdot)$. We initialized h_i and s with zero mean Gaussian random numbers with unit standard deviation. In practice, we may not know the true lengths of h_i and s , K and N , respectively. We do know the length of the observed signals, which is $K + N - 1$. In this example, $K = 894$ and $N = 1,717$, and we guess $K_{\text{est}} = 844$ and $N_{\text{est}} = 1,767$. We used L1General [12] to find a local solution of the augmented problem (6). The initial guesses for h_i and s were drawn from the standard normal distribution.

The results of the recovery are shown in Figure 4. We show the output and channel impulse response for only the first channel; the second channel is similar. Researchers studying the acoustic

$$\nabla_h f = \begin{bmatrix} \bar{s}[0] & \bar{s}[1] & \bar{s}[2] & \bar{s}[3] & \bar{s}[4] \\ & \bar{s}[0] & \bar{s}[1] & \bar{s}[2] & \bar{s}[3] & \bar{s}[4] \\ & & \bar{s}[0] & \bar{s}[1] & \bar{s}[2] & \bar{s}[3] & \bar{s}[4] \end{bmatrix} \times \begin{bmatrix} r[0] \\ r[1] \\ r[2] \\ r[3] \\ r[4] \\ r[5] \\ r[6] \end{bmatrix}$$

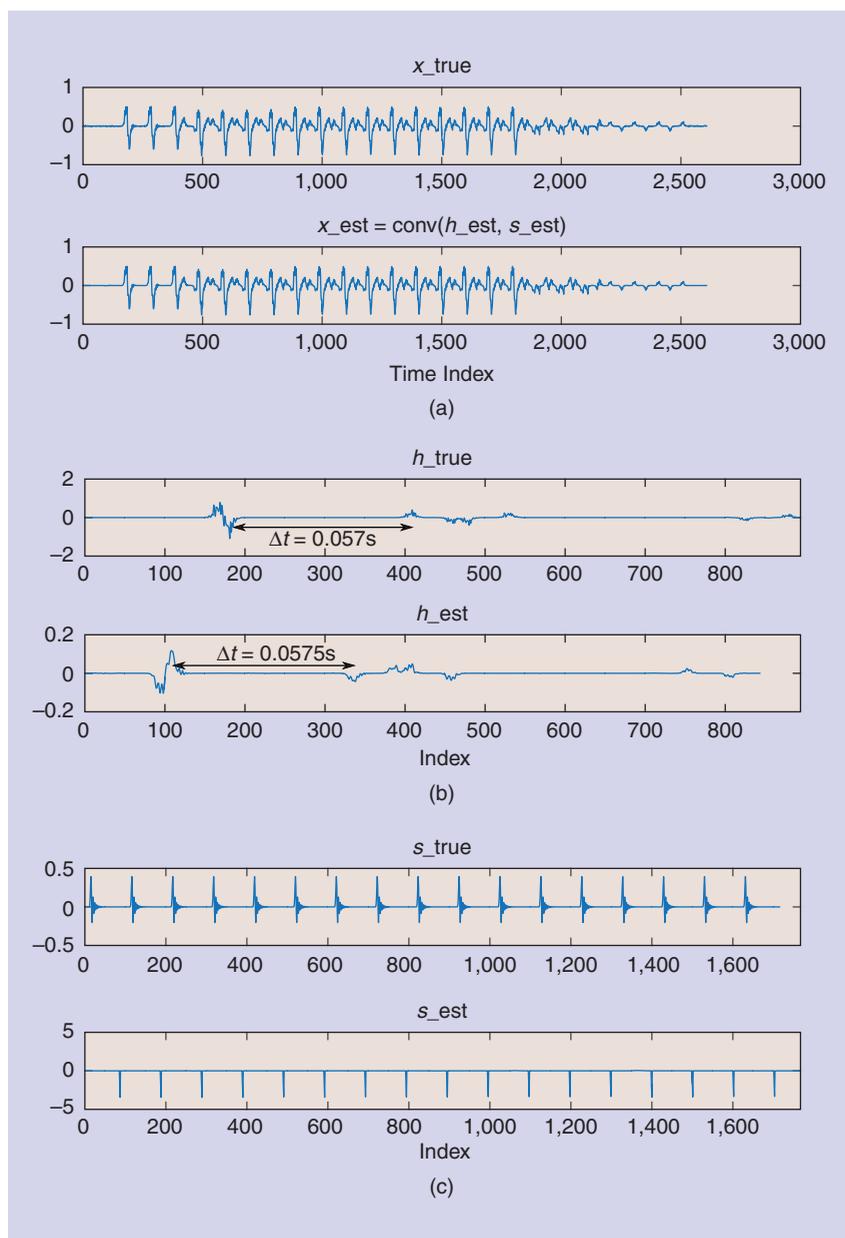


FIGURE 4. The blind channel estimation via problem (6) for a simulated underwater acoustic impulsive source. Shown are the (a) true and estimated first channel output, (b) first channel impulse response, and (c) source. There are obvious magnitude and time shift differences, but the estimated channel output and channel impulse response are remarkably similar in structure and potentially useful for studying the underwater channel.

channel are mainly interested in the time delays between peaks and relative phase shifts [11]. Besides the overall magnitude ambiguity and time shift, it appears we can accurately estimate the time delays and phase shifts between the peaks of the channel response.

Discussion

We have demonstrated a framework for computing the adjoint of discrete wave-

let transforms, which arise in gradient computations in optimization problems. The framework separates the action of the wavelet transform and the signal extension operator. The half-point symmetric extension is a common choice, and we have provided its transpose, but other extension operators can readily be cast into the framework. For many practical purposes, it appears that the adjoint operator is remarkably

close to the pseudoinverse transform, however, implementing the true adjoint is just a simple matter of handling the signal extension operation.

Another interesting adjoint appears in a blind channel estimation optimization problem. The convolution of two variables, which is a nonlinear operation, can be written as a matrix-vector product in two ways. With this viewpoint, deriving the gradient becomes simple and a fast implementation is immediate. For an impulsive source sent through a simulated underwater acoustic channel, the blind channel estimation problem is able to accurately recover important details of the acoustic channel response.

Supplementary material

This article has supplementary downloadable material available at https://github.com/jamesfolberth/ieee_adjoints, provided by the authors. The material includes MATLAB implementations of the adjoint operators discussed in this lecture note. Contact james.folberth@colorado.edu for further questions about this work.

Authors

James Folberth (james.folberth@colorado.edu) is a Ph.D. student in Stephen Becker's Information Extraction group in the Department of Applied Mathematics at the University of Colorado at Boulder. His research interests include signal and image processing, machine learning, and optimization.

Stephen Becker (stephen.becker@colorado.edu) received his bachelor's degree in mathematics and physics from Wesleyan University in 2005 and his Ph.D. degree in 2011 in applied and computational mathematics from the California Institute of Technology. He is an assistant professor in the Department of Applied Mathematics at the University of Colorado at Boulder. He was the recipient of a postdoctoral fellowship to study at Paris 6 and the Goldstine postdoctoral fellowship at IBM Research. His work on the TFOCS algorithm won the triennial Beale-Orchard-Hays Prize from the

(continued on page 147)

Zhuo Wang, Jintao Zhang,
and Naveen Verma

Reducing Quantization Errors for Inner-Product Operations in Embedded Digital Signal Processing Systems

Inner-product operations are used extensively in embedded digital signal processing (DSP) systems. Their applications range from signal processing (filtering/convolution) to inference (classification). In embedded systems, resources (energy, area, etc.) are typically highly constrained, making tradeoffs with computational precision a fundamental concern. Indeed, with increasing requirements on algorithmic performance, many systems are trending toward higher computational precision to ensure accuracy of results.

This article describes an approach to significantly enhance accuracy for inner-product operations at very low bit precisions, significantly improving the energy/area-versus-accuracy tradeoffs traditionally incurred [1]. Low-energy embedded systems often employ linear, fixed-precision representation for operands due to the simplicity and relative efficiency at lower dynamic range. We focus on the use of a floating-point representation. For specific distributions of operands very commonly observed, such a representation substantially enhances the accuracy of most operands. However, particularly at the low bit precisions we focus on, it raises the issue that many operands can incur a large quantization error, much greater than that with standard linear, fixed-precision representation. To address this, a simple optimization is applied

within the quantization process, which is shown to significantly improve accuracy at a given bit precision. By targeting the way in which operands are represented, the approach incurs no added hardware or computational overheads and in fact is shown to reduce energy and area thanks to simplified computation associated with the representation.

A distribution of operands commonly encountered

An inner-product operation can be represented as follows:

$$y = \sum_i h_i x_i. \quad (1)$$

Here, x_i is input data, while h_i is typically predefined coefficients. For example, in the case of a finite impulse response (FIR) filter, x_i are signal samples to be filtered, while h_i are the filter coefficients; in the case of a perceptron (i.e., linear classifier), x_i are input feature elements, while h_i are the weights and bias. Quite often, the distribution of h_i values can result in severe quantization errors when implementing the

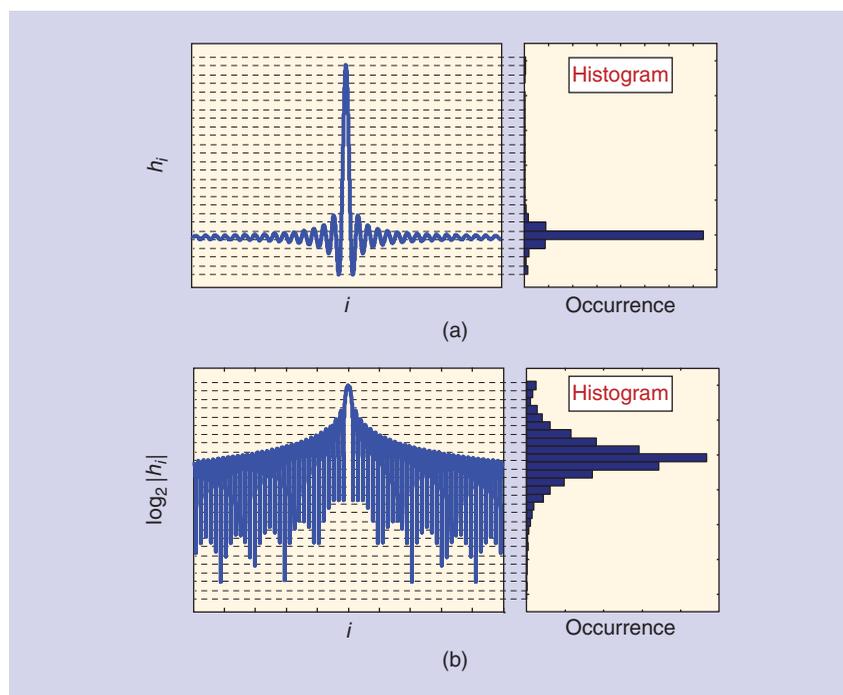


FIGURE 1. The distribution of filter coefficients for a low-pass FIR filter, with (a) standard linear fixed-point quantization and (b) floating-point representation.

Digital Object Identifier 10.1109/MSP.2016.2585278
Date of publication: 4 November 2016

inner-product operation in finite precision, as in an embedded DSP system. As an example, Figure 1 shows the distribution of h_i values for a low-pass FIR filter (plotted is $h_i = \sin c(n_i\omega)$ for $n_i\omega \in [-10\pi, 10\pi]$). The situation is similar for band-pass filters, for instance, achieved by modulating the low-pass impulse response with a sinusoid. With standard linear, fixed-point quantization, as shown in Figure 1(a), the majority of the coefficients fall in low-value bins. This implies inefficient use of the available dynamic range. Even worse, depending on the level of quantization, many of the h_i values fall in the zero-valued bin, resulting in severe output errors, as the number of such elements can generally be quite high. Such a result is commonly observed, e.g., high-order FIR filters (low pass, band pass), linear regression under near colinearity conditions, etc. [2]. To address this, we consider employing a floating-point representation. First, we examine a standard floating-point representation in terms of how it impacts the quantization error. Then, we examine the new opportunity this raises for optimizing the quantization error. Note that we assume that x_i remains in a fixed-point representation, as this may be determined by the preceding signal processing stage and therefore required for system simplicity (i.e., to avoid the overhead of explicitly altering the representation). Thus, the approach actually employs a mixed fixed-/floating-point representation.

Mixed fixed-/floating-point computation

The potential benefit of using floating-point representation for h_i is that increased resolution is allocated where the density of h_i values is greatest, i.e., low-valued bins. Further, no h_i values fall in a zero-valued bin. As illustrated in Figure 1(b), by spanning more quantization bins, such a representation leads to more efficient use of the available dynamic range for the example of the low-pass FIR filter.

To be more precise, the floating-point representation used for h_i is shown in (2)–(5), where l_i represents the sign

of h_i , s_i represents the exponent, and $1 + m_i$ represents the significand in the range of $[1, 2)$. Among these parameters, only m_i can take on continuous values. Thus, values of m_i in the range of $[0, 1)$ must be quantized (\hat{m}_i) for system implementation.

$$h_i = l_i \times 2^{s_i} \times (1 + m_i) \quad (2)$$

$$l_i = \text{sign}(h_i) \quad (3)$$

$$s_i = \lceil \log_2 |h_i| \rceil \quad (4)$$

$$m_i = \frac{|h_i|}{2^{\lceil \log_2 |h_i| \rceil}} - 1. \quad (5)$$

There is a practical consideration in choosing the number of bits assigned to m_i and s_i . For example, (6) and (7) show two approaches to represent the number -100 with 6 bits when using floating-point representation. The first approach assigns more bits to s_i , thus can represent a larger range of h_i (i.e., $(\max |h_i|) / (\min |h_i|) \approx 2^{16}$ versus 2^8). On the other hand, the second approach achieves a higher resolution within the range by assigning more bits to m_i (i.e., $\Delta m_i = 0.5$ versus 0.25). Thus, how to best allocate the bits is determined by the distribution of h_i ; as a rule of thumb, we start by allocating just enough bits for s_i to cover the full range of h_i (i.e., $(\max |h_i|) / (\min |h_i|)$), and then allocate bits to m_i to achieve the highest bin resolution

$$l_i = 1'b0; s_i = 4'b0110; \hat{m}_i = 1'b1 \quad (6)$$

$$l_i = 1'b0; s_i = 3'b110; \hat{m}_i = 2'b10. \quad (7)$$

Now, let's explore the benefits afforded by such a mixed fixed-/floating-point representation. The first benefit, first seen in Figure 1, is shown more explicitly through quantization-error histograms in Figure 2. Here, we consider two types of FIR filters that are used within a seizure-detection system [3], which is employed later in the experimental demonstration for detailed analysis of the approach. The system extracts spectral-energy features from electroencephalogram (EEG) data sampled at 256 Hz. Since seizure-detection features correspond to low-frequency ranges, the system employs front-end decimation, implemented using a

256-order low-pass FIR filter. Following this, eight 64-order band-pass FIR filters (centered from 0 to 21 Hz) are used to isolate the signal in 3-Hz bands, for subsequent energy accumulation over a 2-second epoch. Then, the spectral-energy features from three consecutive epochs are combined to form the overall feature vector. Note that only one band-pass feature-extraction filter is considered for illustration, though all have similar coefficient distributions

As seen in Figure 2, floating-point representation has a quantization error with greatest density at values well below 1 least significant bit (LSB), while fixed-point representation has a quantization error more evenly distributed up to 1 LSB. However, we also see that, with fixed-point representation, the quantization error is limited to 1 LSB; on the other hand, with floating-point representation, the quantization error can exceed 1 LSB. Indeed, for the examples shown, there are many cases where the errors far exceed 1 LSB. This can lead to a significant error. To address this, the error can be further minimized thanks to a simple optimization enabled by the floating-point representation. This is described next, using the concrete example of the seizure-detection system to illustrate the approach and its rationale.

However, before proceeding, we highlight an additional benefit of floating-point representation, which is that it leads to an implementation for multiplier hardware that consumes less energy and area. The reason is that input signals now only need to multiply with \hat{m}_i , which is represented by fewer bits compared to h_i . The remaining operations are simply barrel shifting and sign application, which can be trivially implemented. Given the prominence of multiplication operations in DSP systems, this can present a significant advantage.

Optimizing the quantization error

This section discusses the optimization enabled by floating-point representation to further reduce the quantization error. The opportunity arises from the fact that, in DSP systems, outputs can often be scaled arbitrarily. For instance, in the case of fixed-point representation, this is

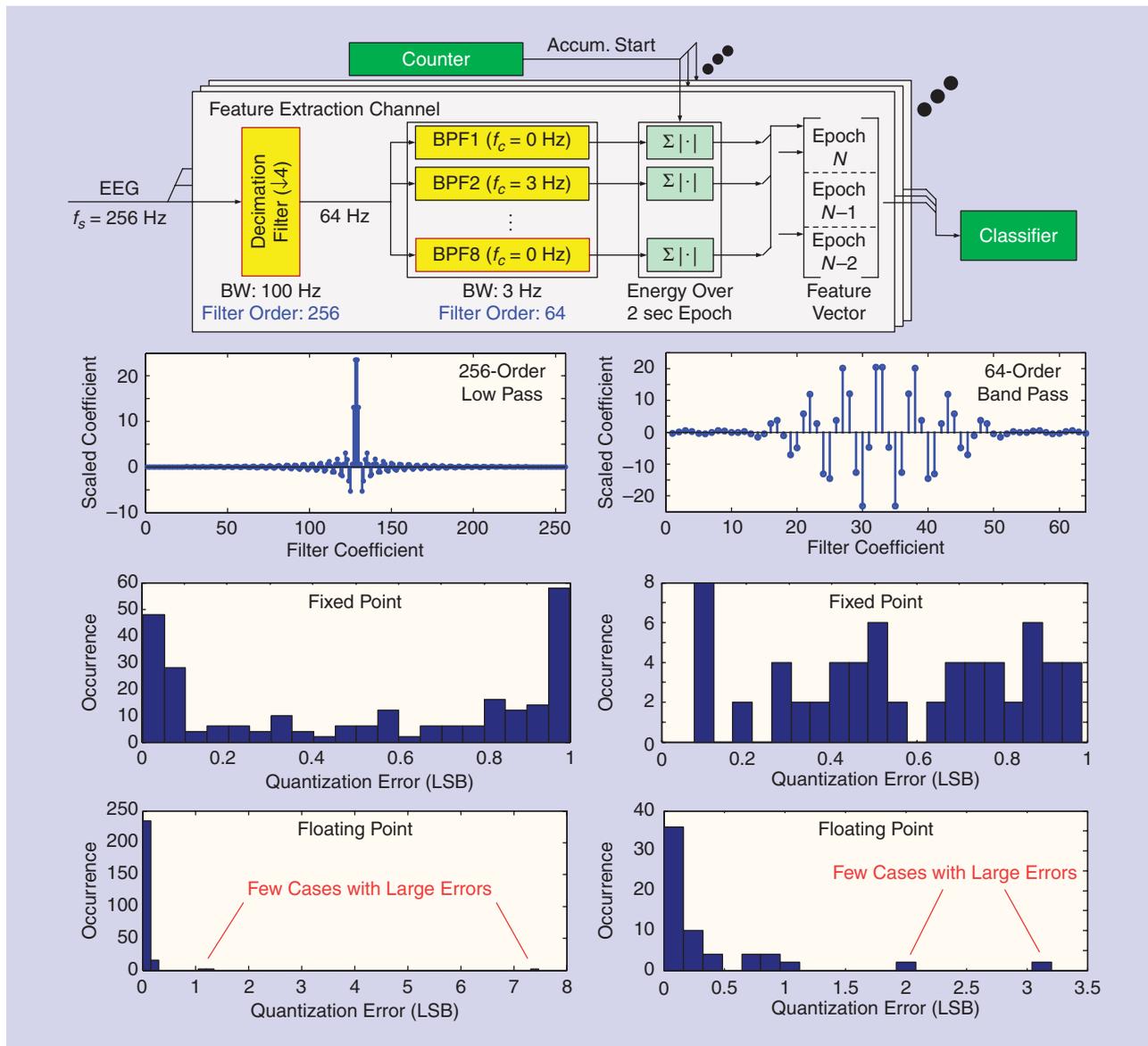


FIGURE 2. A comparison of quantization error for fixed-point and floating-point quantization of FIR filter coefficients. FIR filters are from an EEG-based seizure-detection application consisting of low-pass decimation and band-pass feature extraction.

often done to best utilize the available dynamic range of the system. On the other hand, in the case of a floating-point representation, which affords representation of very large-valued numbers, let's consider how multiplication of all h_i values by a single positive factor α improves utilization of dynamic range, but in a different way. The approach is shown in Figure 3. With floating-point representation of the quantization levels, multiplication by α can be used to move the h_i s closer to the quantization levels. This is made possible because, with the large range that can be covered

using floating-point representation, we can ensure that all values remain within the representation bounds even after the multiplication. With this ensured, the question now is simply how to find an optimal α that leads to minimum output error $\varepsilon_{\hat{y}}$ for the inner-product operation.

With the scaling parameter α applied to each h_i , the resulting outputs y can now be expressed as in (8). Quantization of αh_i ($\widehat{\alpha h_i}$) then affects the output value as shown in (9). Thus, the output error caused by quantization of h_i can be expressed as in (10). Notice here h_i is assumed to be predefined values, while

x_i is assumed to be drawn from a statistical distribution corresponding to the input X_i . As a result, output error $\varepsilon_{\hat{y}}$ is also a statistical distribution, which is parameterized by α .

$$y(\alpha) = \sum_i (\alpha h_i) \cdot x_i \tag{8}$$

$$\hat{y}(\alpha) = \sum_i (\widehat{\alpha h_i}) \cdot x_i \tag{9}$$

$$\varepsilon_{\hat{y}} = \frac{\left| \sum_i (\alpha h_i) x_i - \sum_i (\widehat{\alpha h_i}) x_i \right|}{\alpha} \tag{10}$$

To find the optimal α that minimize output error $\varepsilon_{\hat{y}}$, we proceed by first finding

the distribution of output error $E_{\hat{Y}}$. This requires us to know the distribution of inputs X_i . Note here that capital letters are used for representing the distribution of signal samples, rather than a signal. For simplifying our discussion without loss of generality, we present the optimization approach through an example. We focus on FIR filtering required within the EEG-based seizure-detection system [3], which allows us to specify X_i for the EEG data. For other applications, the same approach may be applied but with a corresponding assumption on the statistical distribution X_i .

First, for FIR filtering, we assume the distribution of X_i over i is identical as all samples are drawn from the same signal source. We also make the assumption here that all X_i is independently distributed. Notice that independence may not hold in many cases. Nonetheless, the assumption is made for convenience of the derivation, and its validity is later verified through simulation of a practical application. The distribution of EEG samples (derived from actual patient signals) is shown in Figure 4(a). We model samples of the EEG data by a normal distribution $X_i \sim N(0, \sigma^2)$. The variable of interest, $E_{\hat{Y}}$, can be derived as shown in (11). Note that we deliberately leave out the scaling parameter α here for simplicity of expression. With the previously stated assumptions on the input X_i , the distribution of $E_{\hat{Y}}$ can be derived, as shown in (12). As seen, the output error is also normally

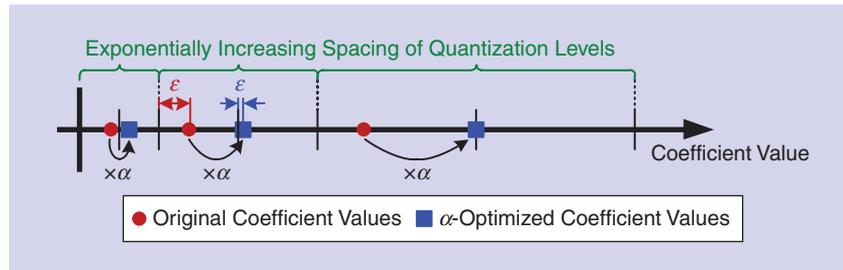


FIGURE 3. Introducing the scaling parameter α to enable coefficient optimization.

distributed, with zero mean and variance dependent on the squared sum of the error on all h_i . Notice, it can be concluded from the previous derivations that the output \hat{Y} itself also follows a zero-mean normal distribution. This implies that the same analysis can be applied in a scenario where inner-product operations are cascaded

$$E_{\hat{Y}} = Y - \hat{Y} = \sum_i h_i \cdot X_i - \sum_i \hat{h}_i \cdot X_i$$

$$= \sum_i (h_i - \hat{h}_i) \cdot X_i = \sum_i \epsilon_{\hat{h}_i} \cdot X_i \tag{11}$$

$$E_{\hat{Y}} \sim N(0, \sum_i \epsilon_{\hat{h}_i}^2 \sigma^2). \tag{12}$$

Using (12), a reasonable objective function $C_{\hat{h}}(\alpha)$ would be one as shown in (13) that minimizes the variance of the scaling factor α . Within this function, $\epsilon_{\hat{h}_i}$ can be derived as shown in (14), where l_i , s_i , and m_i can be directly specified from h_i as previously shown in (3)–(5), and \hat{m}_i can be obtained from quantizing m_i , as

in (15), with k being the quantization level. Consequently, for a given set of h_i , $\epsilon_{\hat{h}_i}$ is only a function of α . Note, however, this function is not convex. It is not even continuous due to the quantization operation applied to \hat{m}_i . Fortunately, there is a structure to this function, making the optimization problem trivial to solve.

$$\min_{\alpha} C_{\hat{h}}(\alpha), \alpha \in R^+, \text{ where } C_{\hat{h}}(\alpha) = \sum_i \epsilon_{\hat{h}_i}^2(\alpha) \tag{13}$$

$$\epsilon_{\hat{h}_i} = h_i - \hat{h}_i = l_i \cdot 2^{s_i} \cdot (1 + m_i) - l_i \cdot 2^{s_i} \cdot (1 + \hat{m}_i)$$

$$= l_i \cdot 2^{s_i} \cdot (m_i - \hat{m}_i) \tag{14}$$

$$\hat{m}_i = \lfloor m_i \cdot 2^k \rfloor / 2^k. \tag{15}$$

To see the structure, (16) modifies (14) by applying the α scaling factor. After plugging in (3)–(5) and (15), it can be seen that α only appears in the cost function in the form of $f(\alpha)$, shown in (17). From this, it can be easily seen that $f(\alpha) = f(2\alpha)$ for $\forall \alpha \in R^+$. This indicates that the cost function in

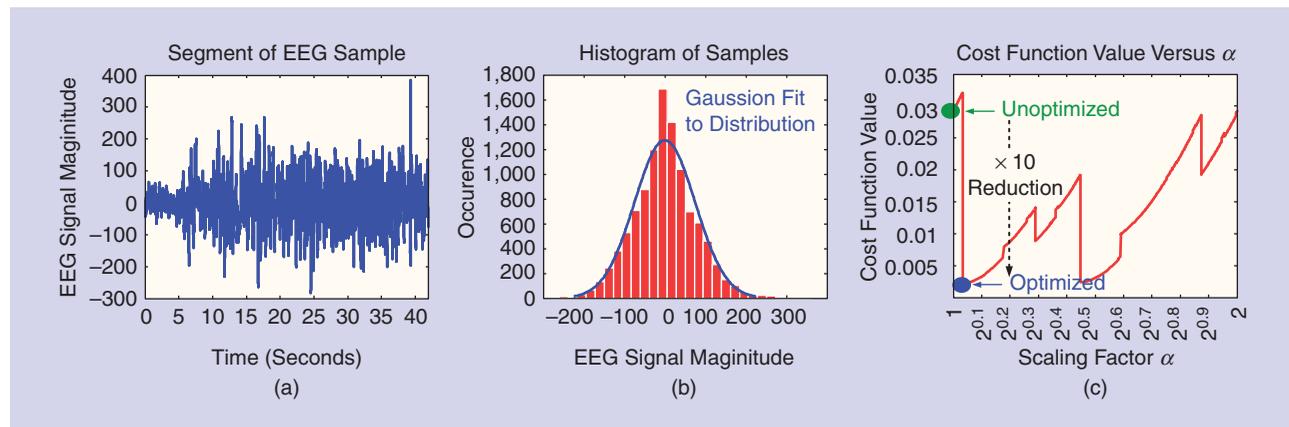


FIGURE 4. The distribution of EEG signals can be modeled as drawn from a Gaussian distribution, as shown by (a) representative time samples from a segment of an EEG channel and (b) the histogram of sample values for the segment, leading to (c) the cost function shown for the α -optimization.

(13) is a repeating function of α . As a result, a global minimum of $C_{\hat{h}_i}(\alpha)$ exists in the range $[1, 2)$. Thus, the optimization is easily solved numerically by sweeping α within this range. For the filter considered, Figure 4 shows the resulting cost function and how its value is improved by the optimal choice of α

$$\begin{aligned} \varepsilon_{\hat{h}_i}(\alpha) &= \frac{\alpha h_i - \widehat{\alpha h}_i}{\alpha} \\ &= l_i \cdot \frac{2^{s_i(\alpha)}}{\alpha} \cdot [m_i(\alpha) - \hat{m}_i(\alpha)] \end{aligned} \quad (16)$$

$$f(\alpha) = 2^{\lceil \log_2(\alpha |h_i|) \rceil} / \alpha. \quad (17)$$

Summary of approach

Below, the specific procedure for applying the optimization generally to minimize quantization error is provided.

- 1) Select a precision requirement for h_i (based on preferred tradeoff of computational accuracy and hardware complexity), e.g., 6 bits in our application.
- 2) Perform floating-point quantization of h_i [assigning bits to significand and exponent based on the corresponding dynamic range $(\max|h_i|)/(\min|h_i|)$].
- 3) Determine the scaling parameter α .
 - a) Model the statistical distribution of inputs X_i .

- b) Compute the distribution of output error $\varepsilon_{\hat{y}}$ due to quantization of h_i [see (10)].
 - c) Define a cost function based on the distribution of output error $\varepsilon_{\hat{y}}$.
 - d) Minimize the cost function to find optimal α .
- 4) Redetermine the quantization for h_i with the scaling parameter α .
 - 5) Only if the absolute value of output y is critical, apply multiplication by $1/\alpha$ to the outputs.

Note that, when required, the energy of Step 5 is greatly amortized since it involves one multiplication, compared to many in the preceding convolution operation.

Experimental demonstration

For demonstration, the presented approach is applied to the EEG-based seizure-detection system of Figure 2. For analysis and demonstration of the approach, three hardware implementations of filters are compared, each employing 6–10 bits for representing h_i . The first implementation is a conventional fixed-point FIR filter. The second implementation is the mixed fixed-/floating-point FIR filter without any α optimization. For this, s_i is assigned 4 b for the decimation filter and 3 b for the band-pass filter to address the required range

of h_i $((\max|h_i|)/(\min|h_i|)) \approx 37k$ for the decimation filter, and 500 for the band-pass filter). m_i is assigned correspondingly for 6–10 bit overall precision of h_i . The third implementation is the presented mixed fixed-/floating-point FIR filter with α optimization. Simulation results are presented using 210 seconds of EEG data (sampled at 256 Hz). All implementations are developed in both MATLAB and Verilog RTL. For analysis of quantization error, the MATLAB implementations are used. For hardware energy and area analysis, the Verilog implementations are synthesized to standard cells and laid out in a 32-nm complementary metal–oxide–semiconductor technology, with transistor-level postlayout simulations performed using NanoSim.

The results in Figure 5 focus on quantization errors. The histograms show the quantization error distributions for the decimation filter and band-pass filter after applying the optimization. Comparing Figure 5 to Figure 2, we can see that with α optimization applied to the mixed fixed-/floating-point approach, the quantization errors are more densely clustered close to zero. The output error resulting from the three implementations for 6-bit coefficient representation are also compared. The samples are ordered based on the magnitude of

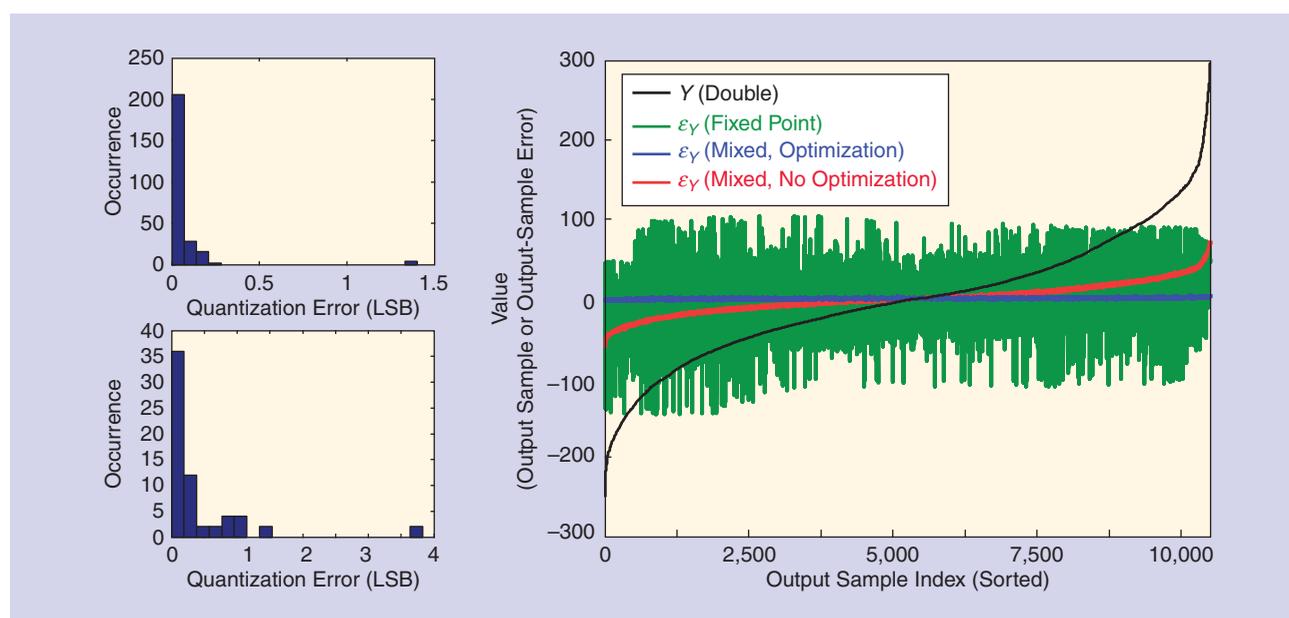


FIGURE 5. The reduction of quantization error by applying the presented optimization.

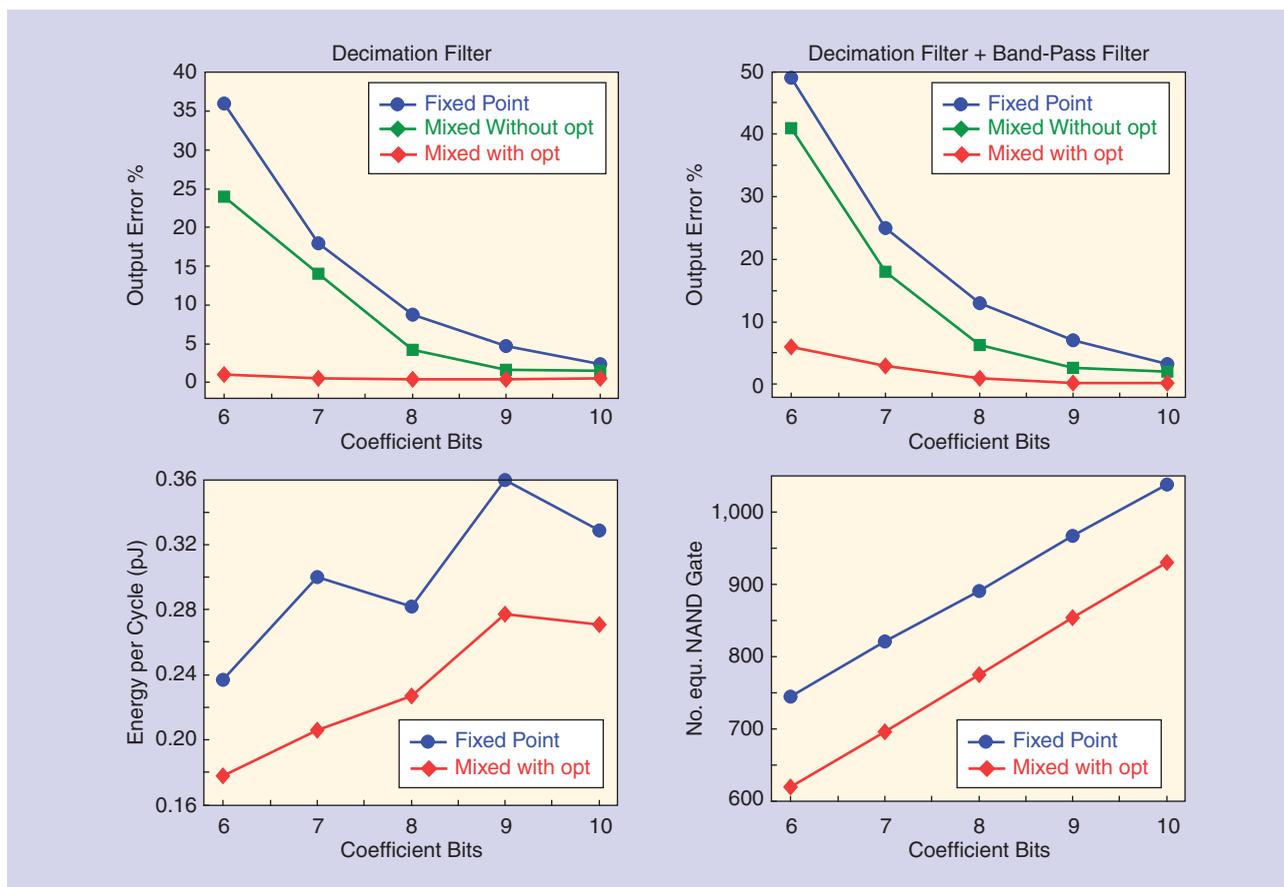


FIGURE 6. A comparison of approaches in terms of quantization error, energy, and hardware complexity.

the output values to aid visualization. As seen, fixed-point coefficient representation results in substantial error. While floating-point coefficient representation reduces this, the error is substantially reduced even further thanks to α optimization.

Figure 6 provides an overall summary, showing the output error (due to quantization), computational energy, and hardware area, versus bit level for representing h_i , for the three implementations. The output error corresponds to the root-mean-square error of the output computed with 6-bit quantized filter coefficients h_i versus double-precision h_i , normalized to the root-mean-square value of the output computed with double-precision h_i . The computational energy is the average energy per inner-product operation determined

The energy-precision tradeoff associated with quantization in a DSP system has always been a critical concern during design.

from postlayout transistor-level (Nano-Sim) simulation of the accelerator, and the hardware area is represented as the equivalent number of NAND gates for the implementation (note that the energy and area of the mixed fixed-/floating-point representation implementation is the same with and without α optimization). The presented implementation with optimization leads to

a substantially lower error than an implementation based on fixed-point coefficient representation, especially at low bit levels. For example, for the decimation filter alone, the presented approach leads to a 37 \times error reduction compared to a conventional fixed-point implementation at the 6-bit level. When considering cascading of the decimation filter and band-pass filter, the corresponding error reduction is 28 \times .

Comparing the energy consumption and hardware complexity, for the fixed-point and mixed fixed-/floating-point implementations, the presented approaches also lead to hardware-resource savings in addition to reducing errors (note, the computational energy and hardware area with and without α optimization are equivalent). As an example, at the 6-bit level, the presented approach leads to 1.4 \times energy savings and 1.2 \times area savings compared to conventional fixed-point implementation. Indeed, that approach is shown to yield such advantages in a broad range of system implementations, involving convolution operations for various tasks (e.g., [4] presents its use for optimizing the implementation of a linear classifier).

Summary

The energy-precision tradeoff associated with quantization in a DSP system has always been a critical concern during design. Considering the

importance of inner-product operations, this article focuses on an approach by which this tradeoff can be substantially improved. The first aspect of the approach is the use of a mixed fixed-/floating-point representation, by which the dynamic range is more efficiently used. The second aspect is a simple optimization enabled by the representation whereby quantized coefficients can all be scaled to minimize the overall error incurred. Filters within an EEG-based seizure-detection system are used to demonstrate the approach. Substantial reduction in the error, computational energy, and hardware area is observed as a result of the approach.

Authors

Zhuo Wang (zhuow@princeton.edu) received his B.S. degree in microelectronics and his M.S. degree from Peking University, Beijing, China, in 2011 and 2013, respectively. Currently, he is working toward his Ph.D. degree in the Department of Electrical Engineering, Princeton University, New Jersey. His research focuses on leveraging statistical

approaches such as machine learning for achieving hardware relaxation in an algorithmic and architectural level in resource-constrained platforms such as embedded sensing systems.

Jintao Zhang (jintao@princeton.edu) received his B.S. degree in electrical engineering from Purdue University, West Lafayette, Indiana, in 2012. Currently, he is working toward his Ph.D. degree at Princeton University, New Jersey. His research is focused on energy-efficient signal analysis (mainly classification and on chip). His primary research interests are very-large-scale integration system design, exploring mixed signal circuit design by using digital controls to enable various analog functionalities of the complementary metal-oxide-semiconductor and, thus, enhancing the system performance with both the benefit from analog circuits and digital circuits.

Naveen Verma (nverma@princeton.edu) received his B.A.S. degree in electrical and computer engineering from the University of British Columbia, Vancouver, Canada, in 2003, and his M.S. and Ph.D. degrees in electrical engineer-

ing from the Massachusetts Institute of Technology, Cambridge, in 2005 and 2009, respectively. Currently, he is an associate professor of electrical engineering at Princeton University, New Jersey, where he has been since 2009. His research focuses on advanced sensing systems, including low-voltage digital logic and static random access memories, low-noise analog instrumentation and data conversion, large-area sensing systems based on flexible electronics, and low-energy algorithms for embedded inference, especially for medical applications.

References

- [1] Z. Wang, J. Zhang, and N. Verma, "Reducing quantization error in low-energy FIR filter accelerators," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Brisbane, Australia, 2015.
- [2] D. Montgomery, E. Peck, and G. Vining, *Introduction to Linear Regression Analysis*. Hoboken, NJ: Wiley, 2015.
- [3] A. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, 2009.
- [4] Z. Wang, J. Zhang, and N. Verma, "Realizing low-energy classification systems by implementing matrix multiplication directly within an ADC," *IEEE Trans. Biomed. Circuit. Syst.*, vol. 9, no. 6, pp. 825–837, 2015.

SP

LECTURE NOTES (continued from page 140)

Mathematical Optimization Society. His current research is centered around deriving efficient optimization methods for large-scale data analysis, with an emphasis on techniques for sparse and low-rank models.

References

- [1] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Spectra, and Filtering*. Philadelphia: SIAM, 2006.
- [2] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse prob-

lems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, Mar. 2009.

[3] *MATLAB and Wavelet Toolbox Release R2015b*. Natick, MA: The MathWorks, Inc., 2015.

[4] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. London, U.K.: Elsevier, 2009.

[5] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, 1996.

[6] A. Weber, "The USC-SIPI image database: Version 5," USC-SIPI Tech. Report 315, Univ. of Southern California, Los Angeles, 1993.

[7] S. Rout, "Orthogonal vs. biorthogonal wavelets for image compression," Ph.D. dissertation, Dept. Electrical and Computer Engineering Virginia Tech, Blacksburg, VA, 2003.

[8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[9] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Trans. Inform. Theory*, vol. 60, no. 3, pp. 1711–1732, 2014.

[10] J. F. Claerbout, *Earth Soundings Analysis: Processing Versus Inversion*, vol. 6, Cambridge, MA: Blackwell, 1992.

[11] B. Rideout and E. M. Nosal. Personal communication.

[12] M. Schmidt, "Graphical model structure learning with ℓ_1 -regularization," Ph.D. dissertation, Dept. Computer Science, Univ. British Columbia, Vancouver, 2010.

SP

DATES AHEAD

Please send calendar submissions to:

Jessica Barragué

E-mail: j.barrague@ieee.org

2016

DECEMBER

Picture Coding Symposium (PCS)

4–7 December, Nuremberg, Germany.

General Chair: André Kaup

URL: <http://www.pcs2016.com/>**Eighth IEEE International Workshop on Information Forensics and Security (WIFS)**

5–7 December, Abu Dhabi, UAE.

General Chairs: Ernesto Damiani and Nasir Memon

URL: <http://wifs2016.mdabaie.com/>**IEEE Global Conference on Signal and Information Processing (GlobalSIP)**

7–9 December, Greater Washington, D.C., USA.

General Chairs: Zhi Tian and Brian Sadler

URL: <http://2016.ieeeglobalsip.org>**IEEE World Forum on Internet of Things (WF-IoT)**

12–14 December, Reston, Virginia, USA.

General Chairs: Geoff Mulligan and Latif Ladid

URL: <http://wfiot2016.ieee-wf-iot.org/>**IEEE Spoken Language Technology Workshop (SLT)**

13–16 December, San Juan, Puerto Rico.

General Chairs: Najim Dehak and Pedro Torres-Carrasquillo

URL: <http://www.slt2016.org/>**Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)**

13–16 December, Jeju, South Korea.

General Chairs: Yo-Sung Ho, C.-C. Jay Kuo, and Haizhou Li

URL: <http://www.apsipa2016.org/>Digital Object Identifier 10.1109/MSRP.2016.2604158
Date of publication: 4 November 2016

IMAGE LICENSED BY INGRAM PUBLISHING

The 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP) will be held 7–9 December 2016 in the greater Washington, D.C., area.

2017

MARCH

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

5–9 March, New Orleans, Louisiana, USA.

General Chair: Magdy Bayoumi

URL: <http://www.ieee-icassp2017.org/>

APRIL

IEEE International Symposium on Biomedical Imaging (ISBI)

18–21 April, Melbourne, Australia.

General Chairs: Olivier Salvado and Gary Egan

URL: <http://biomedicalimaging.org/2017/>**16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)**

18–21 April 2017, Pittsburgh, Pennsylvania, USA.

General Chair: Pei Zhang

URL: <http://ipsn.acm.org/2017/>

JULY

18th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)

3–6 July 2017, Hokkaido, Japan.

General Chairs: Yasutaka Ogawa, Wei Yu, and Fumiyuki Adachi

URL: <http://www.spawc2017.org/>**IEEE International Conference on Multimedia and Expo (ICME)**

10–14 July 2017, Hong Kong, China.

General Chairs: Jörn Ostermann and Kenneth K.M. Lam

URL: <http://www.icme2017.org/>

SEPTEMBER

IEEE International Conference on Image Processing (ICIP)

17–20 September, Beijing, China.

General Chairs: Xinggang Lin, Anthony Vetro, and Min Wu

URL: <http://2017.ieeeicip.org/>

OCTOBER

IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)

15–18 October, New Paltz, New York

General Chairs: Patrick A. Naylor and Meinard Müller

URL: <http://www.waspaa.com/>**19th IEEE International Workshop on Multimedia Signal Processing (MMSp)**

16–18 October, London-Luton, United Kingdom

General Chairs: Vladan Velisavljevic, Vladimir Stankovic, and Zixiang Xiong

URL: <http://mmsp2017.eee.strath.ac.uk/>

DECEMBER

17th IEEE International Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP)

10–13 December, Curacao, Dutch Antilles.

General Chairs: André L.F. de Almeida and Martin Haardt

URL: <http://www.cs.huji.ac.il/conferences/CAMSAP17/>

SP

ICIP 2017

IEEE International Conference on Image Processing
17-20 September 2017, Beijing, China
<http://2017.ieeeicip.org>

Honorary Chairs

Sanjit K. Mitra

Univ. of California at Santa Barbara

Baozong Yuan

Beijing Jiaotong University

General Chair

Xinggang Lin

Tsinghua University

General Co-Chairs

Anthony Vetro

Mitsubishi Electric Research Labs

Min Wu

Univ. of Maryland at College Park

Technical Program Chairs

Jiebo Luo

University of Rochester

Wenjun Zeng

Microsoft Research Asia, China

Yu-Jin Zhang

Tsinghua University

Finance Chair

Yun He

Tsinghua University

Tutorial Chairs

John Apostolopoulos

Cisco

Jiwu Huang

Shenzhen University, China

Special Session Chairs

Pascal Frossard

EPFL

Feng Wu

Univ. of Science and Tech. of China

Awards Chair

Phillip Chou

8, USA

Plenary Chairs

Yao Wang

New York University

Jiangtao Wen

Tsinghua University

Exhibits/Demo Chairs

Qionghai Dai

Tsinghua University

Qi Tian

University of Texas at San Antonio

Industry Chairs

Jian Lu

CTTalk Cloud, China

Debargha Mukherjee

Google

Doctoral Symposium Chair

Dong Xu

University Sydney, Australia

Social Media Chair

Andres Kwansinski

Rochester Institute of Technology

Publicity Chair

Chaker Larabi

Université de Poitiers

Registration Chair

Tiejun Huang

Peking University

Local Arrangement Chairs

Xilin Chen

ICT, Chinese Academy of Sciences

Shengjing Wang

Tsinghua University

Secretary

Zhibo Chen

Univ. of Science and Technology of China

Website Chair

Guijin Wang

Tsinghua University

The 24th IEEE International Conference on Image Processing (ICIP) will be held in the China National Conventional Center, Beijing, China, on 17-20 September 2017. ICIP is the world's largest and most comprehensive technical conference focused on image and video processing and computer vision. The conference will feature world-class speakers, tutorials, exhibits, and a vision technology showcase.

Topics of interest include, but are not limited to:

- Filtering, Transforms, Multi-Resolution Processing
- Restoration, Enhancement, Super-Resolution
- Computer Vision Algorithms and Technologies
- Compression, Transmission, Storage, Retrieval
- Multi-View, Stereoscopic, and 3D Processing
- Multi-Temporal and Spatio-Temporal Processing
- Biometrics, Forensics, and Content Protection
- Biological and Perceptual-based Processing
- Medical Image and Video Analysis
- Document and Synthetic Visual Processing
- Color and Multispectral Processing
- Scanning, Display, and Printing
- Applications to various fields
- Computational Imaging
- Video Processing and Analytics
- Visual Quality Assessment
- Deep learning for Images and Video
- Image and Video Analysis for the Web

Paper Submission:

Authors are invited to submit papers of not more than four pages for technical content including figures and references, with one optional page containing only references. Submission Instructions, templates for the required paper format, and information on "no show" policy are available at 2017.ieeeicip.org.

Tutorials, Special Sessions, and Challenge Sessions Proposals:

Tutorials will be held on September 17, 2017. Tutorial proposals 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. For detailed submission guidelines, please refer to the tutorial proposals page. Special Sessions and Challenge Session Proposals must include a topical title, rationale, session outline, contact information, and a list of invited papers/participants. For detailed submission guidelines, please refer the ICIP 2017 website at 2017.ieeeicip.org.

Important Dates

Special Session Proposal:	Nov. 15, 2016
Notification of Special Session acceptance:	Dec. 15, 2016
Tutorial Proposal:	Dec. 15, 2016
Notification of Tutorial acceptance:	Jan. 15, 2017
Full Paper & Special Session Submission:	Jan. 31, 2017
Notification of acceptance:	Apr. 30, 2017
Camera-Ready Papers:	May. 31, 2017



Digital Object Identifier 10.1109/MSP.2016.2612718



IEEE International Conference on Image Processing

MATLAB SPEAKS WIRELESS DESIGN

You can simulate, prototype, and verify wireless systems right in MATLAB. Learn how today's MATLAB supports RF, LTE, WLAN and 5G development and SDR hardware.

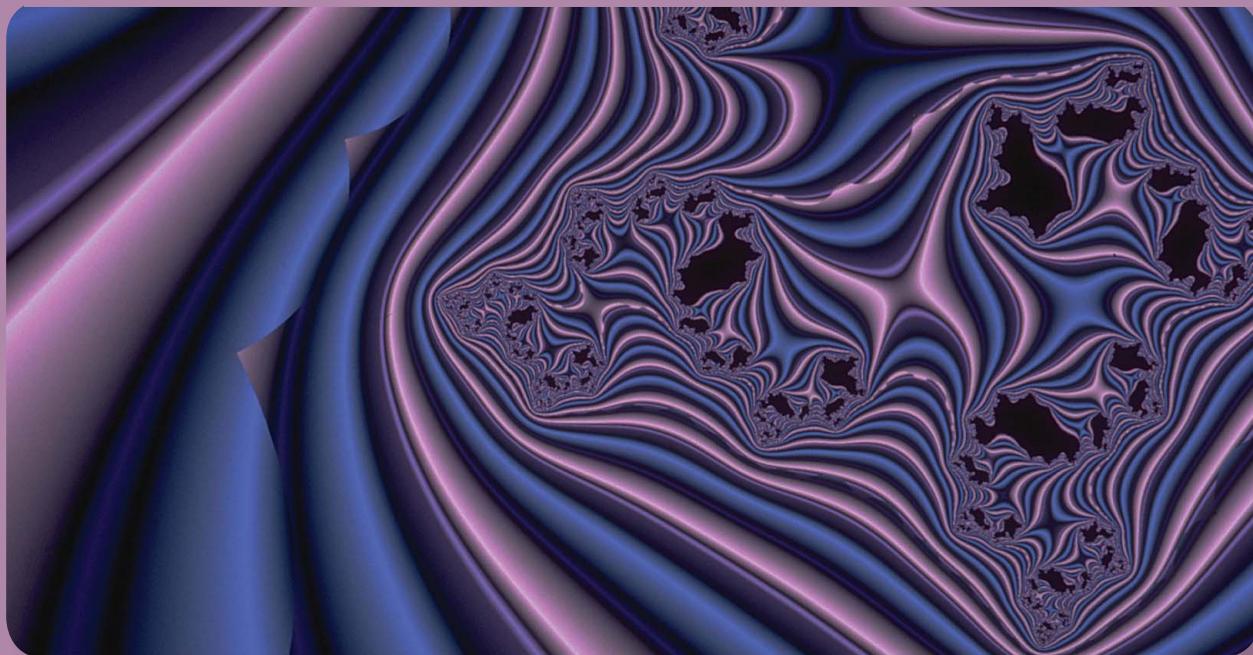
mathworks.com/wireless

IEEE SIGNAL PROCESSING SOCIETY

Content Gazette

NOVEMBER 2016

ISSN 2167-5023

**T-SP October 1 Vol. 64 #19**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7549195>**T-SP October 15 Vol. 64 #20**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7558278>**T-ASLP October Vol. 24 #10**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7574403>**T-IP October Vol. 25 #10**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7569193>**T-IFS October Vol. 11 #10**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7565682>**T-MM October Vol. 18 #10**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7569052>**T-JSTSP October Vol.10 #7**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7575599>**T-SPL October Vol. 23 #10**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7569157>**T-CI September Vol. 2 #3**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7535020>**T-SIPN September Vol. 2 #3**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7534989>

IEEE ASRU 2017

Okinawa, Japan,
December 16-20, 2017



General Chairs:

John R. Hershey, MERL
Tomohiro Nakatani, NTT

Important Dates:

Paper Submission:

June 29, 2017

Paper Notification:

August 31, 2017

Early Registration Period:

August 31 - Oct 5, 2017

Camera Ready Deadline:

Sept 21, 2017

More Information:

<http://asru2017.org>

info@asru2017.org

IEEE Automatic Speech Recognition and Understanding Workshop

The biennial IEEE ASRU workshop has a tradition of bringing together researchers from academia and industry in an intimate and collegial setting to discuss problems of common interest in automatic speech recognition, understanding, and related fields of research. The workshop includes keynotes, invited talks, poster sessions and will also feature challenge tasks, panel discussions, and demo sessions.

We invite papers in all areas of spoken language processing, with emphasis placed on the following topics:

- Automatic speech recognition
- ASR in adverse environments
- New applications of ASR
- Speech-to-speech translation
- Spoken document retrieval
- Multilingual language processing
- Spoken language understanding
- Spoken dialog systems
- Text-to-speech systems



IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



NOVEMBER 1, 2016

VOLUME 64

NUMBER 21

ITPRE D

(ISSN 1053-587X)

REGULAR PAPERS

Diffusion LMS With Correlated Regressors I: Realization-Wise Stability http://dx.doi.org/10.1109/TSP.2016.2576426	<i>M. J. Piggott and V. Solo</i>	5473
Sparse Multinomial Logistic Regression via Approximate Message Passing http://dx.doi.org/10.1109/TSP.2016.2593691	<i>E. Byrne and P. Schniter</i>	5485
Gaussian Mixture Nonlinear Filtering With Resampling for Mixand Narrowing http://dx.doi.org/10.1109/TSP.2016.2595503	<i>M. L. Psiaki</i>	5499
Compressive Sampling for Detection of Frequency-Hopping Spread Spectrum Signals http://dx.doi.org/10.1109/TSP.2016.2597122	<i>F. Liu, M. W. Marcellin, N. A. Goodman, and A. Bilgin</i>	5513
Large-System Estimation Performance in Noisy Compressed Sensing With Random Support of Known Cardinality—A Bayesian Analysis http://dx.doi.org/10.1109/TSP.2016.2591511	<i>R. Boyer, R. Couillet, B.-H. Fleury, and P. Larzabal</i>	5525



Iterative Sparsification-Projection: Fast and Robust Sparse Signal Approximation http://dx.doi.org/10.1109/TSP.2016.2585123	5536
..... <i>M. Sadeghi and M. Babaie-Zadeh</i>	
Optimal Sample Complexity for Blind Gain and Phase Calibration http://dx.doi.org/10.1109/TSP.2016.2598311	5549
..... <i>Y. Li, K. Lee, and Y. Bresler</i>	
Identification of Successive “Unobservable” Cyber Data Attacks in Power Systems Through Matrix Decomposition http://dx.doi.org/10.1109/TSP.2016.2597131	5557
..... <i>P. Gao, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky</i>	
Efficient and Non-Convex Coordinate Descent for Symmetric Nonnegative Matrix Factorization http://dx.doi.org/10.1109/TSP.2016.2591510	5571
..... <i>A. Vandaele, N. Gillis, Q. Lei, K. Zhong, and I. Dhillon</i>	
A Hamiltonian Monte Carlo Method for Non-Smooth Energy Sampling http://dx.doi.org/10.1109/TSP.2016.2585120	5585
..... <i>L. Chaari, J.-Y. Tournet, C. Chaux, and H. Batatia</i>	
Sensor Placement by Maximal Projection on Minimum Eigenspace for Linear Inverse Problems http://dx.doi.org/10.1109/TSP.2016.2573767	5595
..... <i>C. Jiang, Y. C. Soh, and H. Li</i>	
Approximate Message Passing Algorithm With Universal Denoising and Gaussian Mixture Learning http://dx.doi.org/10.1109/TSP.2016.2599484	5611
..... <i>Y. Ma, J. Zhu, and D. Baron</i>	
Low-Rank Matrix Completion in the Presence of High Coherence http://dx.doi.org/10.1109/TSP.2016.2586753	5623
..... <i>G. Liu and P. Li</i>	
Frequency Estimation From Arbitrary Time Samples http://dx.doi.org/10.1109/TSP.2016.2600507	5634
..... <i>K. Mahata and M. M. Hyder</i>	
Dynamic Filtering of Time-Varying Sparse Signals via Minimization http://dx.doi.org/10.1109/TSP.2016.2586745	5644
..... <i>A. S. Charles, A. Balavoine, and C. J. Rozell</i>	
Successive Concave Sparsity Approximation for Compressed Sensing http://dx.doi.org/10.1109/TSP.2016.2585096	5657
..... <i>M. Malek-Mohammadi, A. Koochakzadeh, M. Babaie-Zadeh, M. Jansson, and C. R. Rojas</i>	
Optimal Linear Estimators for Systems With Finite-Step Correlated Noises and Packet Dropout Compensations http://dx.doi.org/10.1109/TSP.2016.2576420	5672
..... <i>S. Sun, T. Tian, and H. Lin</i>	
A TDOA Technique with Super-Resolution Based on the Volume Cross-Correlation Function http://dx.doi.org/10.1109/TSP.2016.2548988 ..	5682
..... <i>H. Shi, H. Zhang, and X. Wang</i>	
The Ray Space Transform: A New Framework for Wave Field Processing http://dx.doi.org/10.1109/TSP.2016.2591500	5696
..... <i>L. Bianchi, F. Antonacci, A. Sarti, and S. Tubaro</i>	
Fundamental Limits in Multi-Image Alignment http://dx.doi.org/10.1109/TSP.2016.2600517	5707
..... <i>C. Aguerrebere, M. Delbracio, A. Bartesaghi, and G. Sapiro</i>	
Bounds for a Mixture of Low-Rank Compound-Gaussian and White Gaussian Noises http://dx.doi.org/10.1109/TSP.2016.2603965	5723
..... <i>O. Besson</i>	
<hr/>	
OVERVIEW ARTICLE	
MIMO Signal Processing in Offset-QAM Based Filter Bank Multicarrier Systems http://dx.doi.org/10.1109/TSP.2016.2580535	5733
..... <i>A. I. Pérez-Neira, M. Caus, R. Zakaria, D. Le Ruyet, E. Kofidis, M. Haardt, X. Mestre, and Y. Cheng</i>	

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



NOVEMBER 15, 2016

VOLUME 64

NUMBER 22

ITPREP

(ISSN 1053-587X)

REGULAR PAPERS

Partial Coherence Estimation via Spectral Matrix Shrinkage under Quadratic Loss http://dx.doi.org/10.1109/TSP.2016.2582464	<i>D. Schneider-Luftman and A. T. Walden</i>	5767
Bounds on the Number of Measurements for Reliable Compressive Classification http://dx.doi.org/10.1109/TSP.2016.2599496	<i>H. Reboledo, F. Renna, R. Calderbank, and M. R. D. Rodrigues</i>	5778
Robust Covariance Matrix Estimation in Heterogeneous Low Rank Context http://dx.doi.org/10.1109/TSP.2016.2599494	<i>A. Breloy, G. Ginolhac, F. Pascal, and P. Forster</i>	5794



Sampling and Reconstruction of Shapes With Algebraic Boundaries http://dx.doi.org/10.1109/TSP.2016.2591505	5807
..... <i>M. Fatemi, A. Amini, and M. Vetterli</i>	
Power Waveforming: Wireless Power Transfer Beyond Time Reversal http://dx.doi.org/10.1109/TSP.2016.2601283	5819
..... <i>M.-L. Ku, Y. Han, H.-Q. Lai, Y. Chen, and K. J. R. Liu</i>	
A New Look at the Radar Detection Problem http://dx.doi.org/10.1109/TSP.2016.2598312	5835
..... <i>E. Grossi, M. Lops, and L. Venturino</i>	
Robust Design of Radar Doppler Filters http://dx.doi.org/10.1109/TSP.2016.2576423	5848
..... <i>A. Aubry, A. De Maio, Y. Huang, and M. Piezso</i>	
Quickest Sequential Multiband Spectrum Sensing With Mixed Observations http://dx.doi.org/10.1109/TSP.2016.2602802	5861
..... <i>J. Geng, W. Xu, and L. Lai</i>	
Old Bands, New Tracks—Revisiting the Band Model for Robust Hypothesis Testing http://dx.doi.org/10.1109/TSP.2016.2600505	5875
..... <i>M. Fauß and A. M. Zoubir</i>	
Dithered Quantization via Orthogonal Transformations http://dx.doi.org/10.1109/TSP.2016.2599482	5887
..... <i>R. Hadad and U. Erez</i>	
Energy-Efficient Resource Optimization for OFDMA-Based Multi-Homing Heterogenous Wireless Networks http://dx.doi.org/10.1109/TSP.2016.2599481	5901
..... <i>W. Wu, Q. Yang, P. Gong, and K. S. Kwak</i>	
Genie Tree and Degrees of Freedom of the Symmetric MIMO Interfering Broadcast Channel http://dx.doi.org/10.1109/TSP.2016.2586747 ..	5914
..... <i>T. Liu and C. Yang</i>	
Weighted ADMM for Fast Decentralized Network Optimization http://dx.doi.org/10.1109/TSP.2016.2602803	5930
..... <i>Q. Ling, Y. Liu, W. Shi, and Z. Tian</i>	
Source Coding in Networks With Covariance Distortion Constraints http://dx.doi.org/10.1109/TSP.2016.2603973	5943
..... <i>A. Zahedi, J. Østergaard, S. H. Jensen, P. A. Naylor, and S. Bech</i>	
A Unifying Approach for Disturbance Cancellation and Target Detection in Passive Radar Using OFDM http://dx.doi.org/10.1109/TSP.2016.2600511	5959
..... <i>G. Gassier, G. Chabriel, J. Barrère, F. Briolle, and C. Jauffret</i>	
New Look on q^2 -Point Fast Fourier Transforms http://dx.doi.org/10.1109/TSP.2016.2598325	5972
..... <i>A. M. Grigoryan and S. S. Aghaian</i>	
Signal Periodic Decomposition With Conjugate Subspaces http://dx.doi.org/10.1109/TSP.2016.2600509	5981
..... <i>S. Deng and J. Han</i>	
Simultaneous Ranging and Self-Positioning in Unsynchronized Wireless Acoustic Sensor Networks http://dx.doi.org/10.1109/TSP.2016.2603972	5993
..... <i>M. Cobos, J. J. Perez-Solano, Ó. Belmonte, G. Ramos, and A. M. Torres</i>	
Optimal Network Beamforming in Collaborative Relay Networks With Centralized Energy Harvesting http://dx.doi.org/10.1109/TSP.2016.2582468	6005
..... <i>A. Gavili and S. ShahbazPanahi</i>	
Signal-Adapted Tight Frames on Graphs http://dx.doi.org/10.1109/TSP.2016.2591513	6017
..... <i>H. Behjat, U. Richter, D. Van De Ville, and L. Sörnmo</i>	
Two Algorithms for Modeling and Tracking of Dynamic Time-Frequency Spectra http://dx.doi.org/10.1109/TSP.2016.2602799	6030
..... <i>D. Carevic and S. Davey</i>	
Robust Adaptive Beamforming Based on Conjugate Gradient Algorithms http://dx.doi.org/10.1109/TSP.2016.2605075	6046
..... <i>M. Zhang, A. Zhang, and Q. Yang</i>	
Low Complexity Iterative Algorithms for Power Estimation in Ultra-Dense Load Coupled Networks http://dx.doi.org/10.1109/TSP.2016.2573764	6058
..... <i>R. L. G. Cavalcante, S. Stańczak, J. Zhang, and H. Zhuang</i>	



CALL FOR PAPERS

MMSP 2017 is the IEEE 19th International Workshop on Multimedia Signal Processing. The workshop is organized by the Multimedia Signal Processing Technical Committee of the IEEE Signal Processing Society. This year's event has a theme of 'Multimedia Processing for Healthcare and Assisted Living'.

Recent advances in multimedia processing and communications have potential to significantly advance current healthcare and assisted living services by enabling remote health monitoring, remote diagnostics, increased patient privacy, robotic-assisted surgery, and home-based treatment. A huge diversity of multimedia processing techniques, ranging from image/audio sensing, compression and networking, denoising, feature extraction, security, distributed processing, depth image processing, cloud and social computing, visualization and multimedia big data analytics, can all find their applications in future healthcare and elderly-care. However, to fully realize this potential and embed multimedia solutions into day-to-day healthcare practice, many challenges need to be overcome that call for significant engineering innovation, which can only happen through close interdisciplinary effort. The MMSP-2017 Workshop will bring together experts from different fields, including signal processing, computer science, communications, medicine, rehabilitation, psychology, to exchange ideas on how multimedia research can support advancements of future healthcare and how best to facilitate interactions between multimedia researchers and healthcare and assisted living providers.

Papers are solicited in (but not limited to) the following areas, covering not only this year's workshop theme, but also the general scope of multimedia signal processing:

- Multimedia big data analytics
- Distributed multimedia for body networks
- Deep learning for health-specific event detection and classification
- Streaming, security and privacy for healthcare
- Sparsity-based and low-rank based sensing of human vital signs
- Multimedia for smart homes and elderly care
- Multimedia processing for tele-rehabilitation
- Computational imaging for healthcare applications
- Healthcare monitoring applications using wearable technologies
- Image/video/speech/audio coding and processing
- Multimedia networking
- Multimedia traffic, communications and heterogeneous interactions
- Multimedia quality assessment
- Internet of Things (IoT)-based multimedia systems and applications
- Multimedia hardware design
- Augmented, mixed and virtual reality

Important dates:

Proposals for Special Sessions and Tutorials: **April 1, 2017**

Notification of Acceptance for Special Session and Tutorial Proposals: **April 15, 2017**

Submission of Regular and Special Session Papers: **June 1, 2017**

Notification of Acceptance for Regular and Special Session Papers: **July 15, 2017**

Submission of Sketch and Demo Papers: **August 1, 2017**

Camera Ready Deadline: **August 1, 2017**

General Co-chairs:

Vladan VELISAVLJEVIC,
University of Bedfordshire, UK
Vladimir STANKOVIC,
University of Strathclyde, UK
Zixiang XIONG,
Texas A&M University, TX, USA

TPC Co-chairs:

Homer H. CHEN,
National Taiwan University, Taiwan
Frederic DUFAUX,
CNRS, France
Raouf HAMZAOU,
De Montfort University, UK

Finance Co-chairs:

Ching-Te CHIU,
National Tsing Hua University, Taiwan
Lina STANKOVIC,
University of Strathclyde, UK

Special Session & Panel Co-chairs:

Gene CHEUNG,
National Institute for Informatics, Japan
Tasos DAGIUKLAS,
London South Bank University, UK

Keynote Co-chairs:

Nikolaos BOULGOURIS,
Brunel University London, UK
Yiannis ANDREOPOULOS,
University College London, UK

Industry & Demo Chair:

Marta MRAK,
BBC R&D, UK

Publication Chair:

Maria MARTINI,
Kingston University London, UK
Safak DOGAN,
Loughborough University London, UK

Publicity Chair:

Ivana TOSIC,
Ricoh, CA, USA

Asia & Australia Liaison:

Tao MEI,
Microsoft Research Beijing, China

North & South America Liaison:

Jacob CHAKARESKEI,
University of Alabama, AL, USA

Local Arrangements Chair:

Nicholas GARDNER,
University of Bedfordshire, UK





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

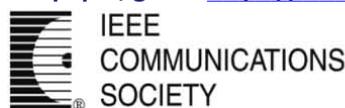
- Processing, analysis, and visualization of big data
- Signal and information processing for crowd computing
- Signal and information processing for the Internet of Things
- Emergence of behavior

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>



<http://www.signalprocessingsociety.org/>



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



OCTOBER 2016

VOLUME 24

NUMBER 10

ITASFA

(ISSN 2329-9290)

REGULAR PAPERS

Estimation of Room Acoustic Parameters: The ACE Challenge http://dx.doi.org/10.1109/TASLP.2016.2577502	1681
Efficient Implementation of Global Variance Compensation for Parametric Speech Synthesis http://dx.doi.org/10.1109/TASLP.2016.2580298	1694
Generative Modeling of Pseudo-Whisper for Robust Whispered Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2580944	1705
A Generalized Nonnegative Tensor Factorization Approach for Distant Speech Recognition With Distributed Microphones http://dx.doi.org/10.1109/TASLP.2016.2580943	1721
Adaptive Filtered-x Algorithms for Room Equalization Based on Block-Based Combination Schemes http://dx.doi.org/10.1109/TASLP.2016.2583065	1732



Variational Bayesian Inference for Source Separation and Robust Feature Extraction http://dx.doi.org/10.1109/TASLP.2016.2583794	1746
..... <i>K. Adiloğlu, and E. Vincent</i>	
Auditory Model-Based Dynamic Compression Controlled by Subband Instantaneous Frequency and Speech Presence Probability Estimates http://dx.doi.org/10.1109/TASLP.2016.2584705	1759
..... <i>S. Kortlang, G. Grimm, V. Hohmann, B. Kollmeier, and S. D. Ewert</i>	
Differentiable Pooling for Unsupervised Acoustic Model Adaptation http://dx.doi.org/10.1109/TASLP.2016.2584700	1773
..... <i>P. Swietojanski, and S. Renals</i>	
Optimal Microphone Array Observation for Clear Recording of Distant Sound Sources http://dx.doi.org/10.1109/TASLP.2016.2585879	1785
..... <i>K. Niwa, Y. Hioka, and K. Kobayashi</i>	
Spherical Harmonic Signal Covariance and Sound Field Diffuseness http://dx.doi.org/10.1109/TASLP.2016.2585862	1796
..... <i>N. Epain, and C. T. Jin</i>	
Near and Far Field Speech-in-Noise Intelligibility Improvements Based on a Time-Frequency Energy Reallocation Approach http://dx.doi.org/10.1109/TASLP.2016.2585864	1808
..... <i>T.-C. Zorilă, Y. Stylianou, T. Ishihara, and M. Akamine</i>	
Similar Word Model for Unfrequent Word Enhancement in Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2585863	1819
..... <i>X. Ma, D. Wang, and J. Tejedor</i>	
Summarizing Meeting Transcripts Based on Functional Segmentation http://dx.doi.org/10.1109/TASLP.2016.2585859	1831
..... <i>M. H. Bokaei, H. Sameti, and Y. Liu</i>	
Abstractive Cross-Language Summarization via Translation Model Enhanced Predicate Argument Structure Fusing http://dx.doi.org/10.1109/TASLP.2016.2586608	1842
..... <i>J. Zhang, Y. Zhou, and C. Zong</i>	
A Morphological Model for Simulating Acoustic Scenes and Its Application to Sound Event Detection http://dx.doi.org/10.1109/TASLP.2016.2587218	1854
..... <i>G. Lafay, M. Lagrange, M. Rossignol, E. Benetos, and A. Roebel</i>	
Parallel Reference Speaker Weighting for Kinematic-Independent Acoustic-to-Articulatory Inversion http://dx.doi.org/10.1109/TASLP.2016.2588340	1865
..... <i>A. Ji, M. T. Johnson, and J. J. Berry</i>	

EDICS—Editor’s Information Classification Scheme http://dx.doi.org/10.1109/TASLP.2016.2607881	1876
Information for Authors http://dx.doi.org/10.1109/TASLP.2016.2607883	1878

Call for Papers

Special Issue on Biosignal-based Spoken Communication

in the IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)

Speech is a complex process emitting a wide range of biosignals, including, but not limited to, acoustics. These biosignals – stemming from the articulators, the articulator muscle activities, the neural pathways, or the brain itself – can be used to circumvent limitations of conventional speech processing in particular, and to gain insights into the process of speech production in general.

Research on biosignal-based speech capturing and processing is a wide and very active field at the intersection of various disciplines, ranging from engineering, electronics and machine learning to medicine, neuroscience, physiology, and psychology. Consequently, a variety of methods and approaches are thoroughly investigated, aiming towards the common goal of creating biosignal-based speech processing devices and applications for everyday use, as well as for spoken communication research purposes.

We aim at bringing together studies covering these various modalities, research approaches, and objectives in a

Special issue of the IEEE Transactions on Audio, Speech, and Language Processing
entitled
Biosignal-based Spoken Communication.

For this purpose we will invite papers describing previously unpublished work in the following broad areas:

- Capturing methods for speech-related biosignals: tracing of articulatory activity (e.g. EMA, PMA, ultrasound, video), electrical biosignals (e.g. EMG, EEG, ECG, NIRS), acoustic sensors for capturing whispered / murmured speech (e.g. NAM microphone), etc.
- Signal processing for speech-related biosignals: feature extraction, denoising, source separation, etc.
- Speech recognition based on biosignals (e.g. silent speech interface, recognition in noisy environment, etc.).
- Mapping between speech-related biosignals and speech acoustics (e.g. articulatory-acoustic mapping)
- Modeling of speech units: articulatory or phonetic features, visemes, etc.
- Multi-modality and information fusion in speech recognition
- Challenges of dealing with whispered, mumbled, silently articulated, or inner speech
- Neural Representations of speech and language
- Novel approaches in physiological studies of speech planning and production
- Brain-computer-interface (BCI) for restoring speech communication
- User studies in biosignal-based speech processing
- End-to-end systems and devices
- Applications in rehabilitation and therapy

Submission Deadline: November 2016

Notification of Acceptance: January 2017

Final Manuscript Due: April 2017

Tentative Publication Date: First half of 2017

Editors:

Tanja Schultz (Universität Bremen, Germany) tanja.schultz@uni-bremen.de (Lead Guest Editor)

Thomas Hueber (CNRS/GIPSA-lab, Grenoble, France) thomas.hueber@gipsa-lab.fr

Dean J. Krusienski (ASPEN Lab, Old Dominion University) dkrusien@odu.edu

Jonathan Brumberg (Speech-Language-Hearing Department, University of Kansas) brumberg@ku.edu

Call for Papers
IEEE Signal Processing Society
IEEE Transactions on Signal and Information Processing over Networks

Special Issue on Distributed Signal Processing for Security and Privacy in Networked Cyber-Physical Systems

GUEST EDITORS:

Arash Mohammadi, Concordia University, Montreal, Canada, arashmoh@encs.concordia.ca
Peng Cheng, Zhejiang University, Hangzhou, China, pcheng@iipc.zju.edu.cn
Vincenzo Piuri, Università degli Studi di Milano, Milan, Italy, vincenzo.piuri@unimi.it
Konstantinos N. Plataniotis, University of Toronto, Toronto, Canada, kostas@ece.utoronto.ca
Patrizio Campisi, Università degli Studi Roma Tre, Italy, patrizio.campisi@uniroma3.it

SCOPE

The focus of this special issue is on distributed information acquisition, estimation, and adaptive learning for security and privacy in the context of networked cyber-physical systems (CPSs) which are engineering systems with integrated computational and communication capabilities that interact with humans through cyber space. The CPSs have recently emerged in several practical applications of engineering importance including aerospace, industrial/manufacturing process control, multimedia networks, transportation systems, power grids, and medical systems. The CPSs typically consist of both wireless and wired sensor/agent networks with different capacity/reliability levels where the emphasis is on real-time operations, and performing distributed, secure, and optimal sensing/processing is the key concern. To satisfy these requirements of the CPSs, it is of paramount importance to design innovative “Signal Processing” tools to provide unprecedented performance and resource utilization efficiency.

A significant challenge for implementation of signal processing solutions in CPSs is the difficulty of acquiring data from geographically distributed observation nodes and storing/processing the aggregated data at the fusion center (FC). As such, there has been a recent surge of interest in development of distributed and collaborative signal processing technologies where adaptation, estimation, and/or control are performed locally and communication is limited to local neighborhoods. Distributed signal processing over networked CPSs, however, raise significant privacy and security concerns as local observations are being shared by neighboring nodes in a collaborative and iterative fashion. On one hand, applications of CPSs are severely safety critical where potential cyber and physical attacks by adversaries on signal processing modules could lead to a variety of severe consequences including customer information leakage, destruction of infrastructures, and endangering human lives. On the other hand, the need for cooperation between neighboring nodes makes it imperative to prevent the disclosure of sensitive local information during distributed information fusion step. At the same time, efficient usage of available resources (communication, computation, bandwidth, and energy) is a prerequisite for productive operation of the CPSs. To accommodate these critical aspects of CPSs, it is of great practical importance and theoretical significance to develop advanced “Secure and Privacy Preserving Distributed Signal Processing” solutions.

The spirit and wide scope of distributed signal processing in revolutionized CPSs calls for novel and innovative techniques beyond conventional approaches to provide precise guarantees on security and privacy of CPSs. The objective of this special issue is to further advance recent developments of distributed signal processing to practical aspects of CPSs for real-time processing and monitoring of the underlying system in a secure and privacy preserving manner while avoiding degradation of the processing performance and preserving the valuable resources. To provide a systematic base for future advancements of CPSs, this special issue aims to provide a research venue to investigate distributed signal processing techniques with adaptation, cooperation, and learning capabilities which are secure against cyber-attacks and protected against privacy leaks. The emphasis of this special issue is on distributed/network aspects of security and privacy in CPSs. Papers with primary emphasis on forensics and security will be redirected to IEEE Transactions on Information Forensics and Security (TIFS). Topics of interest include, but are not limited to:

- Security and Privacy of distributed signal processing in networked CPSs.
- Distributed and secure detection, estimation, and information fusion.
- Security and privacy of consensus and diffusive strategies in networked systems.
- Secure and privacy preserving distributed adaptation and learning.
- Security and privacy of distributed sensor resource management in networked systems.
- Distributed event-based estimation/control in networked CPSs.
- Detection and identification of potential attacks on distributed signal processing mechanisms.
- Application domains including but not limited to, smart grids, camera networks, multimedia network, and vehicular networks.

SUBMISSION GUIDELINES

Authors are invited to submit original research contributions by following the detailed instructions given in the “Information for Authors” at <http://www.signalprocessingsociety.org/publications/periodicals/tsipn/>. Manuscripts should be submitted via Scholar One (Manuscript Central) at <http://mc.manuscriptcentral.com/tsipn-ieee>. Questions about the special issue should be directed to the Guest Editors.

IMPORTANT DATES

Paper submission deadline:	December 15, 2016	Final notification:	September 1, 2017
Notification of the first review:	March 1, 2017	Final manuscript due:	October 15, 2017
Revised paper submission:	April 15, 2017	Publication:	Advance posting in IEEE explore as soon as authors approve galley proofs
Notification of the re-review:	June 15, 2017	Expected inclusion in an issue:	March 2018
Minor revision deadline:	August 1, 2017		

IEEE TRANSACTIONS ON IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



OCTOBER 2016

VOLUME 25

NUMBER 10

IIPRE4

(ISSN 1057-7149)

PAPERS

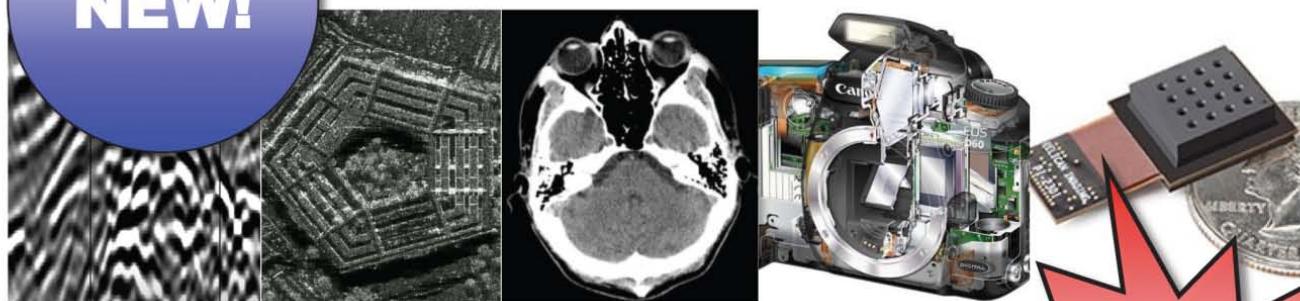
VITRAIL: Acquisition, Modeling, and Rendering of Stained Glass http://dx.doi.org/10.1109/TIP.2016.2585041	4475
..... <i>N. Thanikachalam, L. Baboulaz, P. Prandoni, S. Trümpler, S. Wolf, and M. Vetterli</i>	
Adaptive Image Denoising by Mixture Adaptation http://dx.doi.org/10.1109/TIP.2016.2590318	4489
..... <i>E. Luo, S. H. Chan, and T. Q. Nguyen</i>	
Query-Driven Approach to Face Clustering and Tagging http://dx.doi.org/10.1109/TIP.2016.2592703	4504
..... <i>L. Zhang, X. Wang, D. V. Kalashnikov, S. Mehrotra, and D. Ramanan</i>	
Query-Adaptive Hash Code Ranking for Large-Scale Multi-View Visual Search http://dx.doi.org/10.1109/TIP.2016.2593344	4514
..... <i>X. Liu, L. Huang, C. Deng, B. Lang, and D. Tao</i>	
Scale-Aware Pixelwise Object Proposal Networks http://dx.doi.org/10.1109/TIP.2016.2593342	4525
..... <i>Z. Jie, X. Liang, J. Feng, W. F. Lu, E. H. F. Tay, and S. Yan</i>	
Multimodal Discriminative Binary Embedding for Large-Scale Cross-Modal Retrieval http://dx.doi.org/10.1109/TIP.2016.2592800	4540
..... <i>D. Wang, X. Gao, X. Wang, L. He, and B. Yuan</i>	
Visual Tracking via Coarse and Fine Structural Local Sparse Appearance Models http://dx.doi.org/10.1109/TIP.2016.2592701	4555
..... <i>X. Jia, H. Lu, and M.-H. Yang</i>	
Hyperspectral Unmixing in Presence of Endmember Variability, Nonlinearity, or Mismodeling Effects http://dx.doi.org/10.1109/TIP.2016.2590324	4565
..... <i>A. Halimi, P. Honeine, and J. M. Bioucas-Dias</i>	
Sparse Representation With Spatio-Temporal Online Dictionary Learning for Promising Video Coding http://dx.doi.org/10.1109/TIP.2016.2594490	4580
..... <i>W. Dai, Y. Shen, X. Tang, J. Zou, H. Xiong, and C. W. Chen</i>	
Image Segmentation Using Hierarchical Merge Tree http://dx.doi.org/10.1109/TIP.2016.2592704	4596
..... <i>T. Liu, M. Seyedhosseini, and T. Tasdizen</i>	
Enhanced Use of Mattes for Easy Image Composition http://dx.doi.org/10.1109/TIP.2016.2593345	4608
..... <i>W. Wang, P. Xu, X. Bie, and M. Hua</i>	
Web Image Search Re-Ranking With Click-Based Similarity and Typicality http://dx.doi.org/10.1109/TIP.2016.2593653	4617
..... <i>X. Yang, T. Mei, Y. Zhang, J. Liu, and S. Satoh</i>	



Deformable Registration of Biomedical Images Using 2D Hidden Markov Models http://dx.doi.org/10.1109/TIP.2016.2592702	<i>R. Shenoy, M.-C. Shih, and K. Rose</i>	4631
Low-Rank Tensor Subspace Learning for RGB-D Action Recognition http://dx.doi.org/10.1109/TIP.2016.2589320	<i>C. Jia and Y. Fu</i>	4641
Detecting Salient Objects via Color and Texture Compactness Hypotheses http://dx.doi.org/10.1109/TIP.2016.2594489	<i>P. Hu, W. Wang, C. Zhang, and K. Lu</i>	4653
Automatic Reference Color Selection for Adaptive Mathematical Morphology and Application in Image Segmentation http://dx.doi.org/10.1109/TIP.2016.2586658	<i>H.-C. Shih and E.-R. Liu</i>	4665
Robust Low-Rank Matrix Factorization Under General Mixture Noise Distributions http://dx.doi.org/10.1109/TIP.2016.2593343	<i>X. Cao, Q. Zhao, D. Meng, Y. Chen, and Z. Xu</i>	4677
Neutro-Connectedness Cut http://dx.doi.org/10.1109/TIP.2016.2594485	<i>M. Xian, Y. Zhang, H.-D. Cheng, F. Xu, and J. Ding</i>	4691
False Discovery Rate Approach to Unsupervised Image Change Detection http://dx.doi.org/10.1109/TIP.2016.2593340	<i>V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia</i>	4704
Spatial Mutual Information and PageRank-Based Contrast Enhancement and Quality-Aware Relative Contrast Measure http://dx.doi.org/10.1109/TIP.2016.2599103	<i>T. Celik</i>	4719
Behavior Knowledge Space-Based Fusion for Copy-Move Forgery Detection http://dx.doi.org/10.1109/TIP.2016.2593583	<i>A. Ferreira, S. C. Felipussi, C. Alfaro, P. Fonseca, J. E. Vargas-Muñoz, J. A. dos Santos, and A. Rocha</i>	4729
Image Segmentation via Probabilistic Graph Matching http://dx.doi.org/10.1109/TIP.2016.2590832	<i>A. Heimowitz and Y. Keller</i>	4743
Confidence Preserving Machine for Facial Action Unit Detection http://dx.doi.org/10.1109/TIP.2016.2594486	<i>J. Zeng, W.-S. Chu, F. De la Torre, J. F. Cohn, and Z. Xiong</i>	4753
Universal Background Subtraction Using Word Consensus Models http://dx.doi.org/10.1109/TIP.2016.2598691	<i>P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin</i>	4768
Kernel Learning for Dynamic Texture Synthesis http://dx.doi.org/10.1109/TIP.2016.2598653	<i>X. You, W. Guo, S. Yu, K. Li, J. C. Principe, and D. Tao</i>	4782
Compressive Sensing and Recovery for Binary Images http://dx.doi.org/10.1109/TIP.2016.2598651	<i>J.-H. Ahn</i>	4796
Learning Fast Low-Rank Projection for Image Classification http://dx.doi.org/10.1109/TIP.2016.2598654	<i>J. Li, Y. Kong, H. Zhao, J. Yang, and Y. Fu</i>	4803
Image Enhancement by Entropy Maximization and Quantization Resolution Upconversion http://dx.doi.org/10.1109/TIP.2016.2598485	<i>Y. Niu, X. Wu, and G. Shi</i>	4815
A Spatial Layout and Scale Invariant Feature Representation for Indoor Scene Classification http://dx.doi.org/10.1109/TIP.2016.2599292 ..	<i>M. Hayat, S. H. Khan, M. Bennamoun, and S. An</i>	4829
Weighted Schatten p-Norm Minimization for Image Denoising and Background Subtraction http://dx.doi.org/10.1109/TIP.2016.2599290 ..	<i>Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang, and L. Zhang</i>	4842
Coarse-to-Fine Description for Fine-Grained Visual Categorization http://dx.doi.org/10.1109/TIP.2016.2599102	<i>H. Yao, S. Zhang, Y. Zhang, J. Li, and Q. Tian</i>	4858
A Framework for Depth Video Reconstruction From a Subset of Samples and Its Applications http://dx.doi.org/10.1109/TIP.2016.2598484 ..	<i>L.-K. Liu and T. Q. Nguyen</i>	4873
2D Matching Using Repetitive and Salient Features in Architectural Images http://dx.doi.org/10.1109/TIP.2016.2598612	<i>B. Morago, G. Bui, and Y. Duan</i>	4888
Boosted Dictionary Learning for Image Compression http://dx.doi.org/10.1109/TIP.2016.2598483	<i>M. Nejati, S. Samavi, N. Karimi, S. M. R. Soroushmehr, and K. Najarian</i>	4900
DCT-QM: A DCT-Based Quality Degradation Metric for Image Quality Optimization Problems http://dx.doi.org/10.1109/TIP.2016.2598492	<i>S.-H. Bae and M. Kim</i>	4916
Large-Scale Fiber Tracking Through Sparsely Sampled Image Sequences of Composite Materials http://dx.doi.org/10.1109/TIP.2016.2598640	<i>Y. Zhou, H. Yu, J. Simmons, C. P. Przybyla, and S. Wang</i>	4931
Removing Turbulence Effect via Hybrid Total Variation and Deformation-Guided Kernel Regression http://dx.doi.org/10.1109/TIP.2016.2598638	<i>Y. Xie, W. Zhang, D. Tao, W. Hu, Y. Qu, and H. Wang</i>	4943
Robust Visual Knowledge Transfer via Extreme Learning Machine-Based Domain Adaptation http://dx.doi.org/10.1109/TIP.2016.2598679	<i>L. Zhang and D. Zhang</i>	4959
Exploring Structured Sparsity by a Reweighted Laplace Prior for Hyperspectral Compressive Sensing http://dx.doi.org/10.1109/TIP.2016.2598652	<i>L. Zhang, W. Wei, C. Tian, F. Li, and Y. Zhang</i>	4974
<hr/>		
EDICS—Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIP.2016.2603712		4989
Information for Authors http://dx.doi.org/10.1109/TIP.2016.2603738		4990



IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING



Editor-in-Chief

W. Clem Karl
Boston University

Technical Committee

Charles Bouman
Eric Miller
Peter Corcoran
Jong Chul Ye
Dave Brady
William Freeman

The IEEE Transactions on Computational Imaging publishes 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:

**30% FASTER
TURN AROUND
THAN TIP!**

Computational Imaging Methods and Models

- Coded image sensing
- Compressed sensing
- Sparse and low-rank models
- Learning-based models, dictionary methods
- Graphical image models
- Perceptual models

Computational Image Formation

- Sparsity-based reconstruction
- Statistically-based inversion methods
- Multi-image and sensor fusion
- Optimization-based methods; proximal iterative methods, ADMM

Computational Photography

- Non-classical image capture
- Generalized illumination
- Time-of-flight imaging
- High dynamic range imaging
- Plenoptic imaging

Computational Consumer Imaging

- Mobile imaging, cell phone imaging
- Camera-array systems
- Depth cameras, multi-focus imaging
- Pervasive imaging, camera networks

Computational Acoustic Imaging

- Multi-static ultrasound imaging
- Photo-acoustic imaging
- Acoustic tomography

Computational Microscopy

- Holographic microscopy
- Quantitative phase imaging
- Multi-illumination microscopy
- Lensless microscopy
- Light field microscopy

Imaging Hardware and Software

- Embedded computing systems
- Big data computational imaging
- Integrated hardware/digital design

Tomographic Imaging

- X-ray CT
- PET
- SPECT

Magnetic Resonance Imaging

- Diffusion tensor imaging
- Fast acquisition

Radar Imaging

- Synthetic aperture imaging
- Inverse synthetic aperture imaging

Geophysical Imaging

- Multi-spectral imaging
- Ground penetrating radar
- Seismic tomography

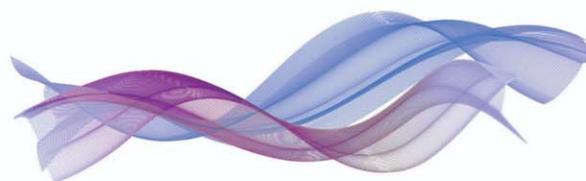
Multi-spectral Imaging

- Multi-spectral imaging
- Hyper-spectral imaging
- Spectroscopic imaging

**NOW WITH
NO PAGE
CHARGES!**

For more information on the IEEE Transactions on Computational Imaging see <http://www.signalprocessingsociety.org/publications/periodicals/tci/>





CAMSAP 2017

Call for Papers

The Seventh IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing

Curaçao, Dutch Antilles

December 10-13, 2017

<http://www.cs.huji.ac.il/conferences/CAMSAP17>



Following the success of the first six editions of the IEEE workshop on Computational Advances in Multi-Sensor Adaptive Processing, we are pleased to announce the seventh workshop in this series. IEEE CAMSAP 2017 will be held in Curaçao, Dutch Antilles, and will feature a number of plenary talks from the world's leading researchers in the area, special focus sessions, and contributed papers. All papers will undergo peer review in order to provide feedback to the authors and ensure a high-quality program.

Topics and applications of interest for the workshop include, but are not limited to, the following.

TOPICS OF INTEREST

- Array processing, waveform diversity, space-time adaptive processing
- Convex optimization and relaxation
- Computational linear & multi-linear algebra
- Computer-intensive methods in signal processing (bootstrap, MCMC, EM, particle filtering, etc.)
- Signal and information processing over networks
- Sparse signal processing

APPLICATIONS

- Big data
- Biomedical signal processing
- Communication systems
- Computational imaging
- Radar
- Sensor networks
- Smart grids
- Sonar

Submission of Papers: Prospective authors are invited to submit original full-length papers, with up to four pages for technical content including figures and references, using the formatting guidelines on the website for reviewing purposes. All accepted papers must be presented at the workshop to appear in the proceedings. Best student paper awards, selected by a CAMSAP committee, will also be presented at the workshop.

Special Session Proposals: In addition to contributed sessions, the workshop will also have a number of special sessions. Prospective organizers of special sessions are invited to submit a proposal form, available on the workshop website, by e-mail to the Special Sessions Chair.

IMPORTANT DEADLINES

Submission of proposals for special sessions	March, 2017
Notification of special session acceptance	March 15, 2015
Submission of papers	July, 2017
Notification of paper acceptance	September, 2017
Final paper submission	



General Chairs

André L. F. de Almeida
 andre@gtel.ufc.br
 Federal University of Ceará,
 Brazil

Martin Haardt
 martin.haardt@tu-ilmenau.de
 Ilmenau University of Technology,
 Germany

Technical Program Chairs

Xiaoli Ma
 xiaoli@gatech.edu
 Georgia Institute of Technology,
 USA

Shahram ShahbazPanahi
 shahram.shahbazpanahi@uoit.ca
 University of Ontario Institute of
 Technology,
 Canada

Finance Chair

Cihan Tepedelenioglu
 cihan@asu.edu
 Arizona State University,
 USA

Publicity and Publications Chair

Ami Wiesel
 ami.wiesel@huji.ac.il
 Hebrew University of Jerusalem,
 Israel

Local Arrangements Chair

Geert Leus
 g.j.t.leus@tudelft.nl
 Delft University of Technology,
 The Netherlands

Robert W. Heath Jr.
 rheath@utexas.edu
 University of Texas at Austin,
 USA



Join Twitter Conversation
#CAMSAP2017



IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

OCTOBER 2016

VOLUME 11

NUMBER 10

ITIFA6

(ISSN 1556-6013)

REGULAR PAPERS

Source Distinguishability Under Distortion-Limited Attack: An Optimal Transport Perspective http://dx.doi.org/10.1109/TIFS.2016.2570739	<i>M. Barni and B. Tondi</i>	2145
On the Fingerprinting of Software-Defined Networks http://dx.doi.org/10.1109/TIFS.2016.2573756	<i>H. Cui, G. O. Karame, F. Klaedtke, and R. Bifulco</i>	2160
A Data Exfiltration and Remote Exploitation Attack on Consumer 3D Printers http://dx.doi.org/10.1109/TIFS.2016.2578285	<i>Q. Do, B. Martini, and K.-K. R. Choo</i>	2174
Achieving Probabilistic Anonymity in a Linear and Hybrid Randomization Model http://dx.doi.org/10.1109/TIFS.2016.2562605	<i>Y. Sang, H. Shen, H. Tian, and Z. Zhang</i>	2187
A Customized Sparse Representation Model With Mixed Norm for Undersampled Face Recognition http://dx.doi.org/10.1109/TIFS.2016.2567318	<i>Z.-M. Li, Z.-H. Huang, and K. Shang</i>	2203
Intent-Based Extensible Real-Time PHP Supervision Framework http://dx.doi.org/10.1109/TIFS.2016.2569063	<i>V. Prokhorenko, K.-K. R. Choo, and H. Ashman</i>	2215
Rethinking Permission Enforcement Mechanism on Mobile Systems http://dx.doi.org/10.1109/TIFS.2016.2581304	<i>Y. Zhang, M. Yang, G. Gu, and H. Chen</i>	2227
Virus Propagation Modeling and Convergence Analysis in Large-Scale Networks http://dx.doi.org/10.1109/TIFS.2016.2581305	<i>X. Wang, W. Ni, K. Zheng, R. P. Liu, and X. Niu</i>	2241
A Unified Resource Allocation Framework for Defending Against Pollution Attacks in Wireless Network Coding Systems http://dx.doi.org/10.1109/TIFS.2016.2581313	<i>W. Tong and S. Zhong</i>	2255
Secure Face Unlock: Spoof Detection on Smartphones http://dx.doi.org/10.1109/TIFS.2016.2578288	<i>K. Patel, H. Han, and A. K. Jain</i>	2268
Design and Fabrication of 3D Fingerprint Targets http://dx.doi.org/10.1109/TIFS.2016.2581306	<i>S. S. Arora, K. Cao, A. K. Jain, and N. G. Paulter, Jr.</i>	2284



Privacy-Preserving Utility Verification of the Data Published by Non-Interactive Differentially Private Mechanisms http://dx.doi.org/10.1109/TIFS.2016.2532839	<i>J. Hua, A. Tang, Y. Fang, Z. Shen, and S. Zhong</i>	2298
Extractable Common Randomness From Gaussian Trees: Topological and Algebraic Perspectives http://dx.doi.org/10.1109/TIFS.2016.2543688	<i>A. Moharrer, S. Wei, G. T. Amariuca, and J. Deng</i>	2312
Integrated Software Fingerprinting via Neural-Network-Based Control Flow Obfuscation http://dx.doi.org/10.1109/TIFS.2016.2555287	<i>H. Ma, R. Li, X. Yu, C. Jia, and D. Gao</i>	2322
Personal Identification Using Minor Knuckle Patterns From Palm Dorsal Surface http://dx.doi.org/10.1109/TIFS.2016.2574309	<i>A. Kumar and Z. Xu</i>	2338
Cryptographic Hierarchical Access Control for Dynamic Structures http://dx.doi.org/10.1109/TIFS.2016.2581147	<i>A. Castiglione, A. De Santis, B. Masucci, F. Palmieri, A. Castiglione, and X. Huang</i>	2349
Privacy-Preserving and Regular Language Search Over Encrypted Cloud Data http://dx.doi.org/10.1109/TIFS.2016.2581316	<i>K. Liang, X. Huang, F. Guo, and J. K. Liu</i>	2365
From Clothing to Identity: Manual and Automatic Soft Biometrics http://dx.doi.org/10.1109/TIFS.2016.2584001	<i>E. S. Jaha and M. S. Nixon</i>	2377

EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIFS.2016.2607043		2391
Information for Authors http://dx.doi.org/10.1109/TIFS.2016.2607059		2392

ANNOUNCEMENTS

Call for Papers-IEEE Transactions on Signal and Information Processing Over Networks Special Issue on Distributed Signal Processing for Security and Privacy in Networked Cyber-Physical Systems http://dx.doi.org/10.1109/TIFS.2016.2607044		2394
Call for Papers-IEEE Journal of Selected Topics in Signal Processing Special Issue on Signal Processing and Machine Learning for Education and Human Learning at Scale http://dx.doi.org/10.1109/TIFS.2016.2607045		2395
Call for Papers-IEEE Transactions on Signal and Information Processing Over Networks Special Issue on Graph Signal Processing http://dx.doi.org/10.1109/TIFS.2016.2607046		2396



The 3rd IEEE International Conference on Network Softwarization (NetSoft 2017) will be held July 3–7, 2017 at the University of Bologna in Bologna, Italy. IEEE NetSoft is the flagship conference of the IEEE SDN Initiative which aims to address “Softwarization” of networks and systemic trends concerning the convergence of Cloud Computing, Software-Defined Networks, and Network Function Virtualization.

TOPICS OF INTEREST

Authors are invited to submit papers that fall into the area of software-defined and virtualized infrastructures. Topics of interest include, but are not limited to, the following:

- SDN and NFV as enabling technologies for 5G
- From Cloud Computing to Edge-Fog Computing
- 5G Functional Decomposition and Infrastructure slicing
- 5G sustainable ecosystems: IoT, Industry 4.0, Pervasive Robotics, Self-driving vehicles, Tactile Internet, Immersive Communications, Artificial Intelligence applications
- Software Defined infrastructures for Public Protection and Disaster Relief (PPDR) network services
- Service Function Chaining for NFV: Modeling, composition algorithms, deployment
- Intent-based interfacing for NFV
- SDN/NFV Network & Service Orchestration and Management
- Management of federated SDN/NFV infrastructure and frameworks
- Real time operations and efficient network/service monitoring in SDN/NFV
- Performance and scalability issues in NFV implementation scenarios
- Traffic Engineering and QoS/QoE in SDN/NFV
- APIs, protocols and languages for programmable networks and Software-Defined Infrastructure
- SDN switch/router architectures/designs
- SDN/NFV issues and opportunities for security, trust and privacy
- Experience reports from experimental testbeds and deployment
- Softwarized platforms for Internet-of-Things (IoT)
- New value chains and business models

SCOPE

The telecommunications landscape will change radically in the next few years. Pervasive ultra-broadband, programmable networks, and cost reduction of IT systems are paving the way to new services and commoditization of telecommunications infrastructure while lowering entrance barriers for new players and giving rise to new value chains. While this results in considerable challenges for service providers, this transformation also brings unprecedented opportunities for the Digital Society and the Digital Economy related to emerging new services and applications. Examples include Tactile Internet of Things, Industry 4.0, Cloud Robotics, and Artificial Intelligence. 5G will both exploit and accelerate this transformation.

NetSoft 2017 aims to capture the theme of “Softwarization Sustaining a Hyper-connected World: en route to 5G” and serve as forum for researchers to discuss the latest advances in this area. NetSoft 2017 will feature technical paper, keynotes, tutorials, and demos and exhibits from world-leading experts representing operators, vendors, research institutes, open source projects, and academia.

PAPER SUBMISSION

Authors are invited to submit original contributions (written in English) in PDF format. Only original papers not published or submitted for publication elsewhere can be submitted. Papers can be of two types: full (up to 9 pages) or short (up to 5 pages) papers. Full Papers accepted as short Papers will be required to be reduced to 5-pages length. Papers should be in IEEE 2-column US-Letter style using IEEE Conference template (http://www.ieee.org/conferences_events/conferences/publishing/templates.html) and submitted in PDF format via JEMS at: <https://jems.sbc.org.br/home.cgi?c=2657>. Papers exceeding these limits, multiple submissions, and self-plagiarized papers will be rejected without further review. All submitted papers will be subject to a peer-review process. The accepted papers will be published in IEEE Xplore, provided that the authors do present their paper at the conference.

IMPORTANT DATES

December 5, 2016: Technical Papers deadline

December 15, 2016: Workshop Submission deadline

March 6, 2017: Paper submission acceptance notification

April 10, 2017: Full Conference Paper; Speaker registration

May 30, 2017: Early-Bird Registration

GENERAL CO-CHAIRS

Antonio Manzalini (Telecom Italia Mobile, Italy)

Roberto Verdone (University of Bologna, Italy)

STEERING COMMITTEE CHAIR

Prosper Chemouil (Orange Labs, France)

TPC CO-CHAIRS

Franco Callegati (University of Bologna, Italy)

Alexander Clemm (Cisco, USA)

Kohei Shiimoto (NTT Labs, Japan)

IEEE TRANSACTIONS ON **MULTIMEDIA**

A PUBLICATION OF
THE IEEE CIRCUITS AND SYSTEMS SOCIETY
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY



<http://www.signalprocessingsociety.org/tmm/>

OCTOBER 2016

VOLUME 18

NUMBER 10

ITMUF8

(ISSN 1520-9210)

SPECIAL SECTION ON MULTIMEDIA-BASED HEALTHCARE

GUEST EDITORIAL

Multimedia-Based Healthcare <http://dx.doi.org/10.1109/TMM.2016.2606738> *J. You, V. Bhagavatula, V. Piuri, and D. Zhang* 1925

SPECIAL SECTION PAPERS

Classification-Based Record Linkage With Pseudonymized Data for Epidemiological Cancer Registries http://dx.doi.org/10.1109/TMM.2016.2598482	<i>Y. Siegert, X. Jiang, V. Krieg, and S. Bartholomäus</i>	1929
ConfidentCare: A Clinical Decision Support System for Personalized Breast Cancer Screening http://dx.doi.org/10.1109/TMM.2016.2589160	<i>A. M. Alaa, K. H. Moon, W. Hsu, and M. van der Schaar</i>	1942
Kernel Combined Sparse Representation for Disease Recognition http://dx.doi.org/10.1109/TMM.2016.2602062	<i>Q. Feng and Y. Zhou</i>	1956
Audiovisual Spatial-Audio Analysis by Means of Sound Localization and Imaging: A Multimedia Healthcare Framework in Abdominal Sound Mapping http://dx.doi.org/10.1109/TMM.2016.2594148	<i>C. A. Dimoulas</i>	1969
Tensor Manifold Discriminant Projections for Acceleration-Based Human Activity Recognition http://dx.doi.org/10.1109/TMM.2016.2597007	<i>Y. Guo, D. Tao, J. Cheng, A. Dougherty, Y. Li, K. Yue, and B. Zhang</i>	1977
Multiple Video Delivery in m-Health Emergency Applications http://dx.doi.org/10.1109/TMM.2016.2597001	<i>S. Cicalò, M. Mazzotti, S. Moretti, V. Tralli, and M. Chiani</i>	1988
Enabling Secure and Fast Indexing for Privacy-Assured Healthcare Monitoring via Compressive Sensing http://dx.doi.org/10.1109/TMM.2016.2602758	<i>X. Yuan, X. Wang, C. Wang, J. Weng, and K. Ren</i>	2002



REGULAR PAPERS

Compression and Coding

- Depth Map Down-Sampling and Coding Based on Synthesized View Distortion <http://dx.doi.org/10.1109/TMM.2016.2594145> C. Yao, J. Xiao, T. Tillo, Y. Zhao, C. Lin, and H. Bai 2015
- λ -Domain Rate Control Algorithm for HEVC Scalable Extension <http://dx.doi.org/10.1109/TMM.2016.2595264> L. Li, B. Li, D. Liu, and H. Li 2023
- Sensing Matrix Optimization Based on Equiangular Tight Frames With Consideration of Sparse Representation Error <http://dx.doi.org/10.1109/TMM.2016.2595261> H. Bai, S. Li, and X. He 2040
- Efficient Residual DPCM Using an l_1 Robust Linear Prediction in Screen Content Video Coding <http://dx.doi.org/10.1109/TMM.2016.2595259> J.-W. Kang, S.-K. Ryu, N.-Y. Kim, and M.-J. Kang 2054

Image/Video/Graphics Analysis and Synthesis

- Learning Cascaded Deep Auto-Encoder Networks for Face Alignment <http://dx.doi.org/10.1109/TMM.2016.2591508> R. Weng, J. Lu, Y.-P. Tan, and J. Zhou 2066
- Animal Detection From Highly Cluttered Natural Scenes Using Spatiotemporal Object Region Proposals and Patch Verification <http://dx.doi.org/10.1109/TMM.2016.2594138> Z. Zhang, Z. He, G. Cao, and W. Cao 2079
- Background Subtraction Using Background Sets With Image- and Color-Space Reduction <http://dx.doi.org/10.1109/TMM.2016.2595262> ... H. Lee, H. Kim, and J.-I. Kim 2093

Subjective and Objective Quality Assessment, and User Experience

- Learning Blind Quality Evaluator for Stereoscopic Images Using Joint Sparse Representation <http://dx.doi.org/10.1109/TMM.2016.2594142> F. Shao, K. Li, W. Lin, G. Jiang, and Q. Dai 2104

Social and Web Multimedia

- Filtering of Brand-Related Microblogs Using Social-Smooth Multiview Embedding <http://dx.doi.org/10.1109/TMM.2016.2581483> Y. Gao, Y. Zhen, H. Li, and T.-S. Chua 2115

COMMENTS AND CORRECTIONS

- Corrections to “Cross-Modal Correlation Learning by Adaptive Hierarchical Semantic Aggregation” <http://dx.doi.org/10.1109/TMM.2016.2593559> Y. Hua, S. Wang, S. Liu, A. Cai, and Q. Huang 2127

- Information for Authors <http://dx.doi.org/10.1109/TMM.2016.2608219> 2128

ANNOUNCEMENTS

- IEEE International Conference on Multimedia and Expo Call for Papers <http://dx.doi.org/10.1109/TMM.2016.2605539> 2130
- Introducing IEEE Collabratec <http://dx.doi.org/10.1109/TMM.2016.2608221> 2131
- IEEE Member Get-A-Member (MGM) Program <http://dx.doi.org/10.1109/TMM.2016.2608239> 2132

2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2017)

October 15–18, 2017



The 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2017) will be held at the Mohonk Mountain House in New Paltz, New York, and is supported by the Audio and Acoustic Signal Processing technical committee of the IEEE Signal Processing Society. The objective of this workshop is to provide an informal environment for the discussion of problems in audio, acoustics and signal processing techniques leading to novel solutions. Technical sessions will be scheduled throughout the day. Afternoons will be left free for informal meetings among workshop participants. Papers describing original research and new concepts are solicited for technical sessions on, but not limited to, the following topics:

Acoustic Signal Processing

- Source separation: single- and multi-microphone techniques
- Acoustic source localization and tracking
- Signal enhancement: dereverberation, noise reduction, echo reduction
- Microphone and loudspeaker array processing
- Acoustic sensor networks: distributed algorithms, synchronization
- Acoustic scene analysis: event detection and classification
- Room acoustics: analysis, modeling and simulation

Audio and Music Signal Processing

- Content-based music retrieval: fingerprinting, matching, cover song retrieval
- Musical signal analysis: segmentation, classification, transcription
- Music signal synthesis: waveforms, instrument models, singing
- Music separation: direct-ambient decomposition, vocal and instruments
- Audio effects: artificial reverberation, amplifier modeling
- Upmixing and downmixing

Audio and Speech Coding

- Waveform and parametric coding
- Spatial audio coding
- Sparse representations
- Low-delay audio and speech coding
- Digital rights

Hearing and Perception

- Hearing aids
- Computational auditory scene analysis
- Auditory perception and spatial hearing
- Speech and audio quality assessment
- Speech intelligibility measures and prediction

Important Dates

Submission of four-page paper
April 20, 2017

Notification of acceptance
June 27, 2017

Early registration until
August 15, 2017

Workshop
October 15–18, 2017

www.waspaa.com

Mohonk Mountain House
New Paltz, New York, USA

Workshop Committee

General Chairs

Patrick A. Naylor
Imperial College London

Meinard Müller
International AudioLabs Erlangen

Technical Program Chairs

Gautham Mysore
Adobe Research

Mads Christensen
Aalborg University

Finance Chair

Michael S. Brandstein
M.I.T. Lincoln Laboratory

Publications Chair

Toon van Waterschoot
KU Leuven

Registration Chair

Tiago H. Falk
INRS, Montréal

Industrial Liaison Chair

Tao Zhang
Starkey Hearing Technologies

Far East Liaison

Shoko Araki
NTT

Local Arrangements Chair

Youngjune Gwon
M.I.T. Lincoln Laboratory

Demonstrations Chair

Christine Evers
Imperial College London

Awards Chair

Sebastian Ewert
Queen Mary University of London



IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING



www.ieee.org/sp/index.html

OCTOBER 2016

VOLUME 10

NUMBER 7

IJSTGY

(ISSN 1932-4553)

ISSUE ON ADVANCED SIGNAL PROCESSING FOR BRAIN NETWORKS

EDITORIAL

- Introduction to the Issue on Advanced Signal Processing for Brain Networks <http://dx.doi.org/10.1109/IJSTSP.2016.2602945>
 V. Van De Ville, V. Jirsa, S. Strother, J. Richiardi, and A. Zalesky 1131

PAPERS

- Blind Source Separation for Unimodal and Multimodal Brain Networks: A Unifying Framework for Subspace Modeling
<http://dx.doi.org/10.1109/IJSTSP.2016.2594945> R. F. Silva, S. M. Plis, J. Sui, M. S. Pattichis, T. Adalı, and V. D. Calhoun 1134
- Bayesian Inference of Task-Based Functional Brain Connectivity Using Markov Chain Monte Carlo Methods
<http://dx.doi.org/10.1109/IJSTSP.2016.2599010> M. F. Ahmad, J. Murphy, D. Vatansever, E. A. Stamatakis, and S. J. Godsill 1150
- Archetypal Analysis for Modeling Multisubject fMRI Data <http://dx.doi.org/10.1109/IJSTSP.2016.2595103>
 J. L. Hinrich, S. E. Bardenfleth, R. E. Røge, N. W. Churchill, K. H. Madsen, and M. Mørup 1160
- A Combined Static and Dynamic Model for Resting-State Brain Connectivity Networks <http://dx.doi.org/10.1109/IJSTSP.2016.2594949>
 A. Liu, X. Chen, X. Dan, M. J. McKeown, and Z. J. Wang 1172
- Network-Based Statistic Show Aberrant Functional Connectivity in Alzheimer's Disease <http://dx.doi.org/10.1109/IJSTSP.2016.2600298>
 Y. Zhan, H. Yao, P. Wang, B. Zhou, Z. Zhang, Y. Guo, N. An, J. Ma, X. Zhang, and Y. Liu 1182
- Graph Frequency Analysis of Brain Signals <http://dx.doi.org/10.1109/IJSTSP.2016.2600859>
 W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro 1189
- Transmodal Learning of Functional Networks for Alzheimer's Disease Prediction <http://dx.doi.org/10.1109/IJSTSP.2016.2600400>
 M. Rahim, B. Thirion, C. Comtat, and G. Varoquaux 1204
- Localizing Sources of Brain Disease Progression with Network Diffusion Model <http://dx.doi.org/10.1109/IJSTSP.2016.2601695>
 C. Hu, X. Hua, J. Ying, P. M. Thompson, G. E. Fakhri, and Q. Li 1214
- Effective Connectivity Analysis in Brain Networks: A GPU-Accelerated Implementation of the Cox Method
<http://dx.doi.org/10.1109/IJSTSP.2016.2601820> V. Andalibi, F. Christophe, T. Laukkarinen, and T. Mikkonen 1226
- Hierarchical Coupled-Geometry Analysis for Neuronal Structure and Activity Pattern Discovery
<http://dx.doi.org/10.1109/IJSTSP.2016.2602061> G. Mishne, R. Talmon, R. Meir, J. Schiller, M. Lavzin, U. Dubin, and R. R. Coifman 1238
- MERLiN: Mixture Effect Recovery in Linear Networks <http://dx.doi.org/10.1109/IJSTSP.2016.2601144>
 S. Weichwald, M. Grosse-Wentrup, and A. Gretton 1254
- Identifying Seizure Onset Zone From the Causal Connectivity Inferred Using Directed Information
<http://dx.doi.org/10.1109/IJSTSP.2016.2601485> R. Malladi, G. Kalamangalam, N. Tandon, and B. Aazhang 1267
- Event-Related Functional Network Identification: Application to EEG Classification <http://dx.doi.org/10.1109/IJSTSP.2016.2602007>
 V. Gonuguntla, Y. Wang, and K. C. Veluvolu 1284
- Multilinear Discriminant Analysis With Subspace Constraints for Single-Trial Classification of Event-Related Potentials
<http://dx.doi.org/10.1109/IJSTSP.2016.2599297> H. Higashi, T. M. Rutkowski, T. Tanaka, and Y. Tanaka 1295
- An Informed Multitask Diffusion Adaptation Approach to Study Tremor in Parkinson's Disease
<http://dx.doi.org/10.1109/IJSTSP.2016.2578878> S. Monajemi, K. Eftaxias, S. Sanei, and S.-H. Ong 1306
- Modeling Effective Connectivity in High-Dimensional Cortical Source Signals <http://dx.doi.org/10.1109/IJSTSP.2016.2600023>
 Y. Wang, C.-M. Ting, and H. Ombao 1315



CALL FOR PAPERS

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (JSTSP)

Special Issue on Light Field Image Processing

As a representative approach for passive depth sensing, light field imaging and processing techniques has attracted great attentions from computer vision, computer graphics and signal processing communities. Specifically, light fields implicitly capture 3D scene geometry and reflectance properties by coding the spatial and angular information of a scene, allowing a wide range of applications. Compared with 2D image, this new visual information representation provides ways for better vision understanding and interacting capability for human. Traditional vision problems like image segmentation, scene parsing, salience detection, video stabilization, motion tracking, object detection, tracking and recognition, etc. can be better solved using the light field information. Moreover, it allows new applications including light field display, light field microscopy, etc. However, compared with commercial 3D depth sensing devices such as Microsoft Kinect using active 3D imaging approaches (structure light and time-of-flight imaging), light field imaging still suffers from crucial problems such as low spatial resolution and poor reconstructed depth quality. Therefore, the light field research is still calling for better theories and methods, targeting for more efficient data capture and better ways for data analysis. This special issue aims at such goal, and will deliver the timely and state-of-the-art research works on light field capture, processing, display and their applications. Topics of interest include, but are not limited to:

- Light Field Camera and Capture
- Light Field Synthesis
- Light Field Display
- Light Field Microscopy
- Light Field Data Compression and Transmission
- Light Field Editing and Application
- Light Field for VR/AR/MR
- Light Field Super-resolution
- Image Processing for Multiview Image
- Camera Array Technique and System
- Depth Estimation Using Light Field
- Computer Vision Using Light Field
- Time-of-Flight Depth Sensing
- Snapshot Depth Acquisition and Processing
- Plenoptic Imaging and Signal Processing
- Deep Learning Methods for Light Field Data

Tutorial or overview papers, creative papers outside the areas listed above but related to the overall scope of the special issue are also welcome. Prospective authors can contact the Guest Editors to ascertain interest on such topics. Submission of a paper to JSTSP is permitted only if the paper has not been submitted, accepted, published, or copyrighted in another journal. Papers that have been published in conference and workshop proceedings may be submitted for consideration to JSTSP provided that (i) the authors cite their earlier work; (ii) the papers are not identical; and (iii) the journal publication includes novel elements (*e.g.*, more comprehensive experiments). For submission information, please consult the IEEE JSTSP Information for Authors: <http://www.signalprocessingsociety.org/publications/periodicals/jstsp/>. Manuscripts should be submitted at <http://mc.manuscriptcentral.com/jstsp-ieee> and will be peer reviewed according to standard IEEE processes.

Important Dates

Initial Paper Submission:	December 1, 2016
1st Review Completed:	February 15, 2017
Revised Manuscript Due:	April 1, 2017
2nd Review Completed:	May 15, 2017
Final Manuscript Due:	June 15, 2017
Publication Date:	October, 2017

Guest Editors

Yebin Liu	Tsinghua University, China	liyuebin@tsinghua.edu.cn
Jingyi Yu	University of Delaware, USA	yu@eecis.udel.edu
Qing Wang	Northwestern Polytechnical University, China	qwang@nwpu.edu.cn
Lu Fang	Hong Kong University of Science and Technology	eefang@ust.hk
Diego Gutierrez	Universidad of Zaragoza, Spain	diegog@unizar.es
Feng Wu	University of Science and Technology of China, China	fengwu@ustc.edu.cn

IEEE Journal on Selected Topics in Signal Processing
IEEE Transactions on Signal and Information Processing over Networks
Special Issues on Graph Signal Processing

Numerous applications rely on the processing of high-dimensional data that resides on irregular or otherwise unordered structures which are naturally modeled as networks (such as social, economic, energy, transportation, telecommunication, sensor, and neural, to name a few). The need for new tools to process such data has led to the emergence of the field of graph signal processing, which merges algebraic and spectral graph theoretic concepts with computational harmonic analysis to process signals on structures such as graphs. This important new paradigm in signal processing research, coupled with its numerous applications in very different domains, has fueled the rapid development of an inter-disciplinary research community that has been working on theoretical aspects of graph signal processing and applications to diverse problems such as big data analysis, coding and compression of 3D point clouds, biological data processing, and brain network analysis.

The purpose of these special issues is to gather the latest advances in graph signal processing and disseminate new ideas and experiences in this emerging field to a broad audience. We encourage the submission of papers with new results, methods or applications in graph signal processing. In particular, the topics of interest include (but are not limited to):

- Sampling and recovery of graph signals
- Graph filter and filter bank design
- Uncertainty principles and other fundamental limits
- Graph signal transforms
- Graph topology inference
- Prediction and learning in graphs
- Statistical graph signal processing
- Non-linear graph signal processing
- Applications to visual information processing
- Applications to neuroscience and other medical fields
- Applications to economics and social networks
- Applications to various infrastructure networks

Submission procedure:

Prospective authors should follow the instructions given on the IEEE JSTSP webpages and submit their manuscript with the web submission system at <https://mc.manuscriptcentral.com/jstsp-ieee>. The decisions on whether the accepted papers will be published in IEEE JSTSP or IEEE TSIPN will depend on the respective themes of the papers and will be made by the Guest Editors.

Schedule (all deadlines are firm)

Manuscript due:	Nov 1, 2016
First Review Completed:	Jan 1, 2017
Revised manuscript due:	Mar 1, 2017
Second Review Completed:	May 1, 2017
Final manuscript due:	June 1, 2017
Publication date:	September 2017

Guest Editors:

Pier-Luigi Dragotti, Imperial College, London (p.dragotti@imperial.ac.uk)
Pascal Frossard, EPFL, Lausanne (pascal.frossard@epfl.ch)
Antonio Ortega, USC, Los Angeles (ortega@sipi.usc.edu)
Michael Rabbat, McGill University, Montreal (michael.rabbat@mcgill.ca)
Alejandro Ribeiro, UPenn, Philadelphia (aribeiro@seas.upenn.edu)

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

OCTOBER 2016

VOLUME 23

NUMBER 10

ISPLEM

(ISSN 1070-9908)

LETTERS

A Low-Complexity Transceiver Design in Sparse Multipath Massive MIMO Channels http://dx.doi.org/10.1109/LSP.2016.2582199	1301
..... Y. Yu, P. Wang, H. Chen, Y. Li, and B. Vucetic	
A Construction of Codebooks Nearly Achieving the Levenstein Bound http://dx.doi.org/10.1109/LSP.2016.2595106	1306
..... P. Tan, Z. Zhou, and D. Zhang	
Deep CNNs Along the Time Axis With Intermap Pooling for Robustness to Spectral Variations http://dx.doi.org/10.1109/LSP.2016.2589962	1310
..... H. Lee, G. Kim, H.-G. Kim, S.-H. Oh, and S.-Y. Lee	
Secure Transmissions via Compressive Sensing in Multicarrier Systems http://dx.doi.org/10.1109/LSP.2016.2595524	1315
..... J. Choi	
Robust Localization Using Time Difference of Arrivals http://dx.doi.org/10.1109/LSP.2016.2569666	1320
..... X. Shi, B. D. O. Anderson, G. Mao, Z. Yang, J. Chen, and Z. Lin	
Exploiting Attribute Dependency for Attribute Assignment in Crowded Scenes http://dx.doi.org/10.1109/LSP.2016.2592689	1325
..... C. Deng, Z. Cao, Y. Xiao, H. Lu, K. Xian, and Y. Chen	
Visual Indoor Localization in Known Environments http://dx.doi.org/10.1109/LSP.2016.2593958	1330
..... Claudio Picciarelli	
Dual Microphone Voice Activity Detection Exploiting Interchannel Time and Level Differences http://dx.doi.org/10.1109/LSP.2016.2597360	1335
..... J. Park, Y. G. Jin, S. Hwang, and J. W. Shin	
A Note on Alternating Minimization Algorithm for the Matrix Completion Problem http://dx.doi.org/10.1109/LSP.2016.2576979	1340
..... D. Gamarnik and S. Misra	
Microphone Array Processing Strategies for Distant-Based Automatic Speech Recognition http://dx.doi.org/10.1109/LSP.2016.2592683 ...	1344
..... S. A. Khoubrouy and J. H. L. Hansen	
A BP-MF-EP Based Iterative Receiver for Joint Phase Noise Estimation, Equalization, and Decoding http://dx.doi.org/10.1109/LSP.2016.2593917	1349
..... W. Wang, Z. Wang, C. Zhang, Q. Guo, P. Sun, and X. Wang	
Impact Analysis of Baseband Quantizer on Coding Efficiency for HDR Video http://dx.doi.org/10.1109/LSP.2016.2597175	1354
..... C.-W. Wong, G.-M. Su, and M. Wu	
High-Resolution Radar Detection in Interference and Nonhomogeneous Noise http://dx.doi.org/10.1109/LSP.2016.2597738	1359
..... Y. Gao, G. Liao, and W. Liu	
Optimal Energy Harvesting-Ratio and Beamwidth Selection in Millimeter Wave Communications http://dx.doi.org/10.1109/LSP.2016.2598151	1364
..... Y. Wu, Q. Yang, Q. He, and K. S. Kwak	
Modeling and Classifying Tip Dynamics of Growing Cells in Video http://dx.doi.org/10.1109/LSP.2016.2598558	1369
..... A. L. Tambo and B. Bhanu	
Discriminative Action States Discovery for Online Action Recognition http://dx.doi.org/10.1109/LSP.2016.2598878	1374
..... B. Hu, J. Yuan, and Y. Wu	

Perceptually Weighted Analysis-by-Synthesis Vector Quantization for Low Bit Rate MFCC Codec http://dx.doi.org/10.1109/LSP.2016.2598226	<i>G. Min, X. Zhang, X. Zou, and J. Yang</i>	1379
Blind Deconvolution From Multiple Sparse Inputs http://dx.doi.org/10.1109/LSP.2016.2599104	<i>L. Wang and Y. Chi</i>	1384
Adaptive Multiscale Decomposition of Graph Signals http://dx.doi.org/10.1109/LSP.2016.2598750	<i>X. Zheng, Y. Y. Tang, J. Pan, and J. Zhou</i>	1389
Gradient Direction for Screen Content Image Quality Assessment http://dx.doi.org/10.1109/LSP.2016.2599294	<i>Z. Ni, L. Ma, H. Zeng, C. Cai, and K.-K. Ma</i>	1394
Localized Low-Rank Promoting for Recovery of Block-Sparse Signals With IntraBlock Correlation http://dx.doi.org/10.1109/LSP.2016.2599255	<i>L. Yang, J. Fang, H. Li, and B. Zeng</i>	1399
NeRD: A Neural Response Divergence Approach to Visual Saliency Detection http://dx.doi.org/10.1109/LSP.2016.2600592	<i>M. J. Shafiee, P. Siva, C. Scharfenberger, P. Fieguth, and A. Wong</i>	1404
A New Algorithm for Optimizing TV-Based PolSAR Despeckling Model http://dx.doi.org/10.1109/LSP.2016.2602299	<i>X. Nie, B. Zhang, Y. Chen, and H. Qiao</i>	1409
UNIQUE: Unsupervised Image Quality Estimation http://dx.doi.org/10.1109/LSP.2016.2601119	<i>D. Temel, M. Prabhushankar, and G. AlRegib</i>	1414
Robust Rank-Two Multicast Beamforming Under a Unified CSI Uncertainty Model http://dx.doi.org/10.1109/LSP.2016.2600618	<i>B. Liu, L. Shi, and X.-G. Xia</i>	1419
Why Does a Hilbertian Metric Work Efficiently in Online Learning With Kernels? http://dx.doi.org/10.1109/LSP.2016.2598615	<i>M. Yukawa and K.-R. Müller</i>	1424
Universal Share for Multisecret Image Sharing Scheme Based on Boolean Operation	<i>Y. K. Meghrajani and H. S. Mazumdar</i>	1429
Adaptive Gain Control for Enhanced Speech Intelligibility Under Reverberation http://dx.doi.org/10.1109/LSP.2016.2598807	<i>P. N. Petkov and Y. Stylianou</i>	1434
Space-Frequency Coded Index Modulation With Linear-Complexity Maximum Likelihood Receiver in the MIMO-OFDM System http://dx.doi.org/10.1109/LSP.2016.2599020	<i>L. Wang, Z. Chen, Z. Gong, and M. Wu</i>	1439
Time-Asynchronous Robust Cooperative Transmission for the Downlink of C-RAN http://dx.doi.org/10.1109/LSP.2016.2601113	<i>S.-H. Park, O. Simeone, and S. Shamai</i>	1444
Flexible Pilot Contamination Mitigation With Doppler PSD Alignment http://dx.doi.org/10.1109/LSP.2016.2601270	<i>X. Luo and X. Zhang</i>	1449
When Correlation Filters Meet Convolutional Neural Networks for Visual Tracking http://dx.doi.org/10.1109/LSP.2016.2601691	<i>C. Ma, Y. Xu, B. Ni, and X. Yang</i>	1454
Sound Source and Microphone Localization From Acoustic Impulse Responses http://dx.doi.org/10.1109/LSP.2016.2601878	<i>D. Salvati, C. Drioli, and G. L. Foresti</i>	1459
A Low-Complexity Multiuser Adaptive Modulation Scheme for Massive MIMO Systems http://dx.doi.org/10.1109/LSP.2016.2601932	<i>Y. Zhou, C. Zhong, S. Jin, Y. Huang, and Z. Zhang</i>	1464
Multisnapshot Sparse Bayesian Learning for DOA http://dx.doi.org/10.1109/LSP.2016.2598550	<i>P. Gerstoft, C. F. Mecklenbräuker, A. Xenaki, and S. Nannuru</i>	1469
Heretical Multiple Importance Sampling http://dx.doi.org/10.1109/LSP.2016.2600678	<i>V. Elvira, L. Martino, D. Luengo, and M. F. Bugallo</i>	1474
Supervised Hashing Based on the Dimensions' Value Cardinalities of Image Descriptors http://dx.doi.org/10.1109/LSP.2016.2599901	<i>D. Rafailidis</i>	1479
Universal Collaboration Strategies for Signal Detection: A Sparse Learning Approach http://dx.doi.org/10.1109/LSP.2016.2601911	<i>P. Khanduri, B. Kailkhura, J. J. Thiagarajan, and P. K. Varshney</i>	1484
Knowledge-Based Adaptive Detection: Joint Exploitation of Clutter and System Symmetry Properties http://dx.doi.org/10.1109/LSP.2016.2601931	<i>C. Hao, D. Orlando, G. Foglia, and G. Giunta</i>	1489
Unbiased Injection of Signal-Dependent Noise in Variance-Stabilized Range http://dx.doi.org/10.1109/LSP.2016.2601689	<i>L. R. Borges, M. A. da Costa Vieira, and A. Foi</i>	1494
Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks http://dx.doi.org/10.1109/LSP.2016.2603342	<i>K. Zhang, Z. Zhang, Z. Li, and Y. Qiao</i>	1499

General Chairs

Zhi Tian

George Mason Univ.

Brian M. Sadler

*Army Research Lab.***Technical Program Chairs**

Philip Regalia

Catholic Univ. of America

Trac D. Tran

Johns Hopkins Univ.

Brian Mark

*George Mason Univ.***Finance Chair**

Jill Nelson

*George Mason Univ.***Local Arrangement Chair**

Nathalia Peixoto

*George Mason Univ.***Publications Chair**

Kathleen Wage

*George Mason Univ.***Publicity Chairs**

Piya Pal

UC San Diego

Seung-Jun Kim

*Univ. MD, Baltimore Cty***Technical Workshop****Liaison Chair**

Min Wu

*Univ. MD, College Park***Government Panel Chair**

Joel Goodman

*Naval Research Lab***Industrial Liaison****Chairs**

Kristine Bell

Metron Inc.

Hang Liu

*Catholic Univ.***International Liaison****Chairs**

Chengyang Yang

BUAA, China

Mounir Ghogho

*Univ. of Leeds, UK***Advisory Committee**

Monson Hayes

George Mason Univ.

IEEE GlobalSIP



Fourth IEEE Global Conference on Signal and Information Processing

December 7–9, 2016, Greater Washington D.C., USA

IEEE
Signal Processing Society<http://2016.ieeeglobalsip.org/>

Call for Participation

The fourth IEEE Global Conference on Signal and Information Processing (GlobalSIP) will be held in Greater Washington, DC, USA on December 7–9, 2016. The conference features world-class plenary and keynote talks, tutorials and government panel discussions, as well as co-located timely symposia selected based on the responses to the Call for Symposium Proposals.

Plenary Speakers

Ben Vigoda (Gamalon Machine Intelligence)**Danielle Bassett** (University of Pennsylvania)**Stéphane Mallat** (Ecole Normale Supérieure)

Government Panel

Brian Sadler (ARL), Moderator**David Aha** (NRL)**Charles Clancy** (Virginia Tech)**Jill Crisman** (IARPA)**Tom Rondeau** (DARPA)**Paul Tilghman** (DARPA)

List of Symposia

- General symposium
- Compressed sensing and deep learning
- Signal and information processing over networks
- Distributed information processing, optimization, and resource management
- Transceivers and signal processing for 5G wireless and mm-wave systems
- Information theoretic approaches to security and privacy
- Signal processing of big data
- Cognitive communications and Radar
- Big data analytics and challenges in medical imaging
- Signal processing for understanding crowd dynamics
- Signal and information processing for smart grid infrastructure
- Non-commutative theory and applications
- Sparse signal processing for communications
- Emerging signal processing applications

Conference highlights

- 14 technical symposia with 24 keynotes by leading experts overviewing emerging SIP topics
- Government panel discussions on funding opportunities and trends
- New industrial symposium on emerging SP applications with demos and exhibitions
- Great venue with vibrant cultural, educational, and scientific identity, housing museums (many are free), monuments, art centers, universities, and federal agencies
- Opportunity to attend both GlobalSIP and Globecom (Dec 4-8, 2016) in one trip

Advance Registration Deadline:**Oct. 28, 2016****Hotel Reservation Cut-Off at Conference Rate:****Nov. 15, 2016**

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING

A PUBLICATION OF
IEEE SIGNAL PROCESSING SOCIETY
IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY
IEEE CONSUMER ELECTRONICS SOCIETY



TECHNICALLY CO-SPONSORED BY
IEEE GEOSCIENCE AND REMOTE SENSING SOCIETY



SEPTEMBER 2016

VOLUME 2

NUMBER 3

ITCIAJ

(ISSN 2333-9403)

PAPERS

Computational Imaging Methods and Models

- Learning Efficient Data Representations With Orthogonal Sparse Coding <http://dx.doi.org/10.1109/TCL.2016.2557065> *H. Schütze, E. Barth, and T. Martinetz* 177
- An Algorithm Architecture Co-Design for CMOS Compressive High Dynamic Range Imaging <http://dx.doi.org/10.1109/TCL.2016.2557073> *W. Guicquero, A. Dupret, and P. Vanderghenst* 190
- Color-Line Regularization for Color Artifact Removal <http://dx.doi.org/10.1109/TCL.2016.2575740> *S. Ono and I. Yamada* 204
- Multi-Resolution Compressed Sensing Reconstruction Via Approximate Message Passing <http://dx.doi.org/10.1109/TCL.2016.2575741> *X. Wang and J. Liang* 218
- An Augmented Lagrangian Method for Complex-Valued Compressed SAR Imaging <http://dx.doi.org/10.1109/TCL.2016.2580498> *H. E. Güven, A. Güngör, and M. Çetin* 235

Computational Image Formation

- Toward Long-Distance Subdiffraction Imaging Using Coherent Camera Arrays <http://dx.doi.org/10.1109/TCL.2016.2557067> *J. Holloway, M. S. Asif, M. K. Sharma, N. Matsuda, R. Horstmeyer, O. Cossairt, and A. Veeraraghavan* 251
- Sensitivity Encoding for Aligned Multishot Magnetic Resonance Reconstruction <http://dx.doi.org/10.1109/TCL.2016.2557069> *L. Cordero-Grande, R. P. A. G. Teixeira, E. J. Hughes, J. Hutter, A. N. Price, and J. V. Hajnal* 266
- Combining Inertial Measurements With Blind Image Deblurring Using Distance Transform <http://dx.doi.org/10.1109/TCL.2016.2561701> .. *Y. Zhang and K. Hirakawa* 281



Data-Driven Learning of a Union of Sparsifying Transforms Model for Blind Compressed Sensing http://dx.doi.org/10.1109/TCL.2016.2567299	<i>S. Ravishankar and Y. Bresler</i>	294
Nonlinear Optimization Algorithm for Partially Coherent Phase Retrieval and Source Recovery http://dx.doi.org/10.1109/TCL.2016.2571669	<i>J. Zhong, L. Tian, P. Varma, and L. Waller</i>	310
<i>Computational Imaging Systems</i>		
A Graph Partitioning Approach to Simultaneous Angular Reconstitution http://dx.doi.org/10.1109/TCL.2016.2557076	<i>G. Pragier, I. Greenberg, X. Cheng, and Y. Shkolnisky</i>	323
Non-Binary Discrete Tomography by Continuous Non-Convex Optimization http://dx.doi.org/10.1109/TCL.2016.2563321	<i>M. Zisler, J. H. Kappes, C. Schnörr, S. Petra, and C. Schnörr</i>	335
Simultaneous Temporal Superresolution and Denoising for Cardiac Fluorescence Microscopy http://dx.doi.org/10.1109/TCL.2016.2579606 ..	<i>K. G. Chan, S. J. Streichan, L. A. Trinh, and M. Liebling</i>	348
A Gaussian Mixture MRF for Model-Based Iterative Reconstruction With Applications to Low-Dose X-Ray CT http://dx.doi.org/10.1109/TCL.2016.2582042	<i>R. Zhang, D. H. Ye, D. Pal, J.-B. Thibault, K. D. Sauer, and C. A. Bouman</i>	359
Optimal Gradient Encoding Schemes for Diffusion Tensor and Kurtosis Imaging http://dx.doi.org/10.1109/TCL.2016.2590301	<i>M. Alipoor, I. Y.-H. Gu, S. E. Maier, G. Starck, A. Mehnert, and F. Kahl</i>	375

EDICS—Editor's Classification Information Scheme http://dx.doi.org/10.1109/TCL.2016.2593963		392
Information for Authors http://dx.doi.org/10.1109/TCL.2016.2593964		393

IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS

A PUBLICATION OF
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY



SEPTEMBER 2016

VOLUME 2

NUMBER 3

ITSIBW

(ISSN 2373-776X)

PAPERS

Adaptation, Detection, Estimation, and Learning

Spectral Graph Wavelets and Filter Banks With Low Approximation Error http://dx.doi.org/10.1109/TSIPN.2016.2581303	230
..... A. Sakiyama, K. Watanabe, and Y. Tanaka	
Theoretical Analysis of the Measurement Transportation Algorithm to Fuse Delayed Data in Distributed Sensor Networks http://dx.doi.org/10.1109/TSIPN.2016.2580461	246
..... R. A. J. Chagas and J. Waldmann	
Distributed Bernoulli Filters for Joint Detection and Tracking in Sensor Networks http://dx.doi.org/10.1109/TSIPN.2016.2580460	260
..... S. S. Dias and M. G. S. Bruno	
Diffusion Estimation Over Cooperative Multi-Agent Networks With Missing Data http://dx.doi.org/10.1109/TSIPN.2016.2570679	276
..... M. R. Gholami, M. Jansson, E. G. Ström, and A. H. Sayed	
An Online Parallel Algorithm for Recursive Estimation of Sparse Signals http://dx.doi.org/10.1109/TSIPN.2016.2561703	290
..... Y. Yang, M. Pesavento, M. Zhang, and D. P. Palomar	
The Importance of Being Earnest: Social Sensing With Unknown Agent Quality http://dx.doi.org/10.1109/TSIPN.2016.2555799	306
..... S. Marano, V. Matta, and P. Willett	
Distributed Two-Step Quantized Fusion Rules Via Consensus Algorithm for Distributed Detection in Wireless Sensor Networks http://dx.doi.org/10.1109/TSIPN.2016.2549743	321
..... E. Nurellari, D. McLernon, and M. Ghogho	
<i>Communication, Networking, and Sensing</i>	
Consensus in the Presence of Multiple Opinion Leaders: Effect of Bounded Confidence http://dx.doi.org/10.1109/TSIPN.2016.2571839	336
..... R. Dabarera, K. Premaratne, M. N. Murthi, and D. Sarkar	
Overhearing-Based Co-Operation for Two-Cell Network With Asymmetric Uplink-Downlink Traffics http://dx.doi.org/10.1109/TSIPN.2016.2549179	350
..... C. Li, J. Wang, F.-C. Zheng, J. M. Cioffi, and L. Yang	
Distributed Long-Term Base Station Clustering in Cellular Networks Using Coalition Formation http://dx.doi.org/10.1109/TSIPN.2016.2548781	362
..... R. Brandt, R. Mochaourab, and M. Bengtsson	
<i>Imaging and Media Applications</i>	
Efficient Wireless Multimedia Multicast in Multi-Rate Multi-Channel Mesh Networks http://dx.doi.org/10.1109/TSIPN.2016.2550806	376
..... W. Tu	



Modeling and Analysis

Detecting Convoys Using License Plate Recognition Data http://dx.doi.org/10.1109/TSIPN.2016.2569426	391
..... <i>S. Lawlor, T. Sider, N. Eluru, M. Hatzopoulou, and M. G. Rabbat</i>	
Active Sensing of Social Networks http://dx.doi.org/10.1109/TSIPN.2016.2555785	406
..... <i>H.-T. Wai, A. Scaglione, and A. Leshem</i>	

EDICS—Editor's Information Classification Scheme http://dx.doi.org/10.1109/TSIPN.2016.2593885	420
Information for Authors http://dx.doi.org/10.1109/TSIPN.2016.2593889	421

General Chairs

Jörn Ostermann, LUH, Germany
Kenneth K.M. Lam, PolyU, Hong Kong

Technical Program Chairs

Thomas Sikora (Coordinator), TU Berlin, Germany
Qian Zhang, HKUST, Hong Kong
Prabha Balakrishnan, U. Texas at Dallas, USA
Winston Hsu, NTU, Taiwan
Yonggang Wen, NTU, Singapore
Ce Zhu, UESTC, China

Finance Chairs

Daniel P.K. Lun, PolyU, Hong Kong
Chris Y.H. Chan, PolyU, Hong Kong

Workshop Chairs

Béatrice Pesquet-Popescu, ParisTech, France
Chong-Wah Ngo, CityU, Hong Kong

Special Sessions Chair

Gene Cheung, National Institute of Informatics, Japan

Zhu Li, University of Missouri – Kansas City, USA

Panel Session Chair

Lap Pui Chau, NTU, Singapore

Tutorial Chair

Atanas Gotchev, TUT, Finland

Grand Challenge Chair

Axel Becker-Lakus, Five Open Books LLC, USA

Student Program Chair

Hong Zhao, Tsinghua University, China

Award Chair

Moncef Gabbouj, Tampere University of Technology, Finland

Sponsorship Chair

Xian-Sheng Hua, Alibaba, China

Registration Chairs

Tan Lee, CUHK, Hong Kong

Publications Chairs

Yui Lam Chan, PolyU, Hong Kong

Publicity Chair

Thomas Fang Zheng, Tsinghua University, China
Ling Guan, Ryerson University, Canada

Web Chair

Thorsten Laude, LUH, Germany

Local Arrangement Chairs

Frank Leung, PolyU, Hong Kong

Secretaries

Bonnie Law, PolyU, Hong Kong
Jiantao Zhou, U. Macau, Macau

Europe Liaison

N.N.

North America Liaison

N.N.

South America Liaison

N.N.

Australia/Liaison

N.N.

Asia Liaison

Xiaopeng Fan, HIT, China
Mark Liao, Academia Sinica, Taiwan

Advisor

Wen Gao, Peking University, China
Wan Chi Siu, PolyU, Hong Kong

**IEEE International Conference on Multimedia and Expo (ICME) 2017****CALL FOR PAPERS****IEEE International Conference on Multimedia and Expo (ICME) 2017****July 10 – 14, 2017 • Hong Kong**

IEEE International Conference on Multimedia & Expo (ICME) has been the flagship multimedia conference sponsored by four IEEE societies since 2000. It serves as a forum to promote the exchange of the latest advances in multimedia technologies, systems, and applications from both the research and development perspectives of the circuits and systems, communications, computer, and signal processing communities. An Exposition of multimedia products, animations and industries will be held in conjunction with the conference. The theme of the 2017 conference will be “The New Media Experience”, enabling next generation 3D/AR/VR experiences and applications, based on which various sessions and events will be organized.

Authors are invited to submit a full paper (two-column format, 6 pages maximum) according to the guidelines available on the conference website at <http://www.icme2017.org/>. Reviewing will be double blind. Only electronic submissions will be accepted. Topics of interest include, but are not limited to:

- Speech, audio, image, video, text and new sensor signal processing
- Signal processing for media integration
- Interactive 3D media and immersive VR/AR environments
- Multi-modal multimedia computing systems and human-machine interaction
- Multimedia Cloud and Big Data
- Multimedia IoT
- Mobile Multimedia Systems and Applications
- Multimedia compression
- Multimedia security, privacy, and content protection
- Multimedia databases and digital libraries
- Multimedia applications and services
- Media content analysis, analytics, and search
- Hardware and software for multimedia systems
- Multimedia standards and related issues
- Multimedia quality assessment

ICME 2017 will showcase high quality oral and poster presentations and demonstration sessions. Best paper awards (one diamond and five platinum awards) will be selected and recognized in the conference. Several awards sponsored by industry and institutions will be given out. Accepted papers have to be registered and presented; otherwise they will not be included in the IEEE Xplore Library.

ICME2017 will feature IEEE societies sponsored Workshops. We welcome researchers, developers and practitioners to organize Workshops on various new emerging topics. Industrial exhibitions are held in conjunction with the main conference. Proposals for Tutorials, Special Sessions, and Demos are also invited. Please visit the ICME 2017 website for submission details.

Regular Paper Abstract Submission:**Regular Paper Submission:****Notification of Regular Paper Acceptance:****Special Session/Workshop Proposals Due:****Notification of Special Session/Workshop Proposal Acceptance:****Tutorial Proposals Due:****Workshop Paper Submission:****Notification of Workshop Paper Acceptance:****Camera-Ready Paper Due:****November 25, 2016****December 2, 2016****February 24, 2017****November 1, 2016****November 7, 2016****January 23, 2017****February 27, 2017****March 27, 2017****April 10, 2017**

www.icme2017.org



Contact Email:
icme@tnt.uni-hannover.de

SPAWC 2017

The 18th IEEE International Workshop on Signal Processing Advances in Wireless Communications, July 3 – 6, Sapporo, Japan

General Chairs

Yasutaka Ogawa
Wei Yu
FumiYuki Adachi

Technical Program Chairs

Tomoaki Ohtsuki
Lutz Lampe
Wing-Kin (Ken) Ma

Special Session Chair

Tony Q. S. Quek

Technical Program Committee

Waheed Bajwa
Chong-Yung Chi
Philippe Ciblat
Min Dong
FeiFei Gao
Joakim Jaldén
Shi Jin
Yindi Jing
Markku Juntti
Eleftherios Karipidis
Erik G. Larsson
David J. Love
Matthew McKay
Urbashi Mitra
Chandra Murthy
Michael Rabbat
Alejandro Ribeiro
Ahmed Sadek
Mathini Sellathurai
Milica Stojanovic
Weifeng Su
Cihan Tepedelenlioglu
Wolfgang Utschick
Sergiy A. Vorobyov
Pengfei Xia
Rui Zhang
Wei Zhang

Finance Chairs

Yasushi Yamao
Hai Lin

Local Arrangement Chairs

Takeo Ohgane
Toshihiko Nishimura

SPAWC 2017 will be held at Hokkaido University in Sapporo, Japan on July 3-6, 2017. The workshop is devoted to advances in signal processing for wireless communications, networking, and information theory. The technical program features plenary talks as well as invited and contributed papers presented in poster format.

Call for Papers

Prospective authors are invited to submit papers in the following areas:

- Smart antennas, MIMO systems, massive MIMO, and space-time processing
- Single-carrier, multi-carrier, and multi-rate systems
- Multiple-access and broadcast channels, multi-user receivers
- Signal processing for ad-hoc, multi-hop, and sensor networks
- Cooperative communication, coordinated multipoint transmission and reception
- Distributed resource allocation and scheduling
- Convex and non-convex optimization; Game theory for communications
- Interference management, dynamic spectrum management
- Heterogeneous networks, small cells
- Millimeter wave, 60GHz communications
- Full duplex systems
- Physical layer security
- Feedback in wireless networks
- Cognitive radio and networks
- Cooperative sensing, compressed sensing, sparse signal processing
- Machine-to-machine, device-to-device communications
- Ultra-wideband radio, localization, RFID
- Modeling, estimation and equalization of wireless channels
- Acquisition, synchronization, and tracking
- Cross-layer issues, joint source-channel coding, delay-limited communication
- Signal processing for optical, satellite, and underwater communications
- Signal processing for nano- and molecular communications
- Energy efficiency and energy harvesting
- Large networks and big data in wireless communications
- Emerging techniques, technologies, and new waveforms for 5G

Full papers up to five-page limit should be submitted via EDAS.

March 1, 2017
April 30, 2017
May 10, 2017

Paper submission deadline
Acceptance notification
Final paper due

Publication Chairs

Kazunori Hayashi
Kenichi Higuchi



Publicity Chairs

Julian Webber
Koichi Adachi

IEEE SignalProcessing

Volume 33 | Number 5 | September 2016

MAGAZINE

COMPUTATIONAL PHOTOGRAPHY AND DISPLAYS

Interfacing
with the
Visual World

Recent Advances
in Phase Retrieval

The MPEG Internet
Video-Coding Standard

Highlights from the IEEE SP Cup
2016 Student Competition

IEEE
Signal Processing Society

IEEE

Contents

Volume 33 | Number 5 | September 2016

FEATURES

- 12 FROM THE GUEST EDITORS**
Amit Agrawal, Richard Baraniuk, Paolo Favaro, and Ashok Veeraraghavan
- 16 A SURVEY OF COMPUTATIONAL PHOTOGRAPHY IN THE SMALL**
Sanjeev J. Koppal
- 23 LENSLESS IMAGING**
Vivek Boominathan, Jesse K. Adams, M. Salman Asif, Benjamin W. Avants, Jacob T. Robinson, Richard G. Baraniuk, Aswin C. Sankaranarayanan, and Ashok Veeraraghavan
- 36 PRACTICAL HIGH DYNAMIC RANGE IMAGING OF EVERYDAY SCENES**
Pradeep Sen and Cecilia Aguerreberre
- 45 SIGNAL PROCESSING FOR TIME-OF-FLIGHT IMAGING SENSORS**
Ayush Bhandari and Ramesh Raskar



See the "President's Message" and "Society News" in this issue!



ON THE COVER

IEEE *Signal Processing Magazine* is uniquely positioned to convey and embrace an evolving scope of signal processing. This issue showcases an example of a cross-disciplinary area—fascinating advances of computational photography and display. A variety of articles showcase the potential for the field to revolutionize imaging and displays and transform the way in which we capture, share, and interact with the visual world around us.

COVER IMAGE: ©ISTOCKPHOTO.COM/SCYTHERS

- 59 PRINCIPLES OF LIGHT FIELD IMAGING**
Ivo Ihrke, John Restrepo, and Loïs Mignard-Debise
- 70 CAPTURING COMPUTATIONAL APPEARANCE**
Kristin J. Dana
- 81 ENHANCED COMPRESSIVE IMAGING USING MODEL-BASED ACQUISITION**
Aswin C. Sankaranarayanan, Pavan Turaga, Matthew A. Herman, and Kevin F. Kelly

- 95 COMPUTATIONAL SNAPSHOT MULTISPECTRAL CAMERAS**
Xun Cao, Tao Yue, Xing Lin, Stephen Lin, Xin Yuan, Qionghai Dai, Lawrence Carin, and David J. Brady

- 109 COMPUTATIONAL MILLIMETER WAVE IMAGING**
Vishal M. Patel, Joseph N. Mait, Dennis W. Prather, and Abigail S. Hedden

- 119 FACTORED DISPLAYS**
Gordon Wetzstein and Douglas Lanman

- 130 COMPUTATIONAL IMAGING FOR CULTURAL HERITAGE**
Xiang Huang, Erich Uffelmann, Oliver Cossairt, Marc Walton, and Aggelos K. Katsaggelos

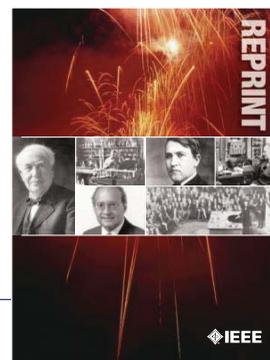
- 139 GAZE-CONTINGENT COMPUTATIONAL DISPLAYS**
Michael Stengel and Marcus Magnor



PG. 149

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 © 2016 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.2016.2590618



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.

Authorized Signature: _____ Date: _____

2016 IEEE MEMBERSHIP APPLICATION

(students and graduate students must apply online)

Start your membership immediately: Join online www.ieee.org/join

Please complete both sides of this form, typing or **printing in capital letters**. Use only English characters and abbreviate only if more than 40 characters and spaces per line. We regret that incomplete applications cannot be processed.

1 Name & Contact Information

Please PRINT your name as you want it to appear on your membership card and IEEE correspondence. As a key identifier for the IEEE database, circle your last/surname.

Male Female Date of birth (Day/Month/Year) ____/____/____

Title First/Given Name Middle Last/Family Surname

▼ **Primary Address** Home Business (All IEEE mail sent here)

Street Address

City State/Province

Postal Code Country

Primary Phone

Primary E-mail

▼ **Secondary Address** Home Business

Company Name Department/Division

Street Address City State/Province

Postal Code Country

Secondary Phone

Secondary E-mail

To better serve our members and supplement member dues, your postal mailing address is made available to carefully selected organizations to provide you with information on technical services, continuing education, and conferences. Your e-mail address is **not** rented by IEEE. Please check box **only** if you do not want to receive these postal mailings to the selected address.

2 Attestation

I have graduated from a three- to five-year academic program with a university-level degree.

Yes No

This program is in one of the following fields of study:

- Engineering
- Computer Sciences and Information Technologies
- Physical Sciences
- Biological and Medical Sciences
- Mathematics
- Technical Communications, Education, Management, Law and Policy
- Other (please specify): _____

This academic institution or program is accredited in the country where the institution is located. Yes No Do not know

I have _____ years of professional experience in teaching, creating, developing, practicing, or managing within the following field:

- Engineering
- Computer Sciences and Information Technologies
- Physical Sciences
- Biological and Medical Sciences
- Mathematics
- Technical Communications, Education, Management, Law and Policy
- Other (please specify): _____



3 Please Tell Us About Yourself

Select the numbered option that best describes yourself. This information is used by IEEE magazines to verify their annual circulation. Please enter numbered selections in the boxes provided.

A. Primary line of business

1. Computers
2. Computer peripheral equipment
3. Software
4. Office and business machines
5. Test, measurement and instrumentation equipment
6. Communications systems and equipment
7. Navigation and guidance systems and equipment
8. Consumer electronics/appliances
9. Industrial equipment, controls and systems
10. ICs and microprocessors
11. Semiconductors, components, sub-assemblies, materials and supplies
12. Aircraft, missiles, space and ground support equipment
13. Oceanography and support equipment
14. Medical electronic equipment
15. OEM incorporating electronics in their end product (not elsewhere classified)
16. Independent and university research, test and design laboratories and consultants (not connected with a mfg. co.)
17. Government agencies and armed forces
18. Companies using and/or incorporating any electronic products in their manufacturing, processing, research or development activities
19. Telecommunications services, telephone (including cellular)
20. Broadcast services (TV, cable, radio)
21. Transportation services (airline, railroad, etc.)
22. Computer and communications and data processing services
23. Power production, generation, transmission and distribution
24. Other commercial users of electrical, electronic equipment and services (not elsewhere classified)
25. Distributor (reseller, wholesaler, retailer)
26. University, college/other educational institutions, libraries
27. Retired
28. Other _____

B. Principal job function

- | | |
|--|--|
| 1. General and corporate management | 9. Design/development engineering—digital |
| 2. Engineering management | 10. Hardware engineering |
| 3. Project engineering management | 11. Software design/development management |
| 4. Research and development management | 12. Computer science |
| 5. Design engineering management—analogue | 13. Science/physics/mathematics |
| 6. Design engineering management—digital | 14. Engineering (not elsewhere specified) |
| 7. Research and development engineering | 15. Marketing/sales/purchasing |
| 8. Design/development engineering—analogue | 16. Consulting |
| | 17. Education/teaching |
| | 18. Retired |
| | 19. Other _____ |

C. Principal responsibility

- | | |
|--|-----------------------|
| 1. Engineering and scientific management | 6. Education/teaching |
| 2. Management other than engineering | 7. Consulting |
| 3. Engineering design | 8. Retired |
| 4. Engineering | 9. Other _____ |
| 5. Software: science/mngmnt/engineering | |

D. Title

- | | |
|--|--------------------------------|
| 1. Chairman of the Board/President/CEO | 10. Design Engineering Manager |
| 2. Owner/Partner | 11. Design Engineer |
| 3. General Manager | 12. Hardware Engineer |
| 4. VP Operations | 13. Software Engineer |
| 5. VP Engineering/Dir. Engineering | 14. Computer Scientist |
| 6. Chief Engineer/Chief Scientist | 15. Dean/Professor/Instructor |
| 7. Engineering Management | 16. Consultant |
| 8. Scientific Management | 17. Retired |
| 9. Member of Technical Staff | 18. Other _____ |

Are you now or were you ever a member of IEEE?
 Yes No If yes, provide, if known:
 Membership Number _____ Grade _____ Year Expired _____

4 Please Sign Your Application

I hereby apply for IEEE membership and agree to be governed by the IEEE Constitution, Bylaws, and Code of Ethics. I understand that IEEE will communicate with me regarding my individual membership and all related benefits. **Application must be signed.**

Signature _____ Date _____ Over Please

(continued on next page)



5 Add IEEE Society Memberships (Optional)

The 39 IEEE Societies support your technical and professional interests. Many society memberships include a personal subscription to the core journal, magazine, or newsletter of that society. **For a complete list of everything included with your IEEE Society membership, visit www.ieee.org/join.** All prices are quoted in US dollars.

Please check the appropriate box.

		BETWEEN 16 AUG 2015- 28 FEB 2016 PAY	BETWEEN 1 MAR 2016- 15 AUG 2016 PAY
IEEE Aerospace and Electronic Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	AES010	25.00 <input type="checkbox"/>	12.50 <input type="checkbox"/>
IEEE Antennas and Propagation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	AP003	15.00 <input type="checkbox"/>	7.50 <input type="checkbox"/>
IEEE Broadcast Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	BT002	15.00 <input type="checkbox"/>	7.50 <input type="checkbox"/>
IEEE Circuits and Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CAS004	22.00 <input type="checkbox"/>	11.00 <input type="checkbox"/>
IEEE Communications <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	COM019	30.00 <input type="checkbox"/>	15.00 <input type="checkbox"/>
IEEE Components, Packaging, & Manu. Tech. <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CPMT021	15.00 <input type="checkbox"/>	7.50 <input type="checkbox"/>
IEEE Computational Intelligence <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CIS011	29.00 <input type="checkbox"/>	14.50 <input type="checkbox"/>
IEEE Computer <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CO16	56.00 <input type="checkbox"/>	28.00 <input type="checkbox"/>
IEEE Consumer Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CE008	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Control Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CS023	25.00 <input type="checkbox"/>	12.50 <input type="checkbox"/>
IEEE Dielectrics and Electrical Insulation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	DEI032	26.00 <input type="checkbox"/>	13.00 <input type="checkbox"/>
IEEE Education <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	E025	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Electromagnetic Compatibility <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	EMC027	31.00 <input type="checkbox"/>	15.50 <input type="checkbox"/>
IEEE Electron Devices <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	ED015	18.00 <input type="checkbox"/>	9.00 <input type="checkbox"/>
IEEE Engineering in Medicine and Biology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	EMB018	40.00 <input type="checkbox"/>	20.00 <input type="checkbox"/>
IEEE Geoscience and Remote Sensing <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	GRS029	19.00 <input type="checkbox"/>	9.50 <input type="checkbox"/>
IEEE Industrial Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IE013	9.00 <input type="checkbox"/>	4.50 <input type="checkbox"/>
IEEE Industry Applications <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IA034	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Information Theory <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IT012	30.00 <input type="checkbox"/>	15.00 <input type="checkbox"/>
IEEE Instrumentation and Measurement <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IM009	29.00 <input type="checkbox"/>	14.50 <input type="checkbox"/>
IEEE Intelligent Transportation Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	ITSS038	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Magnetics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	MAG033	26.00 <input type="checkbox"/>	13.00 <input type="checkbox"/>
IEEE Microwave Theory and Techniques <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	MITT017	17.00 <input type="checkbox"/>	8.50 <input type="checkbox"/>
IEEE Nuclear and Plasma Sciences <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	NPS005	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Oceanic Engineering <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	OE022	19.00 <input type="checkbox"/>	9.50 <input type="checkbox"/>
IEEE Photonics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PHO036	34.00 <input type="checkbox"/>	17.00 <input type="checkbox"/>
IEEE Power Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PEL035	25.00 <input type="checkbox"/>	12.50 <input type="checkbox"/>
IEEE Power & Energy <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PE031	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Product Safety Engineering <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PSE043	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Professional Communication <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PC026	31.00 <input type="checkbox"/>	15.50 <input type="checkbox"/>
IEEE Reliability <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	RL007	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Robotics and Automation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	RA024	9.00 <input type="checkbox"/>	4.50 <input type="checkbox"/>
IEEE Signal Processing <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SP001	22.00 <input type="checkbox"/>	11.00 <input type="checkbox"/>
IEEE Social Implications of Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SIT030	33.00 <input type="checkbox"/>	16.50 <input type="checkbox"/>
IEEE Solid-State Circuits <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SSC037	22.00 <input type="checkbox"/>	11.00 <input type="checkbox"/>
IEEE Systems, Man, & Cybernetics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SMC028	12.00 <input type="checkbox"/>	6.00 <input type="checkbox"/>
IEEE Technology & Engineering Management <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	TEM014	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Ultrasonics, Ferroelectrics, & Frequency Control <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	UFFC020	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Vehicular Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	VT006	18.00 <input type="checkbox"/>	9.00 <input type="checkbox"/>

Legend—Society membership includes:

- One or more Society publications
- Online access to publication
- Society newsletter
- CD-ROM of selected society publications

Complete both sides of this form, sign, and return to:

IEEE MEMBERSHIP APPLICATION PROCESSING
 445 HOES LN, PISCATAWAY, NJ 08854-4141 USA
 or fax to +1 732 981 0225
 or join online at www.ieee.org/join

Please reprint your full name here _____

6 2016 IEEE Membership Rates (student rates available online)

IEEE member dues and regional assessments are based on where you live and when you apply. Membership is based on the calendar year from 1 January through 31 December. All prices are quoted in US dollars.

Please check the appropriate box.

	BETWEEN 16 AUG 2015- 28 FEB 2016 PAY	BETWEEN 1 MAR 2016- 15 AUG 2016 PAY
RESIDENCE		
United States.....	\$197.00 <input type="checkbox"/>	\$98.50 <input type="checkbox"/>
Canada (GST)*.....	\$173.35 <input type="checkbox"/>	\$86.68 <input type="checkbox"/>
Canada (NB, NF and ON HST)*.....	\$185.11 <input type="checkbox"/>	\$92.56 <input type="checkbox"/>
Canada (Nova Scotia HST)*.....	\$188.05 <input type="checkbox"/>	\$94.03 <input type="checkbox"/>
Canada (PEI HST)*.....	\$186.58 <input type="checkbox"/>	\$93.29 <input type="checkbox"/>
Canada (GST and QST Quebec).....	\$188.01 <input type="checkbox"/>	\$94.01 <input type="checkbox"/>
Africa, Europe, Middle East.....	\$160.00 <input type="checkbox"/>	\$80.00 <input type="checkbox"/>
Latin America.....	\$151.00 <input type="checkbox"/>	\$75.50 <input type="checkbox"/>
Asia, Pacific.....	\$152.00 <input type="checkbox"/>	\$76.00 <input type="checkbox"/>

Minimum Income or Unemployed Provision

Applicants who certify that their prior year income did not exceed US\$14,700 (or equivalent) or were not employed are granted 50% reduction in: full-year dues, regional assessment and fees for one IEEE Membership plus one Society Membership. If applicable, please check appropriate box and adjust payment accordingly. Student members are not eligible.

- I certify I earned less than US\$14,700 in 2015
- I certify that I was unemployed in 2015

7 More Recommended Options

- Proceedings of the IEEE..... print \$47.00 or online \$41.00
- Proceedings of the IEEE (print/online combination).....\$57.00
- IEEE Standards Association (IEEE-SA).....\$53.00
- IEEE Women in Engineering (WIE).....\$25.00

8 Payment Amount

Please total the Membership dues, Society dues, and other amounts from this page:

- IEEE Membership dues ⑥.....\$ _____
- IEEE Society dues (optional) ⑤.....\$ _____
- IEEE-SA/WIE dues (optional) ⑦.....\$ _____
- Proceedings of the IEEE (optional) ⑦.....\$ _____
- Canadian residents pay 5% GST or appropriate HST (BC—12%; NB, NF, ON—13%; NS—15%) on Society payments & publications only.....TAX \$ _____

AMOUNT PAID **TOTAL \$** _____

Payment Method

All prices are quoted in US dollars. You may pay for IEEE membership by credit card (see below), check, or money order payable to IEEE, drawn on a US bank.

Check

Credit Card Number _____

MONTH YEAR CARDHOLDER'S 5-DIGIT ZIP/PCODE (BILLING STATEMENT ADDRESS) USA ONLY

Name as it appears on card _____

Signature _____

Auto Renew my Memberships and Subscriptions (available when paying by credit card).
 I agree to the Terms and Conditions located at www.ieee.org/autorenew

9 Were You Referred to IEEE?

- Yes No If yes, provide the following:
- Member Recruiter Name _____
- IEEE Recruiter's Member Number (Required) _____

CAMPAIGN CODE _____ PROMO CODE _____

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/tasl/tasl-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/>

2016 IEEE SIGNAL PROCESSING SOCIETY MEMBERSHIP APPLICATION

Mail to: IEEE OPERATIONS CENTER, ATTN: Matthew Plotner, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA
 or Fax to (732) 981-0225 (credit card payments only.)
 For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org



1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: SEND MAIL TO: Home Address OR Business/School Address
 If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE	FIRST OR GIVEN NAME	MIDDLE NAME	SURNAME/LAST NAME
HOME ADDRESS			
CITY		STATE/PROVINCE	POSTAL CODE COUNTRY

2 Are you now or were you ever a member of IEEE? Yes No
 If yes, please provide, if known:
 MEMBERSHIP NUMBER _____
 Grade _____ Year Membership Expired: _____

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name _____
 Department/Division _____
 Title/Position _____ Years in Current Position _____
 Years in the Profession Since Graduation _____ PE State/Province _____
 Street Address _____
 City _____ State/Province _____ Postal Code _____ Country _____

4. EDUCATION

A baccalaureate degree from an IEEE recognized educational program assures assignment of "Member" grade. For others, additional information and references may be necessary for grade assignment.

A. Baccalaureate Degree Received _____ Program/Course of Study _____
 College/University _____ Campus _____
 State/Province _____ Country _____ Mo./Yr. Degree Received _____

B. Highest Technical Degree Received _____ Program/Course of Study _____
 College/University _____ Campus _____
 State/Province _____ Country _____ Mo./Yr. Degree Received _____

5. Full signature of applicant _____

6. DEMOGRAPHIC INFORMATION – ALL APPLICANTS -

Date of Birth _____ Male Female
 Day _____ Month _____ Year _____

7. CONTACT INFORMATION

Office Phone/Office Fax _____ Home Phone/Home Fax _____
 Office E-Mail _____ Home E-Mail _____

2016 IEEE MEMBER RATES		
IEEE DUES	16 Aug-14-28 Feb 15	1 Mar -15 Aug 15
Residence	Pav Full Year	Pav Half Year**
United States	\$197.00 <input type="checkbox"/>	\$98.50 <input type="checkbox"/>
Canada (incl. GST)	\$173.35 <input type="checkbox"/>	\$86.68 <input type="checkbox"/>
Canada (incl. HST for PEI)	\$186.58 <input type="checkbox"/>	\$93.29 <input type="checkbox"/>
Canada (incl. HST for Nova Scotia)	\$188.05 <input type="checkbox"/>	\$94.03 <input type="checkbox"/>
Canada (incl. HST for NB, NF and ON)	\$185.11 <input type="checkbox"/>	\$92.56 <input type="checkbox"/>
Canada (incl. GST and GST Quebec)	\$188.01 <input type="checkbox"/>	\$94.01 <input type="checkbox"/>
Africa, Europe, Middle East	\$160.00 <input type="checkbox"/>	\$80.00 <input type="checkbox"/>
Latin America	\$151.00 <input type="checkbox"/>	\$75.50 <input type="checkbox"/>
Asia, Pacific	\$152.00 <input type="checkbox"/>	\$76.00 <input type="checkbox"/>

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 12%, 13% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)

Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)

U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin. Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.



2016 SPS MEMBER RATES

	16 Aug-28 Feb	1 Mar-15 Aug
	Pav Full Year	Pav Half Year
Signal Processing Society Membership Fee*	\$ 22.00 <input type="checkbox"/>	\$ 11.00 <input type="checkbox"/>
<small>Fee includes: IEEE Signal Processing Magazine (electronic and digital), Inside Signal Proc. eNewsletter (electronic) and IEEE Signal Processing Society Content Gazette (electronic).</small>		
Add \$17 to enhance SPS Membership and also receive:	\$17.00 <input type="checkbox"/>	\$ 8.50 <input type="checkbox"/>
IEEE Signal Processing Magazine (print) and SPS Digital Library: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.		
<i>Publications available only with SPS membership:</i>		
Signal Processing, IEEE Transactions on:	Print \$209.00 <input type="checkbox"/>	\$104.50 <input type="checkbox"/>
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print \$160.00 <input type="checkbox"/>	\$ 80.00 <input type="checkbox"/>
Image Processing, IEEE Transactions on:	Print \$207.00 <input type="checkbox"/>	\$103.50 <input type="checkbox"/>
Information Forensics and Security, IEEE Trans. on:	Print \$179.00 <input type="checkbox"/>	\$ 89.50 <input type="checkbox"/>
IEEE Journal of Selected Topics in Signal Processing:	Print \$176.00 <input type="checkbox"/>	\$ 88.00 <input type="checkbox"/>
Affective Computing, IEEE Transactions on:	Electronic \$ 36.00 <input type="checkbox"/>	\$ 18.00 <input type="checkbox"/>
Biomedical and Health Informatics, IEEE Journal of:	Print \$ 55.00 <input type="checkbox"/>	\$ 27.50 <input type="checkbox"/>
	Electronic \$ 40.00 <input type="checkbox"/>	\$ 20.00 <input type="checkbox"/>
	Print & Electronic \$ 65.00 <input type="checkbox"/>	\$ 32.50 <input type="checkbox"/>
IEEE Cloud Computing	Electronic and Digital \$ 39.00 <input type="checkbox"/>	\$ 19.50 <input type="checkbox"/>
IEEE Trans. on Cognitive Comm. & Networking	Electronic \$ 27.00 <input type="checkbox"/>	\$ 13.50 <input type="checkbox"/>
IEEE Trans. on Computational Imaging	Electronic \$ 28.00 <input type="checkbox"/>	\$ 14.00 <input type="checkbox"/>
IEEE Trans. on Big Data	Electronic \$ 26.00 <input type="checkbox"/>	\$ 13.00 <input type="checkbox"/>
IEEE Trans. on Molecular, Biological, & Multi-scale Communications	Electronic \$ 25.00 <input type="checkbox"/>	\$ 12.50 <input type="checkbox"/>
IEEE Internet of Things Journal	Electronic \$ 26.00 <input type="checkbox"/>	\$ 13.00 <input type="checkbox"/>
IEEE Trans. on Cloud Computing	Electronic \$ 43.00 <input type="checkbox"/>	\$ 21.50 <input type="checkbox"/>
IEEE Trans. on Computational Social Systems	Electronic \$ 30.00 <input type="checkbox"/>	\$ 15.00 <input type="checkbox"/>
IEEE Trans. on Signal & Info Proc. Over Networks	Electronic \$ 28.00 <input type="checkbox"/>	\$ 14.00 <input type="checkbox"/>
IEEE Biometrics Compendium:	Online \$ 30.00 <input type="checkbox"/>	\$ 15.00 <input type="checkbox"/>
Computing in Science & Engrg. Mag.:	Electronic and Digital \$ 39.00 <input type="checkbox"/>	\$ 19.50 <input type="checkbox"/>
	Print \$ 69.00 <input type="checkbox"/>	\$ 34.50 <input type="checkbox"/>
Medical Imaging, IEEE Transactions on:	Print \$ 74.00 <input type="checkbox"/>	\$ 37.00 <input type="checkbox"/>
	Electronic \$ 53.00 <input type="checkbox"/>	\$ 26.50 <input type="checkbox"/>
	Print & Electronic \$ 89.00 <input type="checkbox"/>	\$ 44.50 <input type="checkbox"/>
Mobile Computing, IEEE Transactions on:		
	ELE/Print Abstract/CD-ROM \$ 41.00 <input type="checkbox"/>	\$ 20.50 <input type="checkbox"/>
Multimedia, IEEE Transactions on:	Electronic \$ 43.00 <input type="checkbox"/>	\$ 21.50 <input type="checkbox"/>
IEEE Multimedia Magazine:	Electronic and Digital \$ 39.00 <input type="checkbox"/>	\$ 19.50 <input type="checkbox"/>
	Print \$ 69.00 <input type="checkbox"/>	\$ 34.50 <input type="checkbox"/>
Network Science and Engrg., IEEE Trans. on:	Electronic \$ 34.00 <input type="checkbox"/>	\$ 17.00 <input type="checkbox"/>
IEEE Reviews in Biomedical Engineering:	Print \$ 25.00 <input type="checkbox"/>	\$ 12.50 <input type="checkbox"/>
	Print & Electronic \$ 40.00 <input type="checkbox"/>	\$ 20.00 <input type="checkbox"/>
IEEE Security and Privacy Magazine:	Electronic and Digital \$ 39.00 <input type="checkbox"/>	\$ 19.50 <input type="checkbox"/>
	Print \$ 69.00 <input type="checkbox"/>	\$ 34.50 <input type="checkbox"/>
IEEE Sensors Journal:	Electronic \$ 40.00 <input type="checkbox"/>	\$ 20.00 <input type="checkbox"/>
Smart Grid, IEEE Transactions on:	Print \$100.00 <input type="checkbox"/>	\$ 50.00 <input type="checkbox"/>
	Electronic \$ 40.00 <input type="checkbox"/>	\$ 20.00 <input type="checkbox"/>
	Print & Electronic \$120.00 <input type="checkbox"/>	\$ 60.00 <input type="checkbox"/>
Wireless Communications, IEEE Transactions on:	Print \$127.00 <input type="checkbox"/>	\$ 63.50 <input type="checkbox"/>
	Electronic \$ 49.00 <input type="checkbox"/>	\$ 24.50 <input type="checkbox"/>
	Print & Electronic \$127.00 <input type="checkbox"/>	\$ 63.50 <input type="checkbox"/>
IEEE Wireless Communications Letters:	Electronic \$ 19.00 <input type="checkbox"/>	\$ 9.50 <input type="checkbox"/>

*IEEE membership required or requested
 Affiliate application to join SP Society only. Amount Paid \$ _____

9. IEEE Membership Affiliate Fee (See pricing in Section 8) \$ _____
Signal Processing Society Fees \$ _____
 Canadian residents pay 5% GST or 13% HST
 Reg. No. 125634188 on Society payment(s) & pubs only Tax \$ _____

AMOUNT PAID WITH APPLICATION TOTAL \$ _____
Prices subject to change without notice.

Check or money order enclosed Payable to IEEE on a U.S. Bank
 American Express VISA MasterCard
 Diners Club

Exp. Date/ Mo./Yr.	_____	_____	_____	_____
Cardholder Zip Code Billing Statement Address/USA Only	_____	_____	_____	_____

Full signature of applicant using credit card _____ Date _____

10. WERE YOU REFERRED?

Yes No If yes, please provide the following information:
 Member Recruiter Name: _____
 IEEE Recruiter's Member Number (Required): _____

2016 IEEE SIGNAL PROCESSING SOCIETY STUDENT MEMBERSHIP APPLICATION

(Current and reinstating IEEE members joining SPS complete areas 1, 2, 8, 9.)

Mail to: IEEE OPERATIONS CENTER, ATTN: Matthew Plotner, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA
or Fax to (732) 981-0225 (credit card payments only.)

For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org



1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: SEND MAIL TO: Home Address OR Business/School Address
If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE	FIRST OR GIVEN NAME	MIDDLE NAME	SURNAME/LAST NAME
HOME ADDRESS			
CITY	STATE/PROVINCE	POSTAL CODE	COUNTRY

2. Are you now or were you ever a member of IEEE? Yes No

If yes, please provide, if known:
MEMBERSHIP NUMBER _____
Grade _____ Year Membership Expired: _____

2016 SPS STUDENT MEMBER RATES

	16 Aug-28 Feb	1 Mar-15 Aug
	Pay Full Year	Pay Half Year
Signal Processing Society Membership Fee*	\$11.00 <input type="checkbox"/>	\$ 5.50 <input type="checkbox"/>
Fee includes: IEEE Signal Processing Magazine (electronic and digital), Inside Signal Processing eNewsletter (electronic) and IEEE Signal Processing Society Content Gazette (electronic).		
Add \$8 to enhance SPS Membership and also receive:	\$ 9.00 <input type="checkbox"/>	\$ 4.50 <input type="checkbox"/>
IEEE Signal Processing Society Magazine (print) and SPS Digital Library: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.		
<i>Publications available only with SPS membership:</i>		
Signal Processing, IEEE Transactions on:	Print \$105.00 <input type="checkbox"/>	\$ 52.50 <input type="checkbox"/>
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print \$ 80.00 <input type="checkbox"/>	\$ 40.00 <input type="checkbox"/>
Image Processing, IEEE Transactions on:	Print \$104.00 <input type="checkbox"/>	\$ 52.00 <input type="checkbox"/>
Information Forensics and Security, IEEE Trans. on:	Print \$ 90.00 <input type="checkbox"/>	\$ 45.00 <input type="checkbox"/>
IEEE Journal of Selected Topics in Signal Processing:	Print \$ 88.00 <input type="checkbox"/>	\$ 44.00 <input type="checkbox"/>
Affective Computing, IEEE Transactions on:	Electronic \$ 18.00 <input type="checkbox"/>	\$ 9.00 <input type="checkbox"/>
Biomedical and Health Informatics, IEEE Journal of:	Print \$ 28.00 <input type="checkbox"/>	\$ 14.00 <input type="checkbox"/>
	Electronic \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
	Print & Electronic \$ 33.00 <input type="checkbox"/>	\$ 16.50 <input type="checkbox"/>
IEEE Cloud Computing	Electronic and Digital \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
IEEE Trans. on Cognitive Comm. & Networking	Electronic \$ 14.00 <input type="checkbox"/>	\$ 7.00 <input type="checkbox"/>
IEEE Trans. on Computational Imaging	Electronic \$ 14.00 <input type="checkbox"/>	\$ 7.00 <input type="checkbox"/>
IEEE Trans. on Big Data	Electronic \$ 13.00 <input type="checkbox"/>	\$ 6.50 <input type="checkbox"/>
IEEE Trans. on Molecular, Biological, & Multi-Scale Communications	Electronic \$ 13.00 <input type="checkbox"/>	\$ 6.50 <input type="checkbox"/>
IEEE Internet of Things Journal	Electronic \$ 13.00 <input type="checkbox"/>	\$ 6.50 <input type="checkbox"/>
IEEE Trans. on Cloud Computing	Electronic \$ 22.00 <input type="checkbox"/>	\$ 11.00 <input type="checkbox"/>
IEEE Trans. on Computational Social Systems	Electronic \$ 15.00 <input type="checkbox"/>	\$ 7.50 <input type="checkbox"/>
IEEE Trans. on Signal & Info Proc. Over Networks	Electronic \$ 14.00 <input type="checkbox"/>	\$ 7.00 <input type="checkbox"/>
IEEE Biometrics Compendium:	Online \$ 15.00 <input type="checkbox"/>	\$ 7.50 <input type="checkbox"/>
Computing in Science & Engrg. Mag.:	Electronic and Digital \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
	Print \$ 35.00 <input type="checkbox"/>	\$ 17.50 <input type="checkbox"/>
Medical Imaging, IEEE Transactions on:	Print \$ 37.00 <input type="checkbox"/>	\$ 18.50 <input type="checkbox"/>
	Electronic \$ 27.00 <input type="checkbox"/>	\$ 13.50 <input type="checkbox"/>
	Print & Electronic \$ 45.00 <input type="checkbox"/>	\$ 22.50 <input type="checkbox"/>
Mobile Computing, IEEE Transactions on:		
	ELE/Print Abstract/CD-ROM \$ 21.00 <input type="checkbox"/>	\$ 10.50 <input type="checkbox"/>
Multimedia, IEEE Transactions on:	Electronic \$ 22.00 <input type="checkbox"/>	\$ 11.00 <input type="checkbox"/>
IEEE MultiMedia Magazine:	Electronic and Digital \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
	Print \$ 35.00 <input type="checkbox"/>	\$ 17.50 <input type="checkbox"/>
Network Science and Engrg., IEEE Trans. on:	Electronic \$ 17.00 <input type="checkbox"/>	\$ 8.50 <input type="checkbox"/>
IEEE Reviews in Biomedical Engineering:	Print \$ 13.00 <input type="checkbox"/>	\$ 6.50 <input type="checkbox"/>
	Print & Electronic \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
IEEE Security and Privacy Magazine:	Electronic and Digital \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
	Print \$ 35.00 <input type="checkbox"/>	\$ 17.50 <input type="checkbox"/>
IEEE Sensors Journal:	Electronic \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
Smart Grid, IEEE Transactions on:	Print \$ 50.00 <input type="checkbox"/>	\$ 25.00 <input type="checkbox"/>
	Electronic \$ 20.00 <input type="checkbox"/>	\$ 10.00 <input type="checkbox"/>
	Print & Electronic \$ 60.00 <input type="checkbox"/>	\$ 30.00 <input type="checkbox"/>
Wireless Communications, IEEE Transactions on:	Print \$ 64.00 <input type="checkbox"/>	\$ 32.00 <input type="checkbox"/>
	Electronic \$ 25.00 <input type="checkbox"/>	\$ 12.50 <input type="checkbox"/>
	Print & Electronic \$ 64.00 <input type="checkbox"/>	\$ 32.00 <input type="checkbox"/>
IEEE Wireless Communications Letters:	Electronic \$ 10.00 <input type="checkbox"/>	\$ 5.00 <input type="checkbox"/>

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name _____
Department/Division _____
Title/Position _____ Years in Current Position _____
Years in the Profession Since Graduation _____ PE State/Province _____
Street Address _____
City _____ State/Province _____ Postal Code _____ Country _____

4. EDUCATION A baccalaureate degree from an IEEE recognized educational program assures assignment of "Member" grade. For others, additional information and references may be necessary for grade assignment.

A. Baccalaureate Degree Received _____ Program/Course of Study _____
College/University _____ Campus _____
State/Province _____ Country _____ Mo./Yr. Degree Received _____
B. Highest Technical Degree Received _____ Program/Course of Study _____
College/University _____ Campus _____
State/Province _____ Country _____ Mo./Yr. Degree Received _____

5. Full signature of applicant _____

6. DEMOGRAPHIC INFORMATION – ALL APPLICANTS -

Date Of Birth _____ Male Female
Day _____ Month _____ Year _____

7. CONTACT INFORMATION

Office Phone/Office Fax _____ Home Phone/Home Fax _____
Office E-Mail _____ Home E-Mail _____

8. 2016 IEEE STUDENT MEMBER RATES

IEEE DUES	16 Aug 14-28 Feb 15	1 Mar -15 Aug 15
Residence	Pay Full Year	Pay Half Year**
United States	\$32.00 <input type="checkbox"/>	\$16.00 <input type="checkbox"/>
Canada (incl. GST)	\$33.60 <input type="checkbox"/>	\$16.80 <input type="checkbox"/>
Canada (incl. HST for NB, NF, and ON)	\$36.16 <input type="checkbox"/>	\$18.08 <input type="checkbox"/>
Canada (incl. HST for Nova Scotia)	\$36.80 <input type="checkbox"/>	\$18.40 <input type="checkbox"/>
Canada (incl. HST for PEI)	\$36.48 <input type="checkbox"/>	\$18.24 <input type="checkbox"/>
Canada (incl. GST and QST Quebec)	\$36.79 <input type="checkbox"/>	\$18.40 <input type="checkbox"/>
Africa, Europe, Middle East, Latin America, Asia, Pacific	\$27.00 <input type="checkbox"/>	\$13.50 <input type="checkbox"/>

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 12%, 13% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)

Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)

U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin. Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.



9. IEEE Membership Fee (See pricing in Section 8) \$ _____

Signal Processing Society Fees \$ _____

Canadian residents pay 5% GST or 13% HST
Reg. No. 125634188 on Society payment(s) & pubs only Tax \$ _____

AMOUNT PAID WITH APPLICATION TOTAL \$ _____
Prices subject to change without notice.

Check or money order enclosed Payable to IEEE on a U.S. Bank
 American Express VISA MasterCard Diners Club

Exp. Date	Mo./Yr.								
-----------	---------	--	--	--	--	--	--	--	--

Cardholder	Zip Code	Billing Statement	Address/USA Only		
------------	----------	-------------------	------------------	--	--

Full signature of applicant using credit card _____ Date _____

10. WERE YOU REFERRED?

Yes No If yes, please provide the follow information:

Member Recruiter Name: _____

IEEE Recruiter's Member Number (Required): _____

