IEEE Signal Processing MAGAZINE

[VOLUME 32 NUMBER 1 JANUARY 2015]

QUANTITATIVE BIOIMAGING

SIGNAL PROCESSING IN LIGHT MICROSCOPY

VIDEO QUALITY ASSESSMENT

HUMAN–MACHINE INTERFACING BY DECODING THE SURFACE ELECTROMYOGRAM

ACCURACY, APPS ADVANCE SPEECH RECOGNITION

IEEE Signal Processing Society

IEEE

CONTENTS

[VOLUME 32 NUMBER 1]

SPECIAL SECTION—QUANTITATIVE BIOIMAGING

FEATURE

COLUMNS

DEPARTMENT

Digital Object Identifier 10.1109/MSP.2014.2364658

IEEE SIGNAL PROCESSING magazine

COVER

COMPUTER AND MICROSCOPE: IMAGES LICENSED BY INGRAM PUBLISHING. IMAGE ON LAPTOP COURTESY OF MICHAEL MCCANN.

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

IEEE SIGNAL PROCESSING MAGAZINE (ISSN 1053-5888) (ISPREG) is published bimonthly by the Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, 17th Floor, New York, NY 10016-5997 USA (+1 212 419 7900). Responsibility for the contents rests upon the authors and not the IEEE, the Society, or its members. Annual member subscriptions included in Society fee. Nonmember subscriptions available upon request. Individual copies: IEEE Members US$20.00 (first copy only), nonmembers US$201.00 per copy.

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.2014.2364660

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.

Call for Papers

www.icip2015.org

The International Conference on Image Processing (ICIP), sponsored by the IEEE Signal Processing Society, is the premier forum for the presentation of technological advances and research results in the fields of theoretical, experimental, and applied image and video processing. ICIP 2015, the twenty second in the series that has been held annually since 1994, brings together leading engineers and scientists in image and video processing from around the world. Research frontiers in fields ranging from traditional image processing applications to evolving multimedia and video technologies are regularly advanced by results first reported in ICIP technical sessions.

Topics include, but are not limited to:

- **Image/video coding and transmission:** Still-image and video coding, stereoscopic and 3-D coding, distributed source coding, source/channel coding, image/video transmission over wireless networks;
- **Image/video processing:** Image and video filtering, restoration and enhancement, image segmentation, video segmentation and tracking, morphological processing, stereoscopic and 3-D processing, feature extraction and analysis, interpolation and super-resolution, motion detection and estimation, color and multispectral processing, biometrics;
- **Image formation:** Biomedical imaging, remote sensing, geophysical and seismic imaging, optimal imaging, synthetic-natural hybrid image systems;
- **Image scanning, display, and printing**: Scanning and sampling, quantization and half toning, color reproduction, image representation and rendering, display and printing systems, image-quality assessment;
- **Image/video storage, retrieval, and authentication:** Image and video databases, image and video search and retrieval, multimodality image/ video indexing and retrieval, authentication and watermarking;
- **Applications:** Biomedical sciences, mobile imaging, geosciences & remote sensing, astronomy & space exploration, document image processing and analysis, other applications.

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.

Call for Tutorials: Tutorials will be held on Sunday, September 27, 2015. Proposals must be submitted by January 15, 2015 via email to tutorials@icip2015.org. For detailed submission guidelines, please visit the *Call for tutorials* section of the web site.

Call for Special Sessions: Proposals should be submitted by November 27, 2014 in a single PDF document sent to specialsessions@icip2015.org. For detailed submission guidelines, please visit the *Call for Special Sessions* section of the web site.

Important Dates

- Special Session Proposals: **27 November 2014**
- Papers Submissions: **15 January 2015**
- Tutorial Proposals: **15 January 2015**
- Authors Notification: **30 April 2015**

Visit icip2015.org for details on paper submission, social events, no-show policy, and more.

Organizing Committee

General Chairs
Jean-Luc DUGELAY
André MORIN

Technical Program Chairs
Fabrice LABEAU
Jean-Philippe THIRAN

Finance
Jean FORTIN

Plenary Sessions
Stéphane COULOMBE
Kenneth ROSE

Special Sessions
Oscar C. AU
Éric DUBOIS

Tutorials
Janusz KONRAD
André ZACCARIN

Local Arrangements
Paul FORTIER

Registration
Xavier MALDAGUE

Exhibit/Industry
Khaled EL-MALEH
Branislav KISACANIN

Publicity
Aishy AMER
Patrick LE CALLET

Publications
Mireille BOUTIN

Electronic Media
Abdulmotaleb EL SADDIK
Benoît HUET

International Liaison
Carlo S. REGAZZONI
Wan-Chi SIU

Student Activities
Sylvie DANIEL
Guoliang FAN

Awards
Phil CHOU

Québec
Office du tourisme de Québec

[from the EDITOR]

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

Taking Up the Torch

It will almost be the start of a new year when you receive this issue of *IEEE Signal Processing Magazine* (*SPM*). Happy New Year to you all!

The new year accompanies major changes in *SPM's* management team. The 2012–2014 editor-in-chief, Abdelhak Zoubir, and his capable executive team of senior area editors—Fulvio Gini (special issues), Marc Moonen (feature articles), and Andrea Cavallaro and Andres Kwanski (columns and forums)—have successfully completed their terms in leading the magazine for the past three years. Given the production lead time required for the magazine, much of the technical material in this and the next several issues are attributed to their efforts. My sincere thanks are given for all of their hard work and also for their generous mentoring and assistance to help me ramp up the learning curve!

It is truly an honor and privilege for me to take on the editor-in-chief position of the magazine for the next three years. Joining me to form the new *SPM* executive team are four new senior area editors: Wade Trappe (special issues), Shuguang Robert Cui (feature articles), and Gwenael Doerr and Kenneth Lam (columns and forums), as well as Christian Debes, who was on board last year for the eNewsletter. Together with a dedicated group of associate editors and members of the senior editorial board, we are ready to take the journey to serve you.

SPM reaches over 20,000 readers worldwide across a diverse range of career sectors and is very different from the transactions-type of research journals with a focused scope and audience. I have been an *SPM* subscriber since my senior year in college. The magazine's tutorial surveys and overview articles gave a student like me a high-quality yet accessible way to take a peek into this incredible field. *SPM* traces its roots back to 1970 as *IEEE Audio and Electroacoustics Newsletter*. Accompanying the expanded technical scope and a then new name of the IEEE Signal Processing Society (SPS), *SPM* in its current name and form was first published in 1991. For a quarter of a century since then, six predecessors leading *SPM* before me have made strong marks in growing *SPM* into a premier publication to serve our signal processing community. As I take up the torch with awe, I ask myself: what more can be done?

WHILE CITATION IMPACT IS AN IMPORTANT METRIC, THE ROLE AND IMPACT OF *SPM* IS MUCH MORE THAN SCORING WELL IN A CITATION RANKING.

SPM has been a leader in citation impact among 200+ publications in electrical and computer engineering as indexed by the Thomson Reuters' annual *Journal Citation Report* (*JCR*). Over the past decade, *SPM* has made it to the top (multiple times) of *JCR's* "popularity measure" known as the *impact factor*, which is based on the average number of received citations in a given year for each article published in the preceding two years. Also by Thomson Reuters, a newer "prestige measure" known as the *article influence score* accounts for both the citation amount and the influences of the citing journals that have contributed to the citations; by excluding self-citations and considering a longer window of five years, the article influence score is more robust to intentional inflation and accounts for a longer-term influence. Currently, *SPM* ranks at the top in article influence score among all 250 electrical and computer engineering publications indexed by Thomson Reuters—what an achievement for *SPM*'s past editorial teams and authors!

To maintain a high citation impact, *SPM* needs to continue to identify topics of keen interest to the R&D community in the broad areas of signal processing and attract authors who can write timely and high-quality tutorial and survey articles on these topics. This is easier said than done. Citation statistics often have a few years' lag time, so actions taken by an editorial team may not be reflected in the citation until toward the end of a three-year term or well after. But help from you, our readers and community colleagues, can complement the effort by our editorial board, so that together we can sustain a high citation impact for *SPM*. For example, let us know the topics you are interested in learning about; offer us your feedback regarding the articles you see in *SPM*—what aspects you love and what aspects you hope we improve or do differently; for the articles that you find beneficial, help us spread the word by recommending them to colleagues and students, and cite these articles in your relevant writings and presentations. And if you are working on topic areas that may stimulate broad interests, consider proposing the topic for a special issue, a feature article, or a column, and publish in *SPM*!

While citation impact is an important metric, the role and impact of *SPM* is much more than scoring well in a citation ranking. A substantial number of our members and readers are not in academia or industrial academia and do not publish regularly as a major part of their jobs. They include, but are not limited to, industry practitioners as well as undergraduate and

Digital Object Identifier 10.1109/MSP.2014.2363499
Date of publication: 5 December 2014

nonresearch-track students. *SPM,* as the Society's only publication going to every member, has a distinguished tradition of offering valuable information to these readers. I will do my best to continue to serve you and make *SPM* an important part of your career advancement and personal enrichment.

In today's well-connected world, what are the best ways for *SPM* to bring signal processing-related information to you? We plan to explore ways both through the content in the traditional magazine format as well as complementary channels such as electronic, online, and social media platforms to bring you useful information. Some of these means may be in cooperation with other SPS publications, boards, and committees. I look forward to learning your thoughts on *SPM* and meeting you in the months to come through SPS conferences, Distinguished Lecturer tours, and other local Chapter events. Until then, I wish you all a healthy and productive new year!

[SP]

ERRATA

In the article "Location-Aware Communications for 5G Networks" by R. Di Taranto et al., *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102–112 [1], math was typeset incorrectly due to a production error in (2) and in the text directly following (5). In (2), a "1" should not appear before the "C" on the left-hand side of the equation. The correct way it should be displayed is as follows:

$$C(\mathbf{x}_i, \mathbf{x}_j) = \mathbb{E}\{\Psi(\mathbf{x}_s, \mathbf{x}_i)\Psi(\mathbf{x}_s, \mathbf{x}_j)\} = \sigma_\psi^2 \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{d_c}\right), \quad (2)$$

In the second line under (5) on page 106, the subscript asterisk should not be followed by a period. The correct way it should be displayed is as follows: in which $\mathbf{k}_*$ is the $N \times 1$ vector of cross-covariances $C(\mathbf{x}_*, \mathbf{x}_i)$ between $\mathbf{x}_*$ and the training inputs $\mathbf{x}_i$.

Also, two authors' names were misprinted. The corrected names are L. Srikar Muppirisetty and Dirk Slock. We apologize for the errors and any confusion they may have caused.

Reference

[1] R. Di Taranto, L. S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks," *IEEE Signal Processing Mag.*, vol. 31, no. 6, pp. 102–112, Nov. 2014.

Digital Object Identifier 10.1109/MSP.2014.2364864
Date of publication: 5 December 2014

Digital Object Identifier 10.1109/MSP.2014.2373471

president's MESSAGE

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

SigView: Video Tutorials in Emerging Signal Processing Topics

In September 2013, the IEEE Signal Processing Society (SPS) launched SigView, an online video library of tutorials in emerging areas of signal processing. SigView acts as a supplemental resource to the SPS's host of educational resources with a focus on new, hot topics in signal processing. As of October 2014, the rapidly growing video library hosts more than 40 videos. The most frequently accessed video, a tutorial on compressive sensing, has more than 6,000 views. You can access that video and many more—including our International Conference on Acoustics, Speech, and Signal Processing 2014 presentations—at http://www.sigview.org. Make sure to sign in using your SPS login information to access member-exclusive content, as well!

When our members want to learn about the very latest in the field, they attend conferences, read conference papers, and read brief *IEEE Signal Processing Letters*; when they want to see a more comprehensive treatment of a subject, they read longer journal articles—a tradeoff for the fact that it may not include the most recent developments.

IEEE Signal Processing Magazine (*SPM*) offers many tutorial-style articles for members who want to learn about new fields, which can be difficult to learn through the Society's more formal journals and conference papers. In addition, *SPM* is highly rated as number three among all IEEE publications in the five-year impact factor of Thomson Reuter's *Journal Citation Report*, a true testament of its quality and value to readers.

But still, some people find it more engaging to watch and listen to an educational lecture. This is why video tutorials are an integral part of the modern era of research and why the SPS wants to continue to engage its members through a variety of resources across platforms. Unfortunately, there are many poor-quality lectures out there. SigView was born as a means to provide well-researched, well-presented signal processing-related tutorial content that meets the Society's standards, treated with similar editorial considerations as our publications, products, and services. SigView offers some interactive features, such as the opportunity for readers to rate the videos and add comments, a key aspect of the modern Web.

Who are the distinguished speakers in these videos? You are! We strongly urge you, the Society's experts, to create SigView tutorials of your own. SigView tutorials are a great way to not only share knowledge and information about new topics but they are a great way for SPS members to exchange ideas, network, and coordinate with one another. Creating a video is easy and only requires the raw materials you want to include—just log in, upload your materials, and sequence and narrate the resulting video. You can mix and match all manner of materials including videos, PowerPoint, PDF, Word documents, images, audio, and much more. SigView automatically integrates all these different media types and allows you to narrate each, requiring nothing more than a microphone. The results are there for viewing and editing immediately, and once you're finished, they are hosted on the SigView site ad-free. For more information on how to create a SigView video, including a video tutorial, visit the SPS Web page: http://www.signalprocessingsociety.org/community/online-tutorial-library. To propose a submission, e-mail your proposed title to ad.sigview@ieee.org.

For any member who would like to get involved as a reviewer or editorial board member, the SigView editorial board asks you to get in touch. These people are responsible for screening submissions for quality and accuracy and assure that the Society's standards are upheld. If you're interested in joining, e-mail ce.sigview@ieee.org to learn how to get involved.

The SPS is always looking for new ways to deliver research, knowledge, and information to its members. If you have any ideas about how to keep SPS at the forefront of providing content, please don't hesitate to reach out to us.

I hope that you'll take the time to check out some of our SigView presentations and maybe even submit your own. Our Society and community is what we make it. Let's keep learning and sharing together.

[SP]

Digital Object Identifier 10.1109/MSP.2014.2363824
Date of publication: 5 December 2014

[special REPORTS]

John Edwards

Signal Processing in Next-Generation Prosthetics

Artificial devices that allow amputees to either walk again or to continue to use an arm and hand have existed since ancient times, the most notable example being the simple peg leg. Modern prosthetics, however, are a world apart from predigital-age artificial limbs. Using advanced robotic, cybernetic, and fabrication technologies, as well as signal processing, researchers are well on their way to making prosthetics that are remarkably useful, realistic, and intuitive.

As prosthetic development advances, signal processing is playing important roles in control and sensing operations. "The sensor is our window to the world," says Veronica J. Santos (Figure 1), an associate professor of mechanical and aerospace engineering at the University of California at Los Angeles. Santos, who is codeveloping a multimodal tactile sensor skin that's designed to assist prosthetic hand users, notes that signal processing is essential to the project. "Without signal processing, we wouldn't be able to measure anything, including the types of inferences we would have to make about the size, shape, or type of object and, therefore, how to grasp and manipulate it."

[FIG1] **Dr. Santos (left) is shown with Biomechatronics Lab members Ryan Manis (right, back) and Randall Hellman. Ryan is mounting a tactile sensor on a robot hand while Randall is adjusting the lab's custom remote actuation system for tendon-driven mechanisms. (Photo courtesy of *ASU Magazine* and Dan Vermillion of Vermillion Studio.)**

AIMING FOR INTUITIVE CONTROL

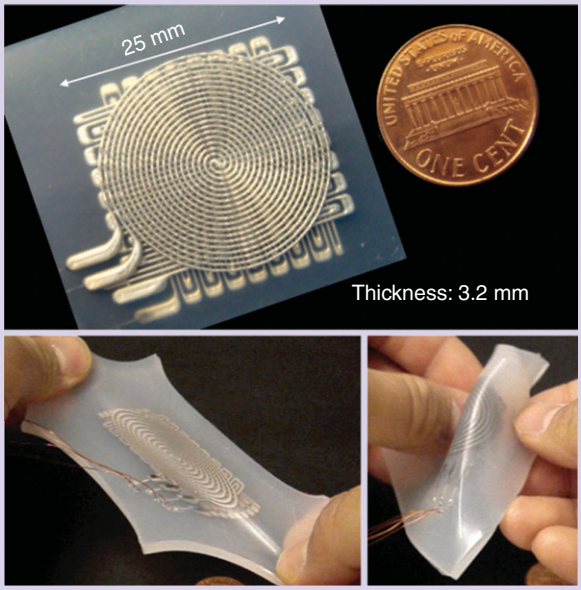
Santos notes that all prosthetic hand researchers face the same fundamental challenge. "Whether a prosthetic hand is a simple body-powered hook or an advanced anthropomorphic device, it will only be useful and desirable to an amputee if it is intuitive to control, functional, and improves the quality of life," she explains. "A prosthesis will be rejected if it poses too great of a cognitive burden on the user."

Santos and colleague Jonathan Posner, an associate professor of mechanical engineering at the University of Washington in Seattle, are developing a multimodal tactile sensor skin that promises to reduce the cognitive burden on prosthetic hand users, making control a faster, intuitive, and more natural process.

The researchers' microfluidic, capacitive-based tactile skin is designed to conformally wrap around curved surfaces, such as a prosthetic hand's fingertips. A current prototype can detect various finger-object interactions, such as normal contact forces and low-frequency dynamic loads. Its pliable, sensor-laden skin, which is closely wrapped around a prosthetic hand's curved digits, aims to help users grip objects by cushioning impacts, increasing the effective contact area during grasp, and enabling activities of daily living that rely upon a sense of touch. "We are developing algorithms to map artificial tactile sensor data to object properties for restoring the sense of touch to amputees," Santos says.

The system's capacitive sensors are created by injecting a flexible material, such as a polydimethyl siloxane (PDMS) polymer, with a liquid metal alloy that's designed to function as deformable wires and plates. "If we monitor the voltage across those plates we can relate that data to the distance between the plates, which we can then relate to the deformation of the skin surface," Santos explains. "The

Digital Object Identifier 10.1109/MSP.2014.2358718
Date of publication: 5 December 2014



[FIG2] An example of soft artificial skin that can detect multiaxis (x and y) strain (i.e., stretch) and contact pressure simultaneously. (Photo courtesy of *IEEE Sensors Journal*.)

raw voltages can be turned into something representing a contact area, an indication of force being applied, or the hardness of an object, for example." Such information can be relayed to amputees via neural interfaces or used to develop intelligent artificial reflexes and grasp patterns within the prosthesis itself.

With funding from the National Science Foundation, Santos and Posner are continuing to add sensing modalities and refine the technology. "We're hoping to take a next-generation prototype and put it on some of our artificial hands and begin experiments with amputees and able-bodied individuals who can help us troubleshoot and develop the grip control algorithms we're interested in," she says.

AN ASSISTIVE TECHNOLOGY

A soft, wearable device that mimics the muscles, tendons, and ligaments of the lower leg promises to aid in the rehabilitation of patients with ankle-foot disorders such as drop foot. The Active Soft Orthotic (ASO) is designed to provide active assistance to ankle motions for people with drop foot symptoms, which can commonly come from various neuromuscular disorders, such as cerebral palsy, multiple sclerosis (MS), amyotrophic lateral sclerosis (ALS), and stroke.

Yong-Lae Park (Figure 2), an assistant professor of robotics at Carnegie Mellon University, developed the device using soft plastics and composite materials, instead of a rigid exoskeleton, while working with collaborators at Harvard University, the University of Southern California, the Massachusetts Institute of Technology, and BioSensics, a Cambridge, Massachusetts, company that specializes in wearable sensors. The soft materials, combined with pneumatic artificial muscles (PAMs), lightweight sensors, and advanced control software, enable the robotic device to achieve more natural motions in the ankle.

Unlike previously developed devices featuring designs that primarily rely on traditional mechanical joints and rigid linkages, Park's system doesn't contain any rigid frame structures that could potentially prohibit natural degrees of joint motions. "Instead, our device employs a bioinspired tendon-ligament-skin architecture combined with pneumatically powered soft artificial muscles," says Park, who performed the research while a postdoctoral researcher at Harvard University's Wyss Institute for Biologically Inspired Engineering. "Another unique feature is the use of highly stretchable strain sensors, which we developed in the lab for measuring ankle joint-angle changes for feedback control of ankle motions."

Among the device's innovations are sensors featuring a touch-sensitive artificial skin—thin rubber sheets containing long microchannels filled with a liquid metal alloy. When the rubber sheets are stretched or pressed, the microchannels' shapes change, which in turn creates changes in the electrical resistance of the alloy. The sensors are positioned on the top and at the side of the ankle.

A significant drawback presented by the soft device is that it's more difficult to control than a rigid exoskeleton. The system demands highly sophisticated sensing to track the position of the ankle and foot and an intelligent plan for controlling foot motion. "We use three different types of sensors," Parks says. "Inertial measurement units (IMUs) and soft strain sensors detect the ankle joint-angle for ankle motion control, and a force sensitive resistor (FSR) foot-pressure sensor detects ground contact events." At the sensing stage, all of the sensors are sampled at 50 Hz.

Each sensor type has its own microcontroller unit (MCU) for independent sampling. After samples are acquired, different processing algorithms for each sensor type are applied in the signal processing stage. "We used an off-the-shelf microcontroller as the MCU for sensor sampling and signal processing," Park continues. "Also, each IMU has its own MCU to run a direct cosine-matrix (DCM)-based orientation estimation algorithm."

Each IMU board computes a three-dimensional (3-D) orientation vector and sends the value to the IMU master board using a standard serial protocol universal asynchronous receiver/transmitter (UART). Then, the IMU master MCU determines the joint-angle by computing the angle between the two orientation vectors. The pressure-sensing MCU produces an on-off output for each pressure sensor. The strain sensor MCU records the analog readings from the strain sensor analog circuit. Although the MCUs for the pressure and strain sensors currently do not perform significant processing, they are reserved for future implementation of more sophisticated filtering algorithms.

"We used this approach because it was simple but accurate enough for our purpose," Park says. "However, there is definitely room to improve in our method since we spent our time mainly on mechanical design and prototyping at this stage."

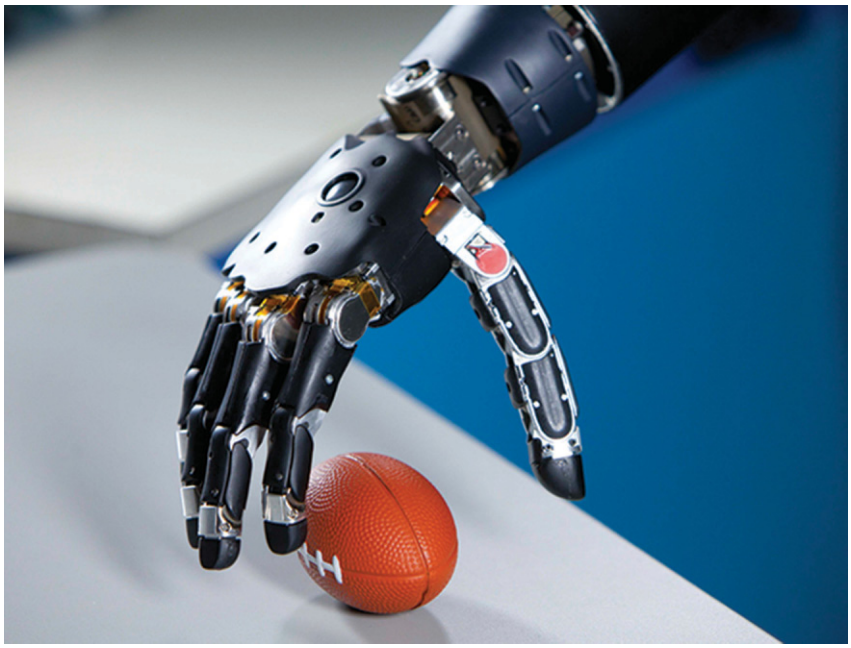
Park and his coresearchers envision a final system that not only increases rehabilitation efficiency by enhancing muscle usage but can also provide assistance during gait by increasing safety and stability. Park says similar approaches could be used to create rehabilitative devices for other joints of the body or even to create soft exoskeletons that increase the strength of the wearer.

The team's next step will be human subject testing. "We need to detect the intention of an individual muscle or a muscle group ... of the wearer so we [can] actuate and control the corresponding artificial muscle to assist the motion without disrupting the existing gait," Park notes. To achieve this goal, the researchers will need to read muscle activation and interpret the signal on both a quantitative and qualitative information basis. "I believe signal processing will be heavily involved and play a significant role in this step," he says.

Park noted that additional work will also be necessary to improve the device's wearability. This includes the development of artificial muscles that are less bulky than the commercially produced PAMs used in the current prototype.

A LIFELIKE MODULAR PROSTHETIC LIMB

Researchers at Baltimore's John Hopkins University have worked for nearly a decade on a prosthetic arm and hand that they hope will eventually bring lifelike functionality to people who have lost a limb to disease or trauma. The project's roots go back to 2005, when the U.S. Defense Advanced Research Projects Agency (DARPA) launched a program designed to produce a technology that would restore natural limb function to soldiers who had suffered amputations in the line of duty. Since prosthetic technology at the time was still rooted in concepts of previous decades, the potential advancement in restorative function that advanced prosthetic devices could provide promised to be revolutionary.



[FIG3] The Modular Prosthetic Limb developed by Johns Hopkins University researchers has an anthropomorphic form factor and appearance as well as humanlike strength and dexterity. (Photo courtesy of Johns Hopkins University.)

DARPA specified that devices created by program participants should look, feel, weigh, perform, and seamlessly integrate with a human user as if it were a natural limb. Following two phases of development, DARPA hoped that by the end of 2009 a fully functional and neutrally integrated prosthetic device would be developed that could accurately mimic the natural function of the human arm.

Michael McLoughlin, chief engineer of research and exploratory development at Johns Hopkins' Applied Physics Laboratory, notes that the DARPA initiative stimulated prosthetics development in what was then a generally neglected area. "The vast majority of amputees have lost a lower extremity; there's probably fewer than 100,000 or so people with upper extremity amputations," McLoughlin says. "So you have a situation where the technology to produce a replacement for the human arm is much more complex than for a leg, yet you don't have as many people who need it so you don't have the commercial pull to stimulate the technology's development."

The Modular Prosthetic Limb (Figure 3) developed by McLoughlin and his research team in response to the DARPA proposal features an anthropomorphic form factor and appearance as well as humanlike strength and dexterity. Other attributes include high-resolution tactile and position sensing capabilities and a neural interface for intuitive and natural closed-loop control.

The prosthesis contains more than 100 sensors. McLoughlin notes that many candidate technologies were prototyped and evaluated by using a custom-designed test bed and various standardized processes. Design constraints included integration, reliability, and manufacturability. At the arm's individual joints, sensors are used to measure angle, velocity, and torque. Additional sensors, located at the fingertips, measure force, vibration, fine point contact, and temperature/heat flux.

To reduce design complexity, the team needed to design motor controllers that were usable at multiple joint locations. A large motor controller (LMC) was developed to provide a circuit design that could be leveraged for use at the four joints of the upper arm and the three joints at the wrist. The LMC offers brushless direct current

(BLDC) motor commutation, sensor signal sampling, and communication with the central limb controller (LC) via a controller area network (CAN) bus.

The LMC was also designed to monitor local joint temperature, torque, position, current, and rotor position sensors for motor commutation. Custom schematic design and multilayer board fabrication allowed direct integration within the drive module. This approach allowed the drives to be designed as single integrated motor and controller packages, helping to shrink the overall mechanical profile and maximizing performance. Each LMC uses an advanced reduced instruction set computer (ARM)-based processor.

"It really comes down to a matter of functionality," McLoughlin says. "If you think about things like turning a doorknob, that's very difficult to do with a prosthesis." The researchers wanted to create a prosthesis capable of extremely fine dexterity and precision allowing users perform tasks ranging from the mundane, such as turning a doorknob up to and including playing a piano. "We want people to be able to do the very complex things with their fingers that those of us who have an arm and hand do naturally," McLoughlin says.

Studying brain signals is crucial to the team's research, since such data is essential for enabling natural control of the artificial limb. "If you think about moving your arm, or opening and closing your hand, there are areas of your brain that will become active," McLoughlin says. "You can actually see those areas of activity if you do an MRI."

The team searches for specific types of signals in different parts of the brain. "The signal processing typically involves things like pattern recognition, in which we look for patterns of neural activity," McLoughlin says. "We can then begin to use pattern recognition techniques to interpret what the user's intent was."

The researchers need to work quickly. "It's all done in real time," he says. "You filter it, process it, pull out the information, and then convert that data into motor commands."

McLoughlin says he's always amazed by the brain's flexibility and adaptability. "We're looking for very specific structures in the signal, and we come up with a very specific model of the signals we're looking at," he says. "I think it's going to open up a whole new realm of possibilities for assistive devices, particularly for the elderly or people with mobility problems who will be able to use machines in ways that are well beyond what we can do now."

One of the challenges facing the team is finding a way of driving down the sophisticated limb's cost. "Right now this is a research tool," McLoughlin says. "We've had ten or 12 different patients utilize the limb with great success, and we have to look at getting it down to a cost point where it's affordable," he says.

AUTHOR

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

Ron Schneiderman

Accuracy, Apps Advance Speech Recognition

Technical breakthroughs in speech recognition have been hard to come by, but the technology continues to improve in accuracy and natural language understanding and find its way into a broad range of enterprise and commercial platforms that include health care, e-commerce, telecommunications, and other vertical markets.

In this third in a series of Q&A interviews for *IEEE Signal Processing Magazine (SPM)*, we talked to Li Deng, the principal researcher and research manager of the Deep Learning Technology Center at Microsoft Research, and Vlad Sejnoha, the chief technology officer at Nuance Communications, about current activities and future developments in speech recognition, text-to-speech (TTS), speech-to-speech translation, and related applications.

IEEE SPM: Accuracy has been an issue in speech technology since its emergence out of Bell Labs in the 1950s. How has it improved?

Vlad Sejnoha: We actually have been improving it quite quickly over the years. We have improved the error rate by a consistent amount and there seems to be no end to this. Each year, it's through a different combination of new algorithms, more data, more computation. Different mixes of that. So, when we talk about speech recognition accuracy, it really is a moving target. I think we have passed a magical threshold of usability that means that you can pick up a device today and speak to it and expect to be understood.

In recent years, the focus has been on deep learning. That is the sort of algorithmic underpinning of why we are continuing to improve. In a few years, it might be something else. So, it's a very rapid and dynamically evolving process.

Li Deng: Progress was relatively slow from 1989 to 2009 compared with the last few years after deep learning made inroads into speech recognition. The introduction of deep learning has been one of the major breakthroughs, mostly in the form of deep neural networks. But

Digital Object Identifier 10.1109/MSP.2014.2358717
Date of publication: 5 December 2014

1053-5888/15©2015IEEE

there's expected to be work that may go beyond neural networks.

The concept of deep learning applies in terms of being able to absorb more data and you can actually impart relevant domain knowledge of speech into the system in a hierarchical manner. All of this new technology, especially from the machine-learning community that is heavily overlapping with our speech/language and signal processing communities, has had a tremendous impact in our field. That's why we're quite hopeful that the accuracy will keep improving, even after neural networks reach their potential limit sometime in the future.

Deep learning started its success drastically raising speech recognition accuracy with deep neural networks [around 2010, soon after Microsoft researchers collaborated with Prof. Geoff Hinton in Redmond, Washington]. In the future, I expect that deep neural network will integrate with other forms of deep models in elegant and theoretically appealing ways to achieve better success, with the new capability of not only absorbing big data but more importantly "big knowledge" that would include semantic knowledge in the top-down fashion.

Sejnoha: One thing that should be pointed out, I think, is the importance of good signal acquisition. If you have the best speech recognition system in the world, if you're using it in a very difficult environment, with a mobile device with a microphone that might be moving and not properly positioned to you, it will not do very well. So, we're finding that these fundamental modeling improvements, like deep learning, are being coupled with more and more sophisticated ways of capturing a signal, and that includes things like using multiple microphones that can be configured into a steerable beam that can track the speaker, in combination with voice biometrics where you actually know who is talking and you can tease out the desired signal from interfering signals. And you can use this with other sensors like audio-visual recognition where you can actually use camera feeds to lock in on the speaker. That's becoming extremely important, but improvements there are equal to multiple generations of modeling improvements.

IEEE SPM: What is deep learning?

Sejnoha: It's an approach that tries to model or make decisions about the nature of input signals by organizing layers of very simple processing units, which are very loosely modeled after our understanding of how neurons work. Each processing unit has a number of inputs, and those inputs are weighted. They come from the outputs of other neurons. These neurons are processing nodes. Some of these are weighted inputs and then pass them through in non-linear function that basically says on or off, or some degree of that, and then passes that on to higher layers.

It can model very complex decision spaces. In the past, these were very difficult to train and got stuck in what we call *local optima*. There have been a number of breakthroughs in the training in recent years that help these layered networks reach more global or overall optima. It's called *compositionality*; that you can have relatively few processing units and they can explain very high dimensional spaces. Ever since they became more trainable, this has been applied to a wide variety of problems—acoustic modeling, language modeling, and assigning meaning to patterns.

Interestingly, you can use them to concurrently learn or optimize multiple parts of the recognition process. Traditionally, we would build a so-called feature extractor that takes the audio signal and maps it usually into spectral space that we think is more amenable to further processing. It turns out that these deep neural nets can simultaneously learn a more optimal set of features that also make decisions on the results. And that's exciting, but I want to caution about thinking that the world going forward is all going to be about machine learning, or learning from examples and patterns.

A lot of current cutting-edge work is about how to combine machine learning with techniques that encapsulate our existing knowledge of the world, either through rules or grammar, or explicit knowledge bases where you describe object, concept, and relationships.

IEEE SPM: What's the role of artificial intelligence (AI) in improving the accuracy of speech recognition? Is it advancing the technology?

Sejnoha: AI is a big term, but some aspects of AI are already manifested in today's virtual assistants. A simple example would be a user asking for a restaurant that serves good spaghetti, and it turns out that the back-end services to which you're connected don't understand that because they only understand cuisine types. You would actually take this input user request, and, using explicit reasoning on a knowledge base that relates specific dishes to specific cuisines, it would learn that spaghetti is a form of Italian cuisine. But that's a trivial example. There are a lot more sophisticated ways of relating an input concept language with a backup concept language and deciding what makes sense and doesn't make sense.

But there could be a lot of subtasks in a request that a virtual assistant should be able to do, but there are lots of contingencies. Being able to specify tasks by expressing goals at a high level with automatic back-off strategies versus prescripting every possible interaction is really important in AI.

IEEE SPM: To what extent is digital signal processing (DSP) playing a role in the development of speech technology?

Deng: The community, including many speech recognition, understanding, and machine-learning researchers, are part of our [IEEE Signal Processing] Society, so we are actually thinking about changing the Society's name to the IEEE Signal and Information Processing Society. This would better describe the recent activities of this community [of the Society's members]. Many big companies, and I am thinking about Microsoft, Apple, Google, Baidu, etc., have researchers and engineers working on speech technology problems much more complex than the DSP topics you would see in Oppenheimer's book. We have moved well beyond traditional DSP.

As for deep learning, the people who brought that into large-scale speech technology applications are mainly from our [IEEE] Signal Processing Society. I recently gave a long lecture at the International Conference on Machine Learning, and I delivered the keynote at the

Interspeech Conference held in Singapore in mid-September, reflecting on part of that history with emphasis on the industry–academic collaboration and on what future direction that history points to. A lot of DSP techniques, such as short-time Fourier transforms, cepstral analysis, and linear prediction, which used to be standard analysis as front ends for speech recognition systems, are now becoming mere initialization of low layers of the full deep-learning system, subject to much more important step of end-to-end learning by what is known as back propagation beyond DSP.

However, one important concept of DSP, convolution, has been playing a crucial role in modern deep learning systems. They are called *deep convolutional neural networks*, very popular in image recognition and recently also gaining popularity in speech recognition.

IEEE SPM: *How much progress has been made in implementing TTS, especially in making it more natural?*

Sejnoha: That's a topic that is near and dear to us. The TTS, or speech synthesis field, has moved through some phases through the years. It started out by building what we call some model-based approaches, some mathematical models that are very compact, and the use of algorithms to express how particular text should be mapped–ultimately into a waveform. But they sounded very robotic.

There was a breakthrough in the late 1990s when the first commercial systems started coming out. People discovered that by extracting snippets of real speech and forming them into structured data bases you could, on the fly, concatenate, or glue together, the appropriate segments and they sounded far more natural. The problem with those systems is that they were big. There were lots of different segments that were needed, and it was very difficult to manipulate pitch, duration, and loudness, but it really carried a lot of emotion and naturalness in human speech.

So a lot of the work now is how to build hybrid systems, just like hybrid systems in understanding and combine machine learning and some explicit knowledge. Building TTS that have the quality of the concatenated system, but the ability to manipulate pitch and loudness and volume—all of the prosodic signals based on understanding of the text. So, what we do is take text and apply natural understanding to what is being said and what it means, and use a combination of prerecorded segments and models to try to generate expressive speech synthesis. Deep learning plays a role there as well.

Deng: Deep learning has also been making an inroad into speech synthesis or TTS research since last year. At ICASSP 2013, there were four nice papers on this topic, from different angles and for different aspects of the synthesis problem. They demonstrated more natural subjective speech sound's quality produced by deep learning systems than the previous state-of-the-art, Gaussian-HMM-based statistical methods. More research papers have come out since then. In a sense, it is very intuitive to adopt an original, generative version of deep learning approaches, called the deep believe network, which is quite different from the deep neural network, to deal with speech generation or synthesis problems.

IEEE SPM: *Are there particular challenges at this point in deploying speech technology globally given the need to support many languages and with a high degree of accuracy?*

Deng: You want to have voice systems perform well in noisy environments. These include the conditions where the voice intended to be recognized are mixed with other speakers' voices, such as when playing Xbox or Kinect games with voice control. As was demoed in May this year, Skype translator will be able to perform real-time speech-to-speech translation. Under the conditions where there is no close-talking microphone, noise robustness, especially the robustness against other speakers' voice, in speech recognition component of the system is very important. Human listeners can use attention to focus on the intended speaker, but so far computer systems cannot simulate such ability easily. Deep learning is moving toward solving such difficult problems, with preliminary promising results already seen in the literature.

Before the rise of deep learning, multilingual speech recognition was very difficult in economic terms due to the need to collect data and design dictionaries from many languages. Deep neural networks have drastically reduced this challenge, thanks to the "transfer learning" capability where the upper hidden layers in the deep networks are shown to represent more abstract acoustic features universal across different language. This capability is made possible because acoustic properties of speech, no matter which language it belongs, are shared across languages since they are all generated by the highly constrained human vocal tract, plus the rest of the speech production system. Only deep learning systems can effectively take advantage of such constraints, not the previous systems without hierarchical feature representations.

For many speech recognition applications that are linked closely to downstream processing, semantic understanding of the recognition output and of the end tasks and the final actions taken by the overall system are the final goal. One particular technical challenge here is how to effectively represent semantics and the backend application-domain knowledge. Recent advances in deep learning for natural language processing have provided a very interesting approach where any semantic linguistic entity and simple relation in the knowledge source can be mapped into a continuous-valued vector, called *embedding*. Embedding has been shown to be quite effective for a word, a phrase, a sentence, a paragraph, or even a whole document. These embedded linguistic units can also be used to represent the output of a speech recognizer. Thus, the designs of downstream text processing and speech recognition systems are intimately connected and can be jointly optimized.

Despite such progress, however, semantic representations for more advanced tasks that would require structured representations and complex relations may not be adequately accomplished with vector

(continued on page 125)

2015 Class of Distinguished Lecturers and Technical Field Award Recipients

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

Candidates for the Distinguished Lecturer Program are solicited from the Society technical committees, editorial boards, Chapters, and other boards and committees by the Awards Board. The Awards Board vets the nominations, and the Board of Governors approves the final selection. Distinguished Lecturers are appointed for a term of two calendar years. Distinguished Lecturers named for 2015 are as follows.

Hynek Hermansky received his doctor of engineering, electrical engineering, University of Tokyo (1979–1983); doctoral graduate studies, Technical University Brno, Czech Republic (1973–1978); and Ing. (M.S. degree equivalent), Technical University Brno, Czech Republic (1967–1972). His appointments include Julian S. Smith Professor in Electrical Engineering, Department of Electrical and Computer Engineering, and director of the Center for Language and Speech Processing, Johns Hopkins University, Baltimore, Maryland (2012–present); research professor (on leave of absence), Brno University of Technology, Czech Republic (2000–present); external fellow, International Computer Science Institute, Berkeley, California (1995–present); professor, Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland (2008–2012); director of research, IDIAP Martigny, Switzerland (2003–2008); professor and director, Center for Information Technology, Oregon Graduate Institute School of Oregon Health and Science University, Portland (1993–2003); senior scientist and fellow, International Computer Science Institute, Berkeley, California (1999–2003); senior member of technical staff, US West Advanced Technologies, Boulder, Colorado (1988–1993); research engineer, Panasonic Technologies, Santa Barbara, California (1983–1988); research fellow, University of Tokyo, Japan (1978–1983); assistant professor, Technical University Brno, Czech Republic (1975–1978); and member of research staff, Technical University Brno, Czech Republic (1972–1975).

Dr. Hermansky is an IEEE Fellow. He was member of the organizing committee in charge of invited lectures, International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (2011); chair of the technical committee, ICASSP (1998); general chair, Automatic Speech Recognition and Understanding Workshop (2013); associate editor, *IEEE Transactions on Speech and Audio Processing* (2000–2001); fellow, International Speech Communication Association; elected member, Board of Directors, International Speech Communication Association (ISCA), (2000–2004, 2013–present); Medal for Scientific Achievements, ISCA (2013); Distinguished Lecturer, ISCA (2013–2014); editorial board member, *Speech Communication*.

He has over 250 publications and has reviewed technical journals, professional conference proceedings, and invited book chapters. He has been granted ten patents. His lecture topics include dealing with unknown unknowns in speech, data-guided features in recognition of speech, information extraction from temporal dynamics of speech, and multistream recognition of speech.

Visa Koivunen received his D.Sc. degree in electrical engineering with honors from the Department of Electrical Engineering, University of Oulu, Finland. He received the Primus Doctor (Best Graduate) Award among doctoral graduates in 1989–1994. He is a member of Eta Kappa Nu. He was a visiting researcher at the University of Pennsylvania, United States, from 1992 to 1995; faculty, Tampere University of Technology, Finland, from 1997 to 1999; and full professor of signal processing, Aalto University (formerly known as Helsinki University of Technology), Finland, since 1999. He holds the academy professor position (distinguished professor nominated by the Academy of Finland). He was a principal investigator with SMARAD Center of Excellence in Research nominated by the Academy of Finland in 2002–2013. He was also an adjunct full professor, University of Pennsylvania, United States (2003–2006); visiting fellow, Princeton University, New Jersey, United States (2007, 2013–2014); and part-time visiting fellow, Nokia Research Center (2006–2012).

He holds six U.S. patents and has coauthored papers that received the Best Paper

Digital Object Identifier 10.1109/MSP.2014.2353431
Date of publication: 5 December 2014

Award at IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (2005); the European Signal Processing Conference (2006); the Conference on Antennas and Propagation (2006); and the International Conference on Advances in Cognitive Radio (2012). He was awarded the IEEE SPS Best Paper Award in 2007 (with J. Eriksson). He served as associate editor, *IEEE Signal Processing Letters* (2002–2004); *IEEE Transactions on Signal Processing* (2011–2012); *Signal Processing,* and *Journal of Wireless Communication and Networking*; and guest editor, *IEEE Journal of Selected Topics in Signal Processing* (Special Issue on Smart Grids). He is an editorial board member, *IEEE Signal Processing Magazine* (2012–present); member, IEEE Signal Processing for Communications and Networking Technical Committee (2005–2010); member, Sensor Array and Multichannel Technical Committee (2009–present); member, SPS Industrial Relations Committee (2011–2013); general chair, IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) (2007); technical program chair, IEEE SPAWC (2015); array processing track chair, ASILOMAR (2014). He has given tutorials at both IEEE ICASSP 2007 and 2013.

His research interests include statistical, communications, sensor array, and multichannel signal processing. His lecture topics include "Optimal Array Signal Processing in the Face of Nonidealities," "Optimization Under Unitary Matrix Constraints: Differential Geometry Approach," "Robust Estimators for Complex-Valued Multichannel Data," "Spectrum Exploration and Exploitation: Joint Optimization of Identifying and Accessing Idle Spectrum," "Analyzing Large-Scale Data: Robust and Sparse Signal Processing," and "Complex Random Vectors and Noncircularity: Statistical Inference and Parameter Estimation."

Hamid Krim received his degrees in electrical engineering. As a member of technical staff at AT&T Bell Labs, he has worked in the area of telephony and digital communication systems/subsystems. In 1991, he became a National Science Foundation (NSF) postdoctoral scholar at the Foreign Centers of Excellence (LSS Supelec/University of Orsay, Paris, France). In 1992, he joined the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, as a research scientist performing/supervising research in his area of interest. In 1998, he joined the Electrical and Computer Engineering Department at North Carolina State University, Raleigh, where he is currently a professor. His editorial activities include editorial board member, *IEEE Transactions on Signal Processing* (2002–2004) and *IEEE Signal Processing Magazine* (2014).

Dr. Krim is an IEEE Fellow and was a fellow, Japanese Foundation for the Advancement of Research in Science and Engineering at the University of Tokyo, Japan. He is a member of SIAM and of Sigma Xi. He is an original contributor and now an affiliate of the Center for Imaging Science, sponsored by the U.S. Army. He is a recipient of the NSF Career Young Investigator Award.

Dr. Krim's research interests are in statistical signal processing and mathematical modeling with an emphasis on applications. He has been particularly interested in introducing geometric and topological tools to statistical signal processing problems and applications. His research has primarily centered on estimation theoretic problems and modeling. Dr. Krim has published extensively on these areas with over 5,000 citations to date.

His lecture topics include "Shape Analysis and Modeling in Video Applications: Activity Analysis," "Convexity, Sparsity, Nullity and All That in Machine Leaning," and "Sensor and Social Networks: A Case for Topological Data Analysis."

Jean-Christophe Olivo-Marin is the head of the Bioimage Analysis Unit and director of the Center for Innovation and Technological Research at Institut Pasteur, Paris. He has chaired the Cell Biology and Infection Department from 2010 to 2014 and was a cofounder of the the Institut Pasteur Korea, Seoul, South Korea, where he held a joint appointment as a chief technology officer (2004–2005). Prior to that, he was a staff scientist at the European Molecular Biology Laboratory, Heidelberg, from 1990 to 1998. He received the Ph.D. degree (1989) and Habilitation (1998) degrees in optics and signal processing from the Institut d'Optique Théorique et Appliquée, University of Paris-Orsay, France.

Dr. Olivo-Marin is an IEEE Fellow. He is the steering committee chair, IEEE International Symposium on Biomedical Imaging (2014–2016); past chair, IEEE SPS Bioimaging and Signal Processing Technical Committee (2009–2011); member, IEEE SPS Conference Board (2010–2011); senior area editor, *IEEE Signal Processing Letters* (2013–2014); and editorial board member, *Medical Image Analysis* and *BMC Bioinformatics*. He has organized several special sessions dedicated to biological imaging at international biomedical conferences including the 2002 International European Light Microscopy Initiative Meeting and Workshop on Advanced Light Microscopy (ELMI), 2003 Extracorporeal Life Support Organization, IEEE ISBI (2004), IEEE ICASSP (2006 and 2011), SPIE Wavelets (2009 and 2013), EMBO (2011), SPIE Wavelets (2009 and 2013), EMBO (2011), and he was the general chair, IEEE ISBI (2008).

Dr. Olivo-Marin's research interests are in image analysis of multidimensional microscopy images, computer vision and motion analysis for cellular dynamics, and mathematical approaches for biological imaging. His lecture topics include "Particle Tracking in Biological Imaging," "Cell Shape and Motility Analysis," and "Quantitative Bioimage Analysis."

Min Wu received the B.E. degree in electrical engineering–automation and the B.A. degree in economics from Tsinghua University, Beijing, China, in 1996 (both with the highest honors), and the M.A. and Ph.D. degrees in electrical engineering from Princeton University, New Jersey, in 1998 and 2001, respectively. She was with NEC Research Institute and Signafy, Inc. in 1998, and with the Media Security Group, Panasonic Information and Networking Laboratories in 1999. Since fall 2001, she has been on the faculty of the Electrical and Computer Engineering Department and the Institute of Advanced Computer Studies at University of Maryland, College Park, where she is now an ADVANCE professor and University Distinguished Scholar-Teacher.

She was a visiting associate professor of Stanford University in 2007–2008. Dr. Wu leads the Media and Security Team at the University of Maryland with main research interests on information security and forensics and multimedia signal processing.

She has been an IEEE Fellow since 2010 "for contributions to multimedia security and forensics." She received an NSF CAREER Award (2002); TR100 Young Innovator Award from *MIT Technology Review Magazine* (2004); ONR Young Investigator Award (2005); Computer World "40 Under 40" IT Innovator Award (2007); IEEE Mac Van Valkenburg Early Career Teaching Award (2009); and the Daily Record Innovator of the Year Award (2012). She has served as vice president–finance, IEEE SPS (2010–2012); chair, IEEE Information Forensics and Security Technical Committee (2012–2013); technical program cochair, ICIP (2013); founding chief editor, IEEE SigPort Initiative (2013–2014); and chair, ChinaSIP Steering Committee as an SPS's outreach effort in China, with a successful launch in Beijing in 2013 and Xi'an in 2014, and the third conference planning under way for Chengdu in 2015. She is the current editor-in-chief of *IEEE Signal Processing Magazine* (2015–2017).

Dr. Wu's IEEE and SPS volunteer activities include: corresponding member, IEEE TAB Finance Committee (2012); editorial board member, *IEEE Signal Processing Magazine* and *IEEE Journal of Selected Topics in Signal Processing* (2012–present); area editor, *IEEE Signal Processing Magazine*; leading the creation and editing of the monthly Inside Signal Processing e-Newsletter (2007–2010); associate editor, *IEEE Transactions on Image Processing* (2009–2011), *IEEE Transactions on Information Forensics and Security* (2008–2011), and *IEEE Signal Processing Letters* (2005–2007); member, Image, Video, and Multidimensional Signal Processing Technical Committee (2007–2012), Multimedia Signal Processing Technical Committee (2002–2005 and 2007–2009), Information Forensics and Security Technical Committee (since 2008), and multimedia-related technical committees of the IEEE Communications Society and IEEE Circuits and Systems Society; general cochair, ICIP (2017); technical program cochair, ICIP (2013); finance chair, ICASSP (2007); publicity chair, ICME (2003); and founding chair, IEEE Signal Processing Washington Chapter.

Dr. Wu's lecture topics include "Seeing the Invisibles: A Backstage Tour of Information Forensics," "Exploring Power Network Signature in Multimedia: From Information Forensics to Digital Humanity," and "Hands-On Education on Multimedia and Security: From Kindergarteners to Undergraduate to Business and English Majors in Continuous Education."

2015 TECHNICAL FIELD AWARD RECIPIENTS

Each year, the IEEE recognizes individuals who have made outstanding contributions or exercised leadership within IEEE-designated technical areas. The IEEE SPS is honored to announce three of its members as recipients of the 2015 IEEE Technical Field Awards.

- IEEE James L. Flanagan Speech and Audio Processing Award: presented to Stephen John Young for "pioneering contributions to the theory and practice of automatic speech recognition and statistical spoken dialogue systems"
- IEEE Fourier Award for Signal Processing: presented to Georgios B. Giannakis for "contributions to the theory and practice of statistical signal processing and its applications to wireless communications"
- IEEE Donald O. Pederson Award in Solid-State Circuits: presented to Robert Whitlock Adams for "contributions to noise-shaping data converter circuits, digital signal processing, and log-domain analog filters."

Congratulations to all of the recipients! The full list of 2015 IEEE Technical Field Awards recipients can be found in [1].

REFERENCE

[1] [Online]. Available: http://www.ieee.org/about/awards/news/2015_ieee_tfa_recipients_and_citations_list.pdf

[SP]

[from the GUEST EDITORS]

Arrate Muñoz-Barrutia,
Jelena Kovačević, Michal Kozubek,
Erik Meijering, and Braham Parvin

Quantitative Bioimaging: Signal Processing in Light Microscopy

Microscopy has historically been an observational technique. In recent years, however, the development of automated microscopes, digital sensing technologies, and novel labeling probes have turned microscopy into a predominantly quantitative technique. In this context, the management and analysis of automatically extracted information calls for the involvement of signal and image processing experts to provide technically sound, quantitative answers to biological questions. This is especially relevant today, due to the widespread use of time-lapse video microscopy, high-throughput imaging, and the development of novel superresolution microscopy techniques. The complexity and size of the multidimensional and often multimodal data produced by those microscopy techniques requires the use of robust computational methods encapsulated in advanced bioimage informatics tools.

Our motivation for publishing this special issue of *IEEE Signal Processing Magazine* is to stimulate the interaction among researchers from the biological, optical, computer science, and signal processing communities by 1) presenting cutting-edge signal processing research in quantitative bioimaging and 2) bringing the vast scope of ongoing open problems and novel applications to the attention of the signal processing community. As we hope to show in this issue, there are many high-impact signal processing challenges at the intersection of quantitative bioimaging and integrative biology where signal processing experts can make a mark. These challenges are described in the context of the imaging modality used, the probes and sensors employed for image acquisition, and the final targeted applications (i.e., development studies, disease diagnosis and prognosis, drug discovery). When possible, works following the reproducible research (http://reproducibleresearch.net) philosophy are highlighted.

The interest that the signal processing community has in quantitative bioimaging is evident from the increasing number of papers submitted on this topic to signal processing-oriented publications, workshops, and conferences. Dedicated issues on molecular and cellular bioimaging were previously published in *IEEE Transactions on Image Processing* [1] and *IEEE Signal Processing Magazine* [2]. The rapid evolution of the field justified the interest of devoting a new special issue to examine all these developments from a signal processing perspective. Furthermore, in the last few years, a number of related "scientific challenges" have been held either as stand-alone or as part of image processing conferences. These activities are very relevant for the community since they facilitate the comparison of various algorithms for a given generic task (e.g., deconvolution, single particle localization, particle tracking, cell tracking) using a normalized framework consisting of annotated data and common evaluation metrics. In terms of funding programs, the importance of quantitative bioimaging research is also apparent. In this respect, the European Strategy Forum on Research Infrastructures roadmap contains a pertinent project, "Euro-Bioimaging," with a dedicated work package on data storage and analysis. The U.S. counterparts of the European initiative are the "Continued Development and Maintenance of Software" program run by the U.S. National Institutes of Health (NIH), since 2002, and the recently announced "Software Infrastructure for Sustained Innovation" program that will be run by the U.S. National Science Foundation (NSF). Apart from those, a number of consortia addressing extraordinarily relevant problems are being or will be funded by the European Union (under the Seventh Framework and the recently opened Horizon 2020 Programmes) and the NIH and NSF. All of these provide ample proof that this issue's theme is timely, and we hope that it offers barrier-breaking material from which the readership will benefit.

From a systems biology perspective, the cell is the principal element of information integration. Profiling cellular responses and clonal organization in its spatiotemporal context are important endpoints for unraveling molecular mechanisms of diseased tissue (e.g., bacterial invasion, cancer). The first article, "Toward a Morphodynamic Model of the Cell," by Ortiz-de-Solórzano et al., is a review of relevant signal processing aspects from the detection of cellular components to the description of the morphodynamics of the entire cell in relation to its extracellular environment. A survey of ongoing efforts to create a credible model of cell behavior is also an integral part of the manuscript. Significantly related, Dufour et al. in "Signal Processing Challenges in Quantitative 3-D Cell Morphology" give an overview of the problems, solutions, and remaining challenges in deciphering the morphology of living cells via computerized approaches, with a particular focus on shape description frameworks and their exploitation, using machine-learning techniques. In their technical article, "Snakes on a Plane," Delgado-Gonzalo et al. present an extended and inclusive taxonomy of different variants of two-dimensional active contours (also known as *snakes*) for the

Digital Object Identifier 10.1109/MSP.2014.2359691
Date of publication: 5 December 2014

segmentation of cells and other biological entities. The authors also lay out general design principles that can help to create new parametric snakes adjusted to different imaging modalities.

Newly developed superresolution microscopy techniques break Abbe's diffraction limit, providing lateral resolution values as high as 10 nm, far below the 250 nm of conventional microscopy. Those techniques, through the visualization of molecular machinery, are helping to answer biological questions about the mechanisms of cellular behavior regulation. Localization microscopy is one of these superresolution techniques. In localization microscopy, the fluorescent labels are photochemically manipulated to switch "on" and "off" stochastically, such that at each instant in time only a sparse subset of all molecules is in the "on" state in which they fluoresce. Assembling the localization data obtained from all frames into the final superresolution image reveals previously hidden details. In "Image Processing and Analysis for Single-Molecule Localization Microscopy," Rieger et al. describe the image processing and workflow involved, from raw camera frames to the visualization and quantitative analysis of the reconstructed superresolution image. Single-molecule approaches place stringent demands on experimental and algorithmic tools due to the low signal levels and the presence of significant extraneous noise sources. This necessitates the use of advanced statistical signal and image processing techniques for the design and analysis of single-molecule experiments. In their article, "Quantitative Aspects of Single-Molecule Microscopy," Ober et al. address this issue and discuss the resolvability of single-molecule localization from an information-theoretic perspective.

The use of time-lapse video microscopy to capture the spatiotemporal dynamics of many biological experiments has significantly increased. The complexity of those experiments is driving continued advances in the incipient field of bioimage informatics [3]. Registration, segmentation, and annotation of microscopy images and respective biological objects (e.g., cells) are distinct challenges often encountered in this field. In "3-D Registration of Biological Images and Models," Qu et al. discuss several studies in widely used model systems such as fruit fly, zebrafish, or *C. elegans* to show how registration methods help solve challenging segmentation and annotation problems for three-dimensional cellular images.

A classical light microscopy application in clinical practice is histopathology. Clinicians evaluate histological preparations for the patient's diagnosis, estimation of prognosis, personalized therapy planning and, in a research context, biomarkers discovery. Tissue processing for histology is increasingly automated, and digitalization using modern computer-driven microscopes or slide scanners is extremely time effective and generates an extensive volume of data. Therefore, as described by McCann et al. in "Automated Histology Analysis," there is a niche for image analysis methods that can automate prohibitively time-consuming tasks for human evaluation. Moreover, as concluded by the authors, a close collaboration and extensive work with pathologists is required for the developed applications to reach an important impact in clinical practice.

The final article, "Optical and Optoacoustic Model-Based Tomography," by Mohajerani et al., describes optical imaging techniques that reach beyond microscopy depths, bringing unique visualization of intact small animals or human tissues in vivo. Light propagation in tissue defines complex nonlinear inversion problems in both optical and optoacoustic model-based tomography. Therefore, the robust localization and quantification of the optical probes is a nontrivial problem opening up a clear opportunity for the signal processing community.

We would like to express our appreciation to the editorial board and staff of *IEEE Signal Processing Magazine* (particularly Special Issue Area Editor Fulvio Gini) for encouraging, reviewing, and facilitating the process of editing this issue. It would not have been possible without the high-quality feedback received from the conscientious reviewers whom we wish to thank for their volunteer efforts and timely responses. We sincerely hope you enjoy reading this issue as much as we enjoyed putting it together.

REFERENCES

[1] R. F. Murphy, E. Meijering, and G. Danuser, "Molecular and cellular bioimaging [Guest editorial]," *IEEE Trans. Image Processing*, vol. 14, no. 9, pp. 1233–1236, Sept. 2005.

[2] J. Kovačević and R. F. Murphy, "Molecular and cellular bioimaging [Guest editorial]," *IEEE Signal Processing Mag.*, vol. 23, no. 3, p. 19, May 2006.

[3] G. Myers, "Why bioimage informatics matters," *Nat. Methods*, vol. 9, no. 7, pp. 659–660, June 2012.

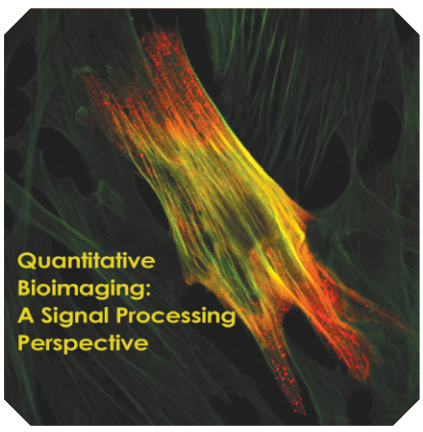
[SP]

Digital Object Identifier 10.1109/MSP.2014.2368013

Carlos Ortiz-de-Solórzano, Arrate Muñoz-Barrutia, Erik Meijering, and Michal Kozubek

Toward a Morphodynamic Model of the Cell

Signal processing for cell modeling



From a systems biology perspective, the cell is the principal element of information integration. Therefore, understanding the cell in its spatiotemporal context is the key to unraveling many of the still unknown mechanisms of life and disease. This article reviews image processing aspects relevant to the quantification of cell morphology and dynamics. We cover both acquisition (hardware) and analysis (software) related issues, in a multiscale fashion, from the detection of cellular components to the description of the entire cell in relation to its extracellular environment. We then describe ongoing efforts to integrate all this vast and diverse information along with data about the biomechanics of the cell to create a credible model of cell morphology and behavior.

INTRODUCTION

Systems biology [1] is a multilevel approach to the study of biological phenomena that integrates structural and functional information at different levels of spatial (molecular, cellular, tissular, organismal) and temporal resolution. Although many noteworthy ongoing efforts aim at computationally describing the structure, function, and even the development of entire organs [2] and simple organisms [3], the cell remains the principal element of information integration and is the key to the design of higher-order models. The cells emanate signals that collectively determine the fate and evolution of organs and, within the cells, signals are directed that elicit inner mechanisms of protein production, replication, differentiation, and death. Understanding how the cell senses, reacts to, and produces these regulatory signals is the key to explaining the principles of life and disease. This is a daunting task that requires the study of the cell from many different perspectives (morphological, biochemical, mechanical, electrical) accounting for both the temporal and spatial dimensions. Accordingly, numerous efforts today are directed toward the creation of multidimensional morphodynamic models of the cell. Feeding into these models are technologies (hardware) and methods (software) that produce quantitative visual information. The evolution of these methods and technologies poses continuous challenges to the signal processing community. In this article, we review the state of the art of computational and signal processing aspects involved in:

- the development of advanced live cell imaging modalities
- the dynamic tracking of cells and subcellular components
- the estimation of forces exerted between the cell and its local environment
- the integration of "visual" information into credible models of cell behavior.

Digital Object Identifier 10.1109/MSP.2014.2358263
Date of publication: 5 December 2014

1053-5888/15©2015IEEE

A brief historical perspective and a discussion of the state of the art of all these fields are presented, along with new challenges that require the involvement of the signal processing community.

IMAGING CELL BEHAVIOR

The first studies of the behavior of living cells date back about 340 years, when the Dutch draper Antoni van Leeuwenhoek decided to turn his interest in lens making to visualizing more interesting objects than his merchandise. In the mid-1670s, possibly inspired by the 1665 groundbreaking publication of Robert Hooke's discovery of plant cells, he was the first to observe micro-organisms, or *little animals* (animalcules) as he called them, in a drop of lake water. The magnification factor of van Leeuwenhoek's single-lens microscopes amounted to a stunning 300×, and the optical resolution of his lenses was already around 1 micron. It took until the 19th century before compound microscopes were developed that surpassed the quality of his microscopes and reached the diffraction limit—roughly half the wavelength of the light—discovered by Ernst Abbe in 1873.

Cells by themselves are fairly transparent and cannot be studied in detail using conventional light microscopy. One trick to improve optical contrast without having to explicitly stain cells is to exploit the fact that when light travels through a medium, it undergoes amplitude and phase changes that are dependent on the properties of the medium. While the human eye is sensitive only to amplitude variations, phase shifts may carry important information about the medium, and can be made visible by conversion to changes in brightness using special optical components. Phase contrast (PC) microscopy [4] was invented in the 1930s by Frits Zernike, which earned him the Nobel Prize in physics in 1953. Unfortunately, the use of this technique is restricted to very thin specimen preparations, and the resulting images suffer from halo artifacts. Another technique invented by Georges Nomarski in the 1950s, is to exploit the interference obtained when recombining two orthogonally polarized and slightly displaced light components after traveling through the specimen. Differential interference contrast (DIC) microscopy [4] yields superior resolution compared to PC microscopy and has excellent optical sectioning capability. However, the effectiveness of DIC is reduced by the specimen's reaction to polarized light. Moreover, the resulting images show typical pseudo-three-dimensional (3-D) artifacts that can be mistakenly interpreted as topographical cell features. Several of the limitations of both PC and DIC microscopy can be avoided by the use of Hoffman modulation contrast microscopy [4], developed by Robert Hoffman in 1975.

These optical contrasting techniques are used particularly in studies that do not require quantification of intracellular components, but that rather aim to characterize the morphodynamics of individual cells or the aggregate migratory behavior of groups of cells. For the study of dynamic processes within a living cell, it is necessary to specifically label the intracellular objects of interest. This has become possible at large by the discovery (1962), gene sequencing and cloning (1992), and expression (1994) of the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* [5]. GFP-labeling enabled the visualization of very specific targets within living cells and opened the door to studying the location and function of intracellular components with unprecedented sensitivity and specificity. This caused a true paradigm shift in biological experimentation [6] to the extent that the inventors of the technique, Osamu Shimomura, Martin Chalfie, and Roger Tsien, were awarded the Nobel Prize in chemistry in 2008. During the 1990s and 2000s, many derivatives of GFP were developed with their own characteristic excitation and emission spectra, which further extended the toolbox of fluorescent labeling [5].

Much of the research in optical microscopy imaging in the past two decades has focused on the development of strategies to break the Abbe resolution limit and achieve "superresolution microscopy." These techniques yield images with a level of detail close to the intrinsic scale of molecular biology. It is especially this endeavor that has led to major new challenges for the signal and image processing community. The most prominent recent examples of microscopy imaging techniques that rely heavily on image processing are photoactivated localization microscopy (PALM) [7] and the related technique of stochastic optical reconstruction microscopy [8]. These exploit the long-known fact that, even though the image of a subresolution particle is diffraction-limited (on the order of hundreds of nanometers), its location can be estimated with much higher accuracy (on the order of nanometers), depending on the signal-to-noise-ratio (SNR) [9]. Instead of acquiring a single image with all labels fluorescing concurrently, by using fluorescent proteins that can be switched on and off, thousands of images of well-separated particles can be acquired and their locations estimated very accurately by particle detection and fitting techniques. The composite image built up from the detections displays very high resolution. An alternative way to acquire better localized images is stimulated emission depletion [10], a nonlinear imaging technique that uses controlled de-excitation of previously excited off-center fluorophores. Another important technique is structured illumination microscopy [11], which computationally combines the images of differently oriented illumination patterns that produce Moiré fringes in the emission, resulting in an image with double resolution in each dimension. Complementary to these developments, improved techniques for imaging intact whole organisms have also been developed in recent years. Selective-plane illumination microscopy (SPIM), for example, in which the specimen is illuminated with a thin sheet of light perpendicular to the direction of observation, has proven to be an extremely valuable technique for long-time observation of embryonic development [12].

UNDERSTANDING THE CELL IN ITS SPATIOTEMPORAL CONTEXT IS KEY TO EXPLAINING MANY OF THE MECHANISMS OF LIFE AND DISEASE.

These and related advanced microscopy imaging techniques have enabled biologists to study the complexity of subresolution

intracellular organelles and the relation of their constituting components down to the molecular level and under physiological conditions in single cells or even developing organisms [13]–[15]. Technical improvements of these fluorescence-based imaging modalities, for instance, taking advantage of the sparsity of the signal being detected [16], go hand-in-hand with new versions of switchable fluorophores that allow nonlinear optical effects to be more efficiently used to push the resolution limit down further. The trends indicate that as imaging techniques become more and more sophisticated, requiring multiple pieces of information to be combined to reconstruct the full image, there is an increasing need for computationally efficient signal and image processing algorithms.

ANALYZING CELL BEHAVIOR

The extraction of biologically relevant information from both classical and novel microscopy imaging modalities requires the use of advanced image processing methods. Here, we will focus on techniques for segmenting and tracking cells and intracellular particles and for estimating cell-matrix tensional forces.

CELL SEGMENTATION AND TRACKING

Accurately defining the boundaries of cells in both static and dynamic images is a classical problem. It has been addressed over the years using a variety of segmentation methods [17]. Traditionally, cell segmentation in high-resolution fluorescence microscopy has been used to establish a spatial reference framework for the quantification of molecular or genetic events inside the cell. Alternatively, low-resolution cell segmentation and tracking, both in fluorescence and brightfield two-dimensional (2-D) microscopy has been applied to study the dynamics of cell populations with an emphasis more on the detection of population changes (e.g., cell mitoses, deaths, fusions) and motility (e.g., organ development, wound healing) than in accurately delineating cell morphology changes. More recently, a growing interest in the mechanobiology of the cell has brought back the focus of the segmentation and tracking field to the accurate delineation of cellular morphology and the changes associated to cell movements on flat substrates, and more recently, in 3-D environments.

CHARACTERIZING HOW THE CELL SENSES, REACTS TO, AND PRODUCES REGULATORY SIGNALS IS A DAUNTING TASK THAT REQUIRES THE INTEGRATION OF MORPHOLOGICAL, BIOCHEMICAL, MECHANICAL, AND ELECTRICAL CLUES.

There are two main approaches to the problem of cell tracking: tracking by detection and tracking by model evolution. The first approach consists of independently segmenting the cells in all the frames of a video and then, using association methods, tracking each segmented cell in all the video frames. The second approach uses evolution of curves or surfaces, either implicitly or explicitly defined, to track the boundaries of the cells along the entire length of the video. The first approach is more suitable for situations of low spatiotemporal resolution—high cell density, large time step—while the second performs best in high spatiotemporal resolution settings where high segmentation accuracy is required and there are few topological changes. Both paradigms can be enhanced by introducing knowledge of the topology changes (e.g., mitosis, apoptosis, fusion) into the data association (tracking by detection) or evolution (tracking by model evolution) phase of the algorithm.

The most recent tracking by detection methods use relatively simple segmentation approaches, such as wavelet decomposition [18], seeded watersheds [19], [20] or thresholding techniques [21], while investing their efforts in sophisticated association methods, such as minimum-coupled cost flow [18], dynamic programming [19], integer programming [22], or multiple-hypothesis [23] tracking. Some of these association methods implicitly incorporate the detection of topological changes [18], [19] while others include preprocessing detection of mitosis [20] or apoptosis to account for them. The state of the art of the tracking by model evolution paradigm uses the evolution of implicit contours (i.e., level sets) [21], [24] to segment and track individual cells. The principal limitation of these methods is the high computational cost involved in evolving one level set function per cell, by finding the numerical solution of its associated partial differential equation (PDE). To address this point, reducing the computational cost, Dufour et al. [25] use a discrete–parametric-active mesh framework and Maska et al. [26] minimize the original Chan–Vese model without solving any PDE, while evolving one single level set function per frame. Finally, there are complex methods that combine these two paradigms by using a fast level set framework combined with local spatiotemporal association [27].

A recurrent problem of the field was the lack of common test data sets and metrics to evaluate the performance of novel and existing algorithms. This prevented a fair and objective evaluation of the segmentation methods leaving the user (normally a noncomputer-proficient biologist) with the decision of choosing between the existing methods, with only the help of complex technical descriptions. In addition, not all published tracking algorithms are publicly available, or they have been released in a format that requires important computer and programming skills. To address this relevant issue, a benchmark for objective evaluation of cell tracking algorithms was recently established [28]. The challenge provides annotated data sets composed of both 2-D and 3-D video microscopy modalities (PC, DIC, fluorescence, confocal), nuclear and cytoplasmic staining, and various cell densities and microscopy resolutions (from high-throughput to high-resolution situations). Realistic simulations of nuclearly stained cells are also provided, for which there is an absolute, unbiased ground truth. The metrics used to compare the algorithms take into account both the accuracy of the segmentation and the accuracy of the tracking (movement and lineage) of the cells.

The outcome of the challenge revealed that the problem of segmenting and tracking cells in microscopy is far from being solved,

especially in the case of cytoplasmic labeled cells or in high-throughput setups (low spatiotemporal resolution and low SNR). In addition, more work needs to be done to segment cells in non-fluorescent microscopy modalities (PC, DIC) where the cellular boundaries present complex gradient patterns or extended artifacts, as well as in novel microscopy techniques such as SPIM.

PARTICLE DETECTION AND TRACKING

Quantitative analysis of healthy cell behavior and how various diseases may alter it often requires the analysis of intracellular dynamic processes. Examples include the motion of proteins or lipids on the cell membrane in relation to cell adhesion and regulation, the dynamics of cytoskeletal filaments involved in cell maintenance and intracellular transport, the interaction of virus particles with the cell machinery, and the intricate molecular processes involved in genome maintenance. Typically, these processes require very large numbers of "particles" (molecules, macromolecular complexes, organelles). In biological experiments, several hundreds to thousands of them are imaged at the same time to allow studying both the characteristic behavior and interaction of individual particles as well as aggregate behavior. Since manual annotation of the image data is infeasible, in addition to being inaccurate, this calls for advanced methods for automatic particle detection and tracking [23], [29]–[33].

QUANTITATIVE ANALYSIS OF HEALTHY CELL BEHAVIOR AND HOW VARIOUS DISEASES MAY ALTER IT OFTEN REQUIRES THE ANALYSIS OF INTRACELLULAR DYNAMIC PROCESSES.

Similar to methods for the analysis of cell dynamic behavior described in the previous section, methods for the tracking of intracellular particles in an image sequence usually consist of two fundamental stages [34]: 1) particle identification within individual image frames and 2) particle association from frame to frame to build trajectories. The goal of the first stage is to distinguish between local image intensity patterns that truly represent particles of interest versus irrelevant image structures and background. Commonly used image analysis methods for this purpose range from simple intensity thresholding, to more advanced linear filtering (in particular, Gaussian and its derivatives) and nonlinear wavelet-based or morphological image processing approaches. After detection, representative coordinates of the underlying particle within its corresponding local image patch are typically estimated by computing the intensity center of mass, by finding the local maximum, or by fitting a theoretical or experimentally obtained intensity model. In the case of spatially well-separated subresolution particles, the ideal model is the point-spread function of the microscope used, which in the case of both widefield and confocal fluorescence microscopy, can be well approximated by a Gaussian [35]. The localization problem is strongly linked to superresolution recovery (SRR) of the underlying true signal. While solid mathematical theory for SRR is now emerging [36], the development of computationally robust and efficient recovery algorithms remains a challenge, especially for multiparticle-tracking applications, where the data usually contains large numbers of (possibly overlapping) diffraction-limited spots drowning in very high levels of Poisson noise. The use of compressed sensing approaches has recently shown promising results in this area [16] and may be further improved by accurate statistical models.

The goal of the second stage in the particle-tracking process is to establish the best possible association of detected particles between image frames. Depending on the density of the particles within the field of view, and whether or not prior knowledge about their dynamic behavior is available, commonly used methods for this purpose range from simple nearest-neighbor linking (connecting each particle in a given frame with the spatially nearest particle in the next frame of the sequence), to more advanced multiframe association schemes, including multiple hypothesis tracking, dynamic programming, and various combinatorial approaches. The use of a motion model is often implemented in the form of Kalman filtering or in the case of nonlinear and non-Gaussian tracking problems, by means of sequential Monte Carlo estimation methods (often confusingly referred to as *particle filtering*). These can be made even more sophisticated by the use of interacting multiple motion models. However, the rise of high-density particle-tracking applications [37] is challenging currently existing methods, increasing the need for dealing with ever-larger amounts of imperfect data. As popular detection and localization methods yield optimal precision and accuracy only in circumstances that are rarely achieved in particle-tracking experiments [38], improved performance can be expected from novel methods that more intimately link the detection, localization, and association aspects of the tracking problem.

Since the early 1990s, many particle-tracking methods have been published based on the mentioned principles. With the increasing encouragement in the field to promote reproducible research, several dozens of software tools implementing these methods have been released [34]. To gain insight into their relative performance in an objective and reproducible manner, an open competition was recently organized [39]. One important finding is that, despite the often-heard claim when a new method is presented in the literature that it beats previous methods, as yet there exists no such thing as a single universal particle-tracking method that works best for all biological experiments. However, overall, certain methods do perform considerably better than others. A shared feature of superiorly performing methods is that they make optimal use of prior knowledge about both the objects of interest and the imaging process, re-emphasizing the importance of domain modeling. Another important finding is that current particle-tracking methods still tend to break down at SNRs representative of typical live-cell fluorescence microscopy imaging experiments. Although the SNR can be easily improved by increasing the illumination level, this has detrimental effects to the cell (photodamage and/or phototoxicity). Thus, in current practice, a careful selection of imaging conditions and analysis methods remains essential.

MEASUREMENT OF FORCES

Mechanobiology is an emerging field at the interface of biology and engineering. It focuses on the processes by which physical forces and cell or tissue mechanics contribute to development, normal physiology and disease. In the early 1920s, Buckminster Fuller first proposed the principle of tensegrity within the realm of architectural design. Tensegrity refers to the stability of a 3-D structure, granted by the opposing equilibrium between a discontinuous set of rigid compression elements and a continuous stabilizing tensile force. It was not until the mid-1970s that the (at that time) Yale undergraduate and current Harvard professor, Donald Ingber, related the behavior of a simple tensegrity stick-and-string model (flat when attached to a flat surface, abruptly becoming rounded when being detached from the flat surface) to the behavior of the cells that he had seen in culture at that semester's cancer lab. Ingber thought that cells might use their recently discovered internal framework, the cytoskeleton, to control their shape, much like a tensegrity structure does by means of a set of compressing elements and force distributing tensile elements [40]. Ingber further reasoned that cells must use their substrate, the extracellular matrix, to anchor themselves. Later on, he proved that mechanical forces exerted at the surface of the cell can be transmitted to the nucleus, resulting in biochemical changes and ultimately genetic changes—causing genes to turn on and off [41].

Understanding the molecular mechanism by which cells sense and respond to physical forces is a major challenge in this field. Traction force microscopy (TFM), a light microscopy technique developed in the mid-1990s [42], can compute traction forces exerted by a cell onto a biomimetic hydrogel substrate. These traction forces are calculated from the displacement of a large number of fluorescent beads embedded in the hydrogel, which in turn can be seen as samples of the deformation field that the forces cause in the substrate. Finally, the traction forces generating the deformations are inferred by direct or inverse methods that work from the expression of the laws of the elasticity of materials. Most of the existing, simplified methods compute the forces exerted on a plane by cells lying flat on a 2-D surface. Legant et al. [43], in a recent breakthrough, estimated traction forces exerted by cells fully encapsulated in a 3-D polymer gel. Using this physiologically relevant model, they discovered that the cells sensed the surrounding gel pulling strongly inward through traction anchors located near the tip of long, thin protruding extensions.

The classical procedure used to recover the forces in 2-D TFM experiments is composed of two steps: first, the displacement of the microbeads is calculated using particle-imaging velocimetry (PIV). Then, the stress field is obtained by considering the substrate as a linear and elastic half-space. The Boussinesq solution of the Green tensor is then computed using Fourier transform traction cytometry (FTTC) [44]. Legant et al. [43] relaxed the half-space constraint and solved the inverse problem within a 3-D geometry using the finite element method (FEM). More recently, the constraint on the linear behavior of the gel has been eliminated by combining multiple nonlinear FEM solutions, thus resulting in higher accuracy in the estimation of the forces [45].

It is clear that there is an urgent need to develop and integrate more efficient, precise, and robust computational methods. In particular, we believe that signal processing could greatly contribute to the technique with 1) robust and accurate cell segmentation algorithms as the ones described in the section "Cell Segmentation and Tracking," 2) sophisticated microbead displacement estimation methods such at the ones described in the section "Particle Detection and Tracking," and 3) fast and robust solutions for the ill-posed problem of recovering the forces (i.e., sparse tensor regularization, sparse reconstruction).

While TFM is well suited for the study of mechanotransduction at the cellular scale, particle-tracking microrheology [46] applies similar approaches to study mechanics at an intracellular scale. In particular, it enables measuring the local viscoelastic properties of the cytoplasm with high spatiotemporal resolution (i.e., nanoscale in seconds intervals). To this end, submicron particles are ballistically injected into the cytoplasm of live cells. After injection, the beads disperse rapidly within the cytoplasm, while being imaged using high-magnification fluorescence microscopy. The random spontaneous movement of the beads is tracked using particle-tracking methods. The trajectories of the cytoplasm-embedded particles are used to compute mean-squared displacements (MSDs). Finally, the time lag-dependent MSDs of the beads are transformed into local estimations of frequency-dependent viscoelastic moduli or the time-dependent creep compliance (deformability) of the cytoplasm. The ongoing efforts to extend the method to 3-D go hand-in-hand with the developments in superresolution microscopy and particle-tracking methods for high-density and low SNR conditions.

MODELING CELL BEHAVIOR

Ultimately, the information obtained using the quantification methods detailed in the section "Analyzing Cell Behavior" should be used to elaborate spatiotemporal models of cell appearance and behavior. The models should not only fit the available data but also lead to new hypotheses that can subsequently be verified experimentally. The truth is that there still remains a long way before a morphodynamic model capturing all the cell's complexity becomes available.

Approaches to modeling cell behavior can be divided into two main categories—top-down and bottom-up—that are being developed independently. Top-down approaches rely primarily on image data starting at the cell level and going down to imaging selected subcellular components trying to infer rules of cell morphology and behavior hidden inside. On the other hand, bottom-up approaches rely primarily on nonimage data (such as bioinformatics databases, signaling pathways, gene expression data, genome sequencing data, measurements of forces, etc.) starting at the level of very basic partial rules of cell behavior and going up to defining more complex rules for specific cellular processes or behavior of small cellular components ultimately leading to the model of how the whole cell works.

TOP-DOWN APPROACHES

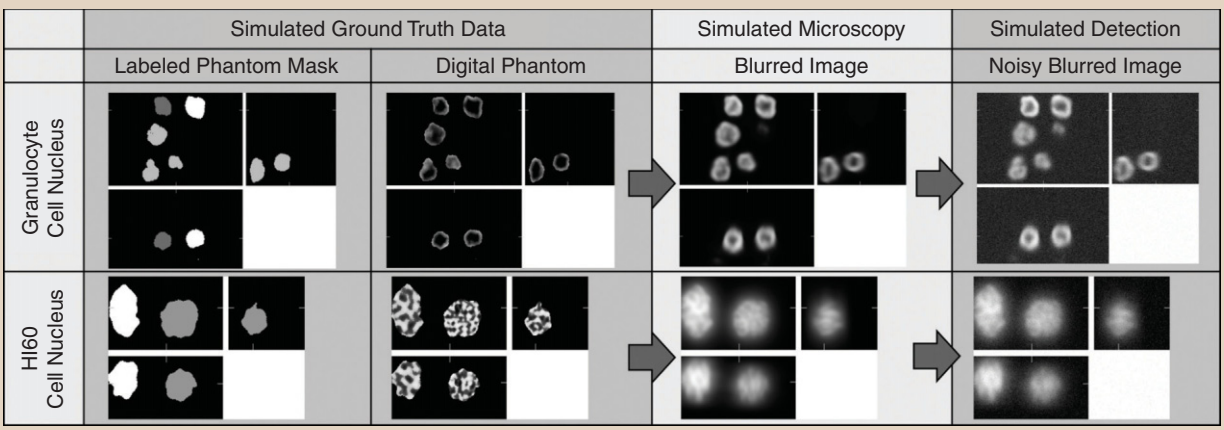
The first models of the cell date back to the 19th century [47], being just rough descriptions based on the limited unspecific observation provided by the microscopes of the time. After oil-immersion lenses became available in the 1870s, people could observe the structure of the membrane, nucleus and cytoplasm (at the time called *protoplasm*). For example, in 1885, C. Rabl [48] published his famous model of the nucleus consisting of nucleoli and chromatin formed by chromosomes. The model even illustrated chromosome behavior during mitosis.

Computers enabled a gradual transition from those original descriptive models based on rather vague verbal explanations of cell components accompanied by drawings to more precisely formulated mathematical models. The earliest mathematical tool to describe cell behavior (interaction of neighboring cells) was the cellular automaton, which became famous in 1970 thanks to John Conway's "Game of Life."

By the end of the 20th century, it became popular to represent cells or cell nuclei using simple mathematical shapes (spheres, ellipsoids, discs, or rods) and create virtual microscopy images that could be used as digital phantoms. These were dedicated to test the limits of image segmentation algorithms to different noise levels, blur degradations, or phantom densities. These digital phantoms enable the comparison of the algorithm results with a known ground truth [28]. Later on, more sophisticated artificial objects were developed: shapes were modeled as randomly deformed spheres or ellipsoids, and texture was added to simulate staining of cell or cell nucleus. There is an abundant work aimed at describing and classifying both subcellular structures and whole cells based on the analysis of protein distribution (i.e., image texture) and morphological descriptors of the cell. This information can be readily incorporated to the digital descriptions of the cell. An excellent review of these machine-learning approaches both from a theoretical and practical point view has been recently presented by Conrad and Gerlich [49].

The virtual microscopy "observation" of these artificial cells also improved: Gaussian blurring was replaced by the convolution with a real point spread function, more noise types were considered (Poisson, Gaussian, dark current, fixed pattern) and imaging artifacts introduced (uneven illumination, depth-related aberrations in 3-D samples, etc.). For a particular cell type, virtual microscopy images can be made almost indistinguishable from reality (not only visually but also based on computed image characteristics) [50]. See Figure 1 for an example of digital phantoms and corresponding virtual microscopy images.

Learning-based cell modeling is another modality that infers algorithmic parameters from training image data by employing supervised or nonsupervised machine-learning techniques [51]. Parameters need not to be just single values but may be also expressed as probability density functions. Learning techniques build either discriminative models dedicated to object classification (of unknown test data) or generative models able to synthetize new artificial images belonging to a given class. Learning-based approaches can be used not only for testing image analysis algorithms but also to characterize the differences between healthy and pathological cells or for structure-function relationship studies.



[FIG1] Modeling cells and their components using 3-D digital phantoms. First, a digital phantom is created for each modeled biological object in the field of view as a synthetic solid object of a precisely defined shape filled with a certain texture. The shape defines the binary mask serving as ground truth segmentation result and the image with texture serves as the input for virtual (simulated) microscopy. The image of the phantom produced by the optical system of a virtual microscope is typically generated by adding blur using convolution with a suitable point spread function. The blurred image is further subject to virtual image detection by adding adequate noise of various types. The noisy blurred images must be indistinguishable from real images of modeled biological objects and can serve for testing performance of image analysis algorithms against known ground truth. Examples of 3-D digital phantoms for two types of cell nuclei are shown. Each 3-D image is shown as a triplet of three mutually orthogonal cuts through the object: xy view (upper left), xz view (bottom), and yz view (right).

Another extended approach is to create a model of a specific cell segmented in a particular image (more precisely a model of its stained components visible in the image data). This approach lacks generalization but can help revealing hidden properties using simulations. For instance, to measure the diffusion coefficient in fluorescence recovery after photobleaching (FRAP), diffusion is simulated on the segmentation-based model [52]. This approximation to the study of biological systems and the learning-based approach presented above are sometimes denoted as image-based systems biology to distinguish them from modeling in computational biology where image data is not considered.

> **THE ADVENT OF NOVEL IMAGING TECHNIQUES, COUPLED WITH THE USE OF ADVANCED COMPUTATIONAL AND SIGNAL PROCESSING METHODS, HAS OPENED THE DOOR TO UNDERSTANDING MANY CRUCIAL ASPECTS OF THE CELL.**

Several software tools for modeling cells and their components have been made publicly available: SIMCEP (2-D digital phantoms), available at http://www.cs.tut.fi/sgn/csb/simcep/; CytoPacq (3-D digital phantoms), available at http://cbia.fi.muni.cz/simulator/; or CellOrganizer (learning-based models), which can be found at http://cellorganizer.org. Moreover, one can use pregenerated benchmark data sets offered on the Web pages of these software packages. The synthetic cell images are still available only for just a few cell types and several cell components but can be generated in large quantities with different levels of noise, various cell densities, and are accompanied by ground truth data. Lately, also time-lapse sequences of such synthetic image data have become available and have been used, for example, in the Cell Tracking Challenge, available at http://www.codesolorzano.com/celltrackingchallenge/.

BOTTOM-UP APPROACHES

Parallel to these black box-modeling efforts, there are attempts to mathematically model the intricacies of the signaling pathways that govern the cellular function of the cell. These models, if properly populated with a complete list of substances (e.g., genes, ribonucleic acids (RNAs), proteins), rules (e.g., transformation of molecular species, reaction kinetics) and cellular spatial or functional compartments, can simulate the molecular machinery of the cell. E-CELL, developed by Tomita et al. [53] is a software environment that simulates the behavior of a cell from the activity of gene sets derived from entire genomes. As a proof of principle, the authors presented a model of a minimal cell based on a subset of genes of *Mycoplasma genitalium*, whose complete 580-kbit genome was sequenced in 1995. This simplified model simulates how proteins interact within the living cell. Specifically, it models how changes in the amount of a protein (by knocking out the corresponding gene or altering its expression level), or the medium (e.g., starving the cell by removing glucose) may affect its behavior (e.g., mitotic rate, probability of entering in apoptosis) and its survival.

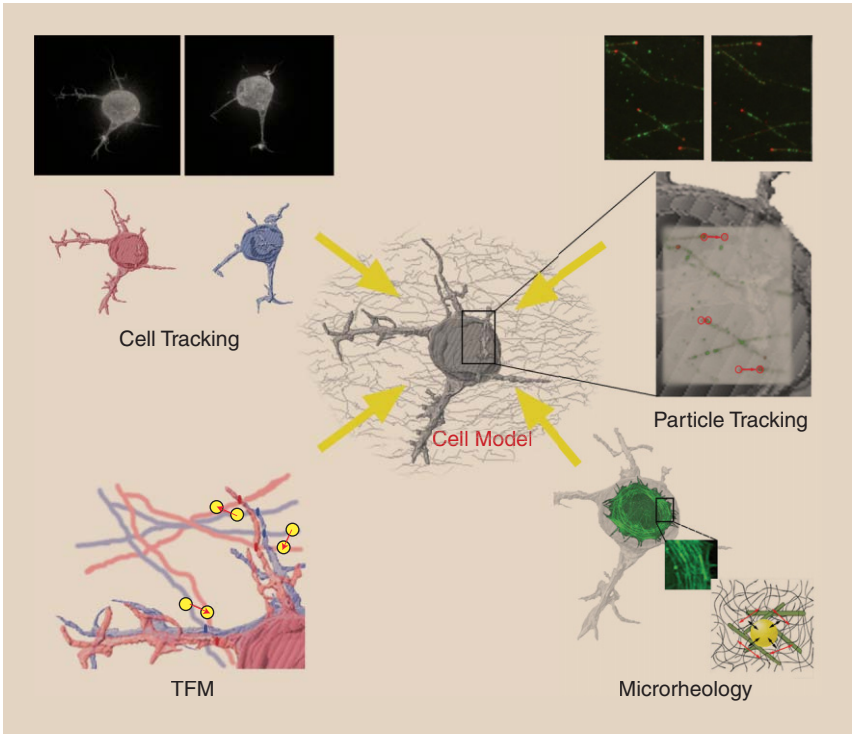
These mathematical models are becoming increasingly complex to account for higher organisms and more temporal scales. Simultaneously, computational models [54] are being developed which, instead of representing cell processes with equations, present recipes (algorithms) that mimic natural phenomena. Instead of searching for a mathematical solution to a complicated list of equations, it provides algorithms that steer into different states or configurations of a cell. The rules of navigation are operational, hence the name *executable biology*. In both types of models (mathematical and executable biology), there is a close connection with the experimental image-based data that feeds the models and are used to validate, and when necessary, update them.

Finally, the recent developments in the field of mechanobiology allow integrating morphological and molecular aspects with the mechanical interactions between the cell and its environment, thus creating mechanical models of the cell. These efforts are fed with information about morphological changes (see the section "Cell Segmentation and Tracking"), traction force and viscosity data (obtained using among others, some of the tools described in the section "Measurement of Forces"), coupled to the trafficking of mechanosensitive and mechanotransductive biomolecules (see the section "Particle Detection and Tracking"). Most of these methods have been used to model cell motility, mostly of cells crawling on a surface [55], from the treadmill of actin that implies persistent front-to-back asymmetry, through a synchronized assembly–disassembly directional process. These models explain the formation and release of lamellas and protrusions, as well as the adhesion-mediated contraction that facilitates the push-and-pull mechanism required for the cell to move on its substrate. Other models focus on mechanosensing in general [56], and on how the mechanical properties of the cell, defined by the composition and structure of its cytoskeleton self-adjust as a reaction to the mechanical properties of the extracellular environment.

FUSION OF AVAILABLE KNOW-HOW

Top-down and bottom-up approaches are complementary to each other. As the coverage of the former ones goes down the scale and deeper into the cell (thanks to the development of imaging techniques enabling observations of subresolution targets with increasing spatial as well as temporal resolution) while the coverage of the latter ones goes up the scale (due to the advances in molecular biology, cellular biochemistry, or the development of high-throughput screening methods), they tend to meet and cover certain cell components or events both from the rules side and from the imaging side. For example, within the MitoCheck project (http://www.mitocheck.org/), systematic analysis of genes and proteins that are required for chromosome segregation and cell division in human cells was performed by inactivating all 22,000 human genes one by one in cultured human cells using RNA interference (RNAi) and recording cellular phenotypes by high-throughput live-cell imaging.

Besides advances in imaging and biotechniques, advances in computer vision and artificial intelligence may also help integrating all pieces of know-how together into a single model of the cell. Besides the ability of learning from image data (as in the case of learning-based modeling methods), they can offer multimodal fusion of knowledge from different sources [57] independent of the application—the data can be related to studying human appearance and behavior in a video sequence as well as studying cell appearance and behavior in a time-lapse series. Multimodal fusion techniques can help integrate pieces of information obtained from different sources depicted in Figure 2, which could be the key for defining an accurate cell model.



[FIG2] A summary of image analysis techniques described in the article. From the top left, counterclockwise: Cell tracking provides dynamic information about morphological changes of a moving cell; TFM calculates tensional forces between a cell and its surrounding environment from the displacement (red arrows) of fluorescence beads (yellow dots) embedded in the extracellular substrate; microrehology informs about the viscoelastic properties of the intracellular space from the microscopic movements (black arrows) of ballistically injected fluorescent nanobeads (yellow dot) under the stress of fibers (green lines), which can be represented as soft and stiff spring series (read arrows); and particle-tracking algorithms provide information about the movement (red arrows) and trafficking of subcellular elements (red dots).

DISCUSSION AND PERSPECTIVE

The cell is an extremely complex machine. Its changing morphology and its dynamic spatial relationship with the surrounding environment depend on the biochemical composition of the latter, internal and external mechanical stresses, electrical signals, gravity, etc. Furthermore, the cell's metabolic production (i.e., its phenotype) is regulated by genetic and epigenetic factors that depend and have an impact on its morphology. Consequently, a faithful model of the cell should take into account the interplay of all these factors in their precise spatiotemporal context.

Needless to say, such a model cannot be based on static observation of the cell, much like the complexity of the universe cannot be explained based on a single snapshot of the skies. The cell is in a particular environment, and the cell is in a precise developmental stage. That explains why the exact genetic content gives rise to such a diverse display of cell types and phenotypes that coexist in a living organism. Therefore, a model of the cell should integrate all the factors involved (e.g., genes, RNAs, molecular/metabolic signaling pathways, structural elements of the cell nucleus and cytoskeleton, forces, biochemical factors) in its precise time and location. Light microscopy, as described in this article, provides visually quantifiable information that feeds into these models. Figure 2 presents a graphic summary of how quantitative image analysis provides information about the cell and its dynamic processes.

Simultaneously visualizing and quantifying all these internal and external players during the entire life of a cell, within its native tissue context is beyond all possible imagination, due to technical and physical limitations. However, the advent of novel imaging techniques, coupled with the use of advanced computational and signal processing methods has opened the door to understanding certain aspects of the cell that can be used to populate a computational model. In this article, we have reviewed the history and the state of the art of both hardware and software that are contributing to this enterprise.

In the hardware arena, the existing microscopy techniques can capture a few events in a relatively limited spatiotemporal framework, mostly in 2-D in vitro setups. The use of synthetic hydrogels of controlled biomechanical properties has recently facilitated the study of cells in more realistic 3-D environments, thus taking full advantage of the sectioning and 3-D imaging capabilities of the diverse flavors of confocal and multidimensional microscopy. Simultaneously, two complementary technological efforts are being pursued. In particular, the development of novel superresolution microscopy methods (especially, those that may work in 3-D and time lapse) and the development of large-scale, whole organ, or whole animal imaging systems, where resolution is sacrificed for the benefit of spatial completeness. Both efforts require significant input from the signal processing community in the areas of efficient data sampling, single-molecule detection, fast sparse-image reconstruction, compression, and data handling.

Software development necessarily follows the advances in image acquisition, thus leading the way to novel 3-D particle and

cell tracking routines. The existence of high-sensitivity sensors has placed the emphasis on the need for fast, optimized tracking methods, while the increasing use of high-throughput systems for big-data analysis pushes toward the development of very robust segmentation and tracking algorithms that may work in low SNR situations. Similarly, the incipient field of mechanotransduction demands novel, more efficient methods for the calculation of cell traction forces, especially in 3-D environments.

Finally, partial, in silico models of the cell, based on simple genetic and molecular approximations are already available. Those are being complemented with morphomechanistic models of cell behavior as it is visualized using an optical microscope of tunable properties. Mechanical models that incorporate the role of forces, and viscoelastic properties in the homeostasis and dynamics of cells are also being developed. These models are far from being complete, and work only as partial descriptions of some cellular processes during limited temporal steps. Furthermore, it remains to be defined how the models of the cells will be incorporated into similarly complex models of both subcellular–molecular models and dynamic models of complete organs or even entire organisms. The tremendous challenge posed to the signal processing and modeling community is how to integrate all the information about the cell—biochemical, structural, and mechanical—into a single unifying, multiscale, and spatiotemporal model that may open the door to the explanation and engineering of life. This, far from being a science fiction exercise, is the goal of the field of synthetic biology. For instance, in what could be considered a breakthrough in the field, Annaluru et al. [58], have reported the synthesis of a functional 272,871–base pair designer eukaryotic chromosome, based on the 316,617–base pair native Saccharomyces cerevisiae chromosome III. The future is indeed here, since the descriptive models of the cell will provide mechanistic information eventually leading to the production of functional cells. The use of this artificial life, properly empowered by bioethical principles, may clear the way to a new era in the field of tissue engineering and regenerative medicine.

ACKNOWLEDGMENTS

Carlos Ortiz-de-Solorzano and Arrate Muñoz-Barrutia were supported by the Spanish Ministry of Economy and Competitiveness grants with reference DPI2012-38090-C03-02 and TEC2013-48552-C02, respectively. Michal Kozubek was supported by the Czech Science Foundation (302/12/G157).

AUTHORS

Carlos Ortiz-de-Solórzano (codesolorzano@unav.es) received a Ph.D. degree in telecommunication engineering from the Universidad Politécnica de Madrid, Spain, in 1996. Since 2004, he has been the head of the Cancer Imaging Laboratory, Center for Applied Medical Research, in Pamplona, Spain. He is also full professor of signal processing at the School of Engineering of the University of Navarra, Pamplona, Spain. His research interests include the fields of quantitative fluorescence microscopy and image analysis, with applications to study the molecular mechanisms of normal tissue development and cancer. He has authored over 80 manuscripts in a wide range of peer-reviewed journals. He has been a member of the IEEE Engineering in Medicine and Biology Society as well as the IEEE Signal Processing Society since 1999. He is a Senior Member of the IEEE.

Arrate Muñoz-Barrutia (arrmunoz@unav.es) received a Ph.D. degree from the Swiss Federal Institute of Technology Lausanne (EPFL) in 2002. Currently, she is a staff scientist at the Cancer Imaging Laboratory of the Center for Applied Medical Research of the University of Navarra, Pamplona, Spain, and also an associate professor in the engineering school of the same university. She has authored more than 30 peer-reviewed international journals in biomedical image processing. She has been a regular member of the Bioimaging and Signal Processing Technical Committee of the IEEE Signal Processing Society since 2011. She has been an associate editor of *IEEE Signal Processing Letters* since 2012. She is a Senior Member of the IEEE.

Erik Meijering (meijering@imagescience.org) is an associate professor at the Erasmus University Medical Center in Rotterdam, The Netherlands, where he leads a research group devoted to cellular and molecular image analysis. He was a special sessions chair for the 2002 and 2004 IEEE International Symposium on Biomedical Imaging, and technical program chair for that meeting in 2006 and 2010. He was/is an associate editor of *IEEE Transactions on Medical Imaging* (since 2004), *IEEE Transactions on Image Processing* (2008–2011), and *International Journal on Biomedical Imaging* (2006–2009), and was a guest editor of *IEEE Transactions on Image Processing* for its September 2005 special issue on "Molecular and Cellular Bioimaging." He is a Senior Member of the IEEE.

Michal Kozubek (kozubek@fi.muni.cz) received the Ph.D. degree in informatics from the Faculty of Informatics, Masaryk University, Brno, Czech Republic, in 1998. He is a group leader at the Centre for Biomedical Image Analysis, Masaryk University, where he was an associate professor from 2002 to 2011, becoming a full professor in 2012. He has been the dean at the Faculty of Informatics, Masaryk University, since 2011. His research interests include image acquisition and analysis.

REFERENCES

[1] K. Hiroaki, "Systems biology: A brief overview," *Science*, vol. 295, no. 5560, pp. 1662–1664, 2002.

[2] D. Noble, "Modeling the heart—From genes to cells to the whole organ," *Science*, vol. 295, no. 5560, pp. 1678–1682, 2012.

[3] Y. Setty, "Multi-scale computational modeling of developmental biology," *Bioinformatics*, vol. 28, no. 25, pp. 2022–2028, 2012.

[4] M. W. Davidson and M. Abramowitz. (1999). Optical microscopy. Online technical note. [Online]. Available: http://www.olympusmicro.com/primer/opticalmicroscopy.html

[5] B. N. G. Giepmans, S. R. Adams, M. H. Ellisman, and R. Y. Tsien, "The fluorescent toolbox for assessing protein location and function," *Science*, vol. 312, no. 5771, pp. 217–224, 2006.

[6] C. Vonesch, F. Aguet, J.-L. Vonesch, and M. Unser, "The colored revolution of bioimaging," *IEEE Signal Processing Mag.*, vol. 23, no. 3, pp. 20–31, 2006.

[7] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1654, 2006.

[8] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science*, vol. 319, no. 5864, pp. 810–813, 2008.

[9] S. Ram, E. S. Ward, and R. J. Ober, "Beyond Rayleigh's criterion: A resolution measure with application to single-molecule microscopy," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 12, pp. 4457–4462, 2006.

[10] S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.*, vol. 19, no. 11, pp. 780–782, 1994.

[11] M. G. L. Gustafsson, L. Shao, P. M. Carlton, C. J. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," *Biophys. J.*, vol. 94, no. 12, pp. 4957–4970, 2008.

[12] J. Huisken and D. Y. R. Stainier, "Selective plane illumination microscopy techniques in developmental biology," *Development*, vol. 136, no. 12, pp. 1963–1975, June 2009.

[13] M. Lakadamyali, "Super-resolution microscopy: Going live and going fast," *ChemPhysChem*, vol. 15, no. 4, pp. 630–636, 2014.

[14] P. Dedecker, G. C. Mo, T. Dertinger, and J. Zhang, "Widely accessible method for superresolution fluorescence imaging of living systems," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 109, no. 27, pp. 10909–10914, 2012.

[15] S. Abrahamson, J. Chen, B. Hajj, S. Stallinga, A. Y. Katsov, J. Wisniewski, G. Mizuguchi, P. Soule, F. Mueller, C. Dugast Darzacq, C. Wu, C. I. Bargmann, D. A. Agard, M. Dahan, and M. G. Gustafsson, "Fast multicolor 3D imaging using aberration-corrected multifocus microscopy," *Nat. Methods*, vol. 10, no. 1, pp. 60–63, 2013.

[16] L. Zhu, W. Zhang, D. Elnatan, and B. Huang, "Faster STORM using compressed sensing," *Nat. Methods*, vol. 9, no. 7, pp. 721–723, 2012.

[17] E. Meijering, "Cell segmentation: 50 years down the road," *IEEE Signal Processing Mag.*, vol. 29, no. 5, pp. 140–145, 2012.

[18] D. Padfield, J. Rittscher, and B. Roysam, "Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis," *Med. Image Anal.*, vol. 15, no. 4, pp. 650–668, 2011.

[19] K. E. G. Magnusson and J. Jaldén, "A batch algorithm using iterative application of the Viterbi algorithm to track cells and construct cell lineages," in *Proc. 9th IEEE Int. Symp. Biomedical Imaging (ISBI'12)*, Barcelona, Spain, 2012, pp. 382–385.

[20] N. Harder, F. Mora-Bermúdez, W. J. Godiznez, J. Ellenberg, R. Eils, and K. Rohr, "Automated analysis of the mitotic phases of human cells in 3-D fluorescence microscopy," in *Medical Image Computing and Computer-Assisted Intervention*. New York: Springer, 2006, pp. 840–848.

[21] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillén-Aghion, J. C. Olivo-Marin, and C. Zimmer, "Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces," *IEEE Trans. Image Processing*, vol. 14, no. 9, pp. 1396–1410, 2005.

[22] F. Li, X. Zhou, J. Ma, and S. T. C. Wong, "Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 96–105, Jan. 2010.

[23] N. Chenouard, I. Bloch, and J. C. Olivo-Marin, "Multiple hypothesis tracking for cluttered biological image sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35. pp. 2736–2750, Nov. 2013.

[24] O. Dzyubachyk, W. A. van Cappellen, J. Essers, W. J. Niessen, and E. Meijering, "Advanced level-set based cell tracking in time-lapse fluorescence microscopy," *IEEE Trans. Med. Imaging*, vol. 29, no. 3, pp. 852–867, 2010.

[25] A. Dufour, R. Thibeaux, E. Labruyère, N. Guillén, and J. C. Olivo-Marin, "3-D active meshes: Fast discrete deformable models for cell tracking in 3-D time-lapse microscopy," *IEEE Trans. Image Processing*, vol. 20, no. 7, pp. 1925–1937, 2011.

[26] M. Maška, O. Danek, S. Garasa, A. Rouzaut, A. Muñoz-Barrutia, and C. Ortiz-de-Solorzano, "Segmentation and shape tracking of whole fluorescent cells based on the Chan-Vese model," *IEEE Trans. Med. Imaging*, vol. 32, no. 6, pp. 995–1006, 2013.

[27] K. Li, E. D. Miller, M. Chen, T. Kanade, L. E. Weiss, and P. G. Campbell, "Cell population tracking and lineage construction with spatiotemporal context," *Med. Image Anal.*, vol. 12, no. 5, pp. 546–566, 2008.

[28] M. Maska, et al., "A benchmark for comparison of cell tracking algorithms," *Bioinformatics*, vol. 30, no. 11, pp. 1609–1617, 2014.

[29] I. Smal, M. Loog, W. Niessen, and E. Meijering, "Quantitative comparison of spot detection methods in fluorescence microscopy," *IEEE Trans. Med. Imaging*, vol. 29, no. 2, pp. 282–301, 2010.

[30] A. Genovesio, T. Liedl, V. Emiliani, W. J. Parak, M. Coppey-Moisan, and J. C. Olivo-Marin, "Multiple particle tracking in 3-D+t microscopy: Methods and application to the tracking of endocytosed quantum dots," *IEEE Trans. Image Processing*, vol. 15, no. 5, pp. 1062–1070, 2006.

[31] I. Smal, K. Draegestein, N. Galjart, W. Niessen, and E. Meijering, "Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: Application to microtubule growth analysis," *IEEE Trans. Med. Imaging*, vol. 27, no. 6, pp. 789–804, 2008.

[32] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. L. Schmid, and G. Danuser, "Robust single-particle tracking in live cell time-lapse sequences," *Nat. Methods*, vol. 5, no. 8, pp. 695–702, 2008.

[33] W. J. Godinez, M. Lampe, S. Wörz, B. Müller, R. Eils, and K. Rohr, "Deterministic and probabilistic approaches for tracking virus particles in time-lapse fluorescence microscopy image sequences," *Med. Image Anal.*, vol. 13, no. 2, pp. 325–342, 2009.

[34] E. Meijering, O. Dzyubachyk, and I. Smal, "In imaging and spectroscopy analysis of living cells," in *Methods Enzymol.*, vol. 504, P. M. Conn, Ed. New York: Elsevier, 2012, ch. 9, pp. 183–200.

[35] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Appl. Opt.*, vol. 46, no. 10, pp. 1819–1829, 2007.

[36] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Applied Math.*, vol. 67, no. 6, pp. 906–956, 2014.

[37] J.-B. Sibarita, "High-density single-particle tracking: Quantifying molecule organization and dynamics at the nanoscale," *Histochem. Cell Biol.*, vol. 141, no. 6, pp. 587–595, 2014.

[38] H. Deschout, F. C. Zanacchi, M. Mlodzianoski, A. Diaspro, J. Bewersdorf, S. T. Hess, and K. Braeckmans, "Precisely and accurately localizing single emitters in fluorescence microscopy," *Nat. Methods*, vol. 11, no. 3, pp. 253–266, 2014.

[39] N. Chenouard, et al., "Objective comparison of particle tracking methods," *Nat. Methods*, vol. 11, no. 3, pp. 281–289, 2014.

[40] D. E. Ingber, "Cellular tensegrity: Defining new rules of biological design that govern the cytoskeleton," *J. Cell Sci.*, vol. 104, no. 3, pp. 613–627, 1993.

[41] D. E. Ingber, "Cellular mechanotransduction: Putting all the pieces together again," *FASEB J.*, vol. 220, pp. 811–827, May 2006.

[42] T. Olivier, M. Dembo, and K. Jacobson, "Traction forces in locomoting cells," *Cell Motil. Cytoskeleton*, vol. 31, no. 3, pp. 225–240, 1995.

[43] R. Legant, J. S. Miller, B. L. Blakely, D. M. Cohen, G. M. Genin, and C. S. Chen, "Measurement of mechanical tractions exerted by cells in three-dimensional matrices," *Nat. Methods*, vol. 7, no. 12, pp. 969–971, 2010.

[44] J. Boussinesq, "Application des potentiels á l'étude de l'équlibre et du movement des solides élastiques," Gauthier-Villars, Paris, 1985.

[45] J. Palacio, A. Jorge-Peñas, A. Muñoz-Barrutia, C. Ortiz-de-Solorzano, E. de Juan-Pardo, and J. M. García-Aznar, "Numerical estimation of 3D mechanical forces exerted by cells," *J. Biomech.*, vol. 46, no. 1. pp. 50–55, 2013.

[46] Y. Tseng, T. P. Kole, and D. Wirtz, "Micromechanical mapping of live cells by multiple-particle-tracking micro-rheology," *Biophys. J.*, vol. 83, no. 6. pp. 3162–3176, 2002.

[47] P. Mazzarello, "A unifying concept: The history of cell theory," *Nat. Cell Biol.*, vol. 1, pp. E13–E15, May 1999.

[48] C. Rabl, "Über Zelltheilung," in *Morphologisches Jahrbuch*, vol. 10, C. Gegenbaur, Ed. Leipzig, Germany: Willhelm Engelmann, 1885, pp. 214–330.

[49] C. Conrad and D. W. Gerlich, "Automated microscopy for high-content RNAi screening," *J. Cell Biol.*, vol. 188, no. 4, pp. 453–461, 2010.

[50] D. Svoboda, M. Kozubek, and S. Stejskal, "Generation of digital phantoms of cell nuclei and simulation of image formation in 3D image cytometry," *Cytometry*, vol. 75A, no. 6, pp. 494–509, 2009.

[51] T. E. Buck, J. Li, G. K. Rohde, and R. F. Murphy, "Towards the virtual cell: Automated approaches to building models of subcellular organization 'learned' from microscopy images," *Bioessays*, vol. 34, no. 9, pp. 791–799, 2012.

[52] I. Sbalzarini, "Modeling and simulation of biological systems from image data," *Bioessays*, vol. 35, no. 5, pp. 482–490, 2013.

[53] M. Tomita, K. Hashimoto, K. Takahashi, T. S. Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. C. Venter, and C. A. Hutchison III, "E-CELL: Software environment for whole-cell simulation," *Bioinformatics*, vol. 15, no. 1, pp. 72–84, 1999.

[54] J. Fisher and T. A. Henzinger, "Executable cell biology," *Nat. Biotechnol.*, vol. 25, no. 11, pp. 1239–1249, 2007.

[55] A. Mogilner, "Mathematics of cell motility: Have we got its number?" *J. Math. Biol.*, vol. 58, no. 1–2, pp. 105–134, 2009.

[56] C. Borau, T. Kim, T. Bidone, J. M. García-Aznar, and R. D. Kamm, "Dynamic mechanisms of cell rigidity sensing: Insights from a computational model of actomyosin networks," *PLoS ONE*, vol. 7, no. 11, p. e49174, 2012.

[57] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, "Multimodal fusion for multimedia analysis: A survey," *Multimedia Syst.*, vol. 16, no. 6, pp. 345–379, 2010.

[58] N. Annaluru, et al., "Total synthesis of a functional designer eukaryotic chromosome," *Science*, vol. 44, no. 6179, pp. 55–58, 2014.

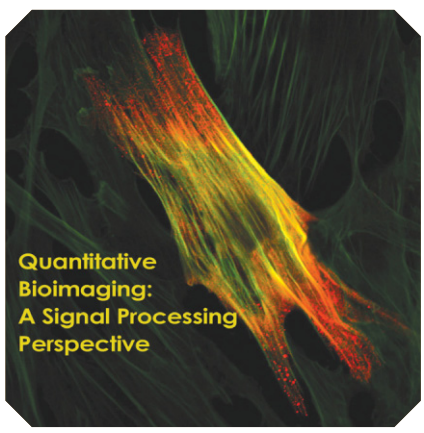
[SP]

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

[Alexandre Dufour, Tzu-Yu Liu, Christel Ducroz, Robin Tournemenne, Beryl Cummings, Roman Thibeaux, Nancy Guillen, Alfred Hero III, and Jean-Christophe Olivo-Marin]

Signal Processing Challenges in Quantitative 3-D Cell Morphology

[More than meets the eye]



Modern developments in light microscopy have allowed the observation of cell deformation with remarkable spatiotemporal resolution and reproducibility. Analyzing such phenomena is of particular interest for the signal processing and computer vision communities due to the numerous computational challenges involved, from image acquisition all the way to shape analysis and pattern recognition and interpretation. This article aims at providing an up-to-date overview of the problems, solutions, and remaining challenges in deciphering the morphology of living cells via computerized approaches, with a particular focus on shape description frameworks and their exploitation using machine-learning techniques. As a concrete illustration, we use our recently acquired data on amoeboid cell deformation, motivated by its direct implication in immune responses, bacterial invasion, and cancer metastasis.

MOTIVATION AND CHALLENGES

Cell deformation and migration are dynamic processes regulated by a complex machinery with major implications on a number of key processes in biology including development, immune responses, and invasive processes [1]. A method of choice for studying this mechanism lies in light microscopy, whereby living cells evolving in their three-dimensional (3-D) environment (both in vitro and in vivo) can be imaged over prolonged periods of time with limited invasiveness, producing time-lapse sequences of volumetric 3-D images [2]. Due to the considerable complexity of cell deformation and migration, visual analysis of such processes is no longer limited just by user bias and fatigue but also fails to apprehend large-scale, population-wise patterns that may otherwise appear random or disorganized. Systematic quantitative analysis and understanding of cellular dynamics is becoming a major interest for the signal processing and computer vision communities, given the wide range of computational challenges to overcome. These challenges principally fall into one of the following five categories, covering many aspects of the experimental pipeline (cf. Figure 1).

IMAGE RECOVERY

Modern optical light microscopy techniques have substantially expanded the diversity and reliability of live cell imaging applications, constantly improving on speed, penetration depth, and spatial resolution, though usually at the expense of the signal-to-noise ratio. An important part of the literature therefore focuses on the development of deconvolution and denoising techniques adapted to the peculiarities of bioimaging data (e.g.,

Digital Object Identifier 10.1109/MSP.2014.2359131
Date of publication: 5 December 2014

1053-5888/15©2015IEEE

mixed Poisson–Gaussian noise and anisotropic lateral-to-axial resolution [3]–[5]), while recent advances in computational optics (notably compressive sensing and superresolution techniques) have given rise to new challenges in signal reconstruction and inverse problems [6]–[9].

SEGMENTATION AND TRACKING

The increasing diversity and complexity of environments in which motile cells can be observed has made their reliable detection and tracking most challenging, notably in crowded or cluttered environments [10]. Added to the sheer amount of routinely produced 3-D imaging data, the need to develop fast and semi- to fully automated approaches remains a long-standing challenge in the community [11], [12].

SHAPE REPRESENTATION

Although informative, raw shape and trajectory information are usually too large and complex to produce interpretable results, unless their dimensionality is sufficiently and adequately reduced. Unfortunately, the natural variability of shape configurations observable within so-called homogeneous cell populations pose significant challenges in defining descriptors that are both robust to noise and retain enough specificity across populations. While much effort has been conducted to develop such descriptors in two-dimensional (2-D) [13] or pseudo-3-D [14], 3-D-shape descriptors that permit robust morphological analysis and facilitate human interpretation are still under active investigation [15]–[21].

SYSTEMATIC QUANTITATIVE ANALYSIS AND UNDERSTANDING OF CELLULAR DYNAMICS IS BECOMING OF MAJOR INTEREST FOR THE SIGNAL PROCESSING AND COMPUTER VISION COMMUNITIES, GIVEN THE WIDE RANGE OF COMPUTATIONAL CHALLENGES TO OVERCOME.

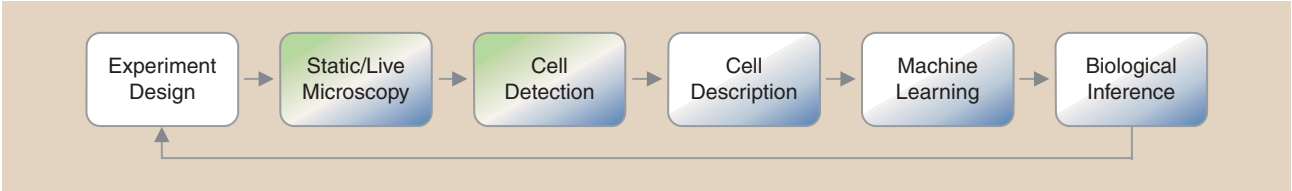
LEARNING AND INTERPRETATION

Linked to the issue of shape description is that of invariant analysis of cell populations across various experimental conditions. The difficulty here lies in two aspects: 1) developing pattern recognition and machine-learning approaches able to capture the differences between populations while remaining robust to intraclass variability [22], [23] and 2) highlighting such differences in a human-readable form and ultimately leading to the inference of standardized computational models with the aim of deriving novel biological hypotheses [24].

AVAILABILITY AND REPRODUCIBILITY

The vast majority of developments in the community generally appears in the literature in the form of theoretical workflows that facilitate understanding and software (or hardware) implementation. Unfortunately, the increasing complexity of these protocols renders their implementation and validation very tedious for nonspecialists, hindering both their adoption and reproducibility. Ironically, while many scientific findings are based on computerized analysis, the associated computer codes are only rarely made public in contrast to reproducible research practices in other scientific domains [25]. Community efforts such as the Reproducible Research Initiative strive to make both code and data publicly available, although more support from publishers and/or research sponsors is required [26].

In the specific context of 3-D bioimaging, the image recovery and segmentation aspects have received extensive focus from the signal processing community over the last several decades, as illustrated by the recent introduction of challenges at the IEEE International Symposium on Biomedical Imaging, with special sessions on image deconvolution (http://bigwww.epfl.ch/deconvolution/challenge), particle tracking [27], and cell segmentation and tracking [12]. The shape representation and machine-learning aspects have been comparatively less thoroughly investigated, even though they provide essential keys to decipher the cell machinery. Here we review recent developments in the fields of cell-shape description and associated machine-learning approaches, highlighting the current state of the art and the challenges ahead toward a comprehensive understanding of cellular dynamics. We also review a number of open-source software solutions that permit reliable and reproducible quantification of cellular images. We shall illustrate this review using the example of amoeboid cell deformation, which is a mechanism of strong interest in the life science community due to its importance in immune response, infectious diseases, and cancer metastasis. Amoeboid motion is characterized by the emission of localized protrusions at the cell surface that permit environment scanning and motion initiation (cf. Figure 2) [28], [29] and poses significant challenges in terms of quantitative characterization and comparative phenotyping.



[FIG1] Typical experimental pipeline for cellular (notably phenotypic) studies. Boxes marked with green and blue labels indicate the availability of associated open hardware and software developments, respectively.

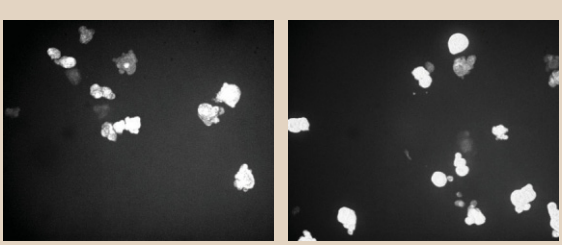
3-D-SHAPE DESCRIPTORS AND THE CELL

Proper description of the cell shape in a numerical form is inherently dependent on its underlying representation, i.e., how it was extracted from the image data [13]. While most segmentation methods usually produce voxel-based masks or distance maps as an indicator of the cell location and cell interior (analogous to the input data), outline-based approaches produce a description that is homotopic to the cell surface (e.g., using control points or parameterized curves [14], [30], [31]). From any of these representations, various low-level descriptors can be extracted (e.g., volume, surface, elongation, ellipticity, compacity, etc.), providing a coarse appreciation of the global shape conformation (most of these descriptors can be extracted using the approaches further reviewed below). However, when dealing with more complex shapes such as cells or organs, such low-level descriptors quickly become sensitive to increasing amounts of noise and usually fail to capture subtle shape variations occurring at different spatial scales, from large conformation changes (e.g., elongation, contraction) to smaller variations at the cell surface (e.g., protrusions) [32]. Higher-lever descriptors based on multiscale decomposition become particularly interesting in this context, as they decompose the shape in a coarse-to-fine manner, allowing one to restrict the analysis to the scale of interest and thereby increasing robustness to negligible or unlikely shape variations. These approaches typically fall into one of the three following categories, classified in increasing order of complexity.

IRONICALLY, WHILE MANY SCIENTIFIC FINDINGS ARE BASED ON COMPUTERIZED ANALYSIS, THE ASSOCIATED COMPUTER CODES ARE ONLY RARELY MADE PUBLIC IN CONTRAST TO REPRODUCIBLE RESEARCH PRACTICES IN OTHER SCIENTIFIC DOMAINS.

LANDMARK-BASED APPROACHES

Landmark-based approaches are a popular choice for morphological studies [33], with numerous applications in medical imaging [34], [35], evolutionary biology [21], and face recognition [36]. The common denominator in these fields is the availability of a reliable low-dimensional model for the shapes of interest, allowing one to reduce the description of the shape (or its tolerated deformations) to a small number of control points (or parameters). This simplified representation in turn permits efficient registration, statistical analysis, and template modeling [37]. Applications in cellular morphology are, however, not as common, mostly due to the fact that deforming cells generally have many more degrees of freedom that cannot be accurately captured using such methods (counterexamples can be found in specific biological applications, e.g., [38]).



[FIG2] Planar slices of two field of views representing (a) wild type and (b) chemically modified parasites. Distinguishing between these populations based on shape information is particularly challenging, even for the trained eye, and requires robust quantitative tools for shape description and machine learning.

GRAPH-BASED REPRESENTATIONS

Graph-based representations fall in two subcategories, depending on whether they describe the interior or the outline of the shape of interest. In the former case, the cell body is converted (e.g., from a initial binary mask) into a hierarchical treelike graph connecting virtual landmarks inside the cell. Typical examples include morphological skeletons, medial axis transforms, or Voronoi tessellations [39], [40]. Once the graph is obtained, local shape features at the cell surface (the leaves of the tree) are semantically segregated from large shape conformation (closer to the root). As the graph generation process may be subject to noise, adequate graph pruning algorithms are required to differentiate structures at the cell edge (e.g., filopodia from erratic spikes), thereby permitting an unbiased analysis of the cell deformation over time [40]. However, such approaches remain limited to 2-D analysis, and their extension to 3-D is computationally challenging.

The latter category considers a surface-based graph representation of the shape of interest. While the topic of signal processing on arbitrary graphs is only in its early days [41], several methods have been developed for the specific case of closed surfaces (homeomorphic to the two-sphere), such as energy-minimizing graph matching (developed for protein surface alignment [42]) and graph-based spherical wavelets (applied to cell-shape analysis in [18]).

MOMENT-BASED APPROACHES

These approaches consider the shape of interest as an arbitrary spatial distribution function that is then mathematically represented as a sum of known polynomial functions, thereby permitting the extraction of geometrical moments with suitable invariants [43]. Such methods generalize traditional Fourier analysis to arbitrary distributions and therefore share the same descriptive properties: low-order moments describe the coarse conformation, while high-order moments retain information at higher frequency. For this reason, these approaches have been utilized in many areas of image processing, with popular choices of

bases including Lengendre, Zernike, Tchebichef polynomials (see [44] for a comparison), and splines [45].

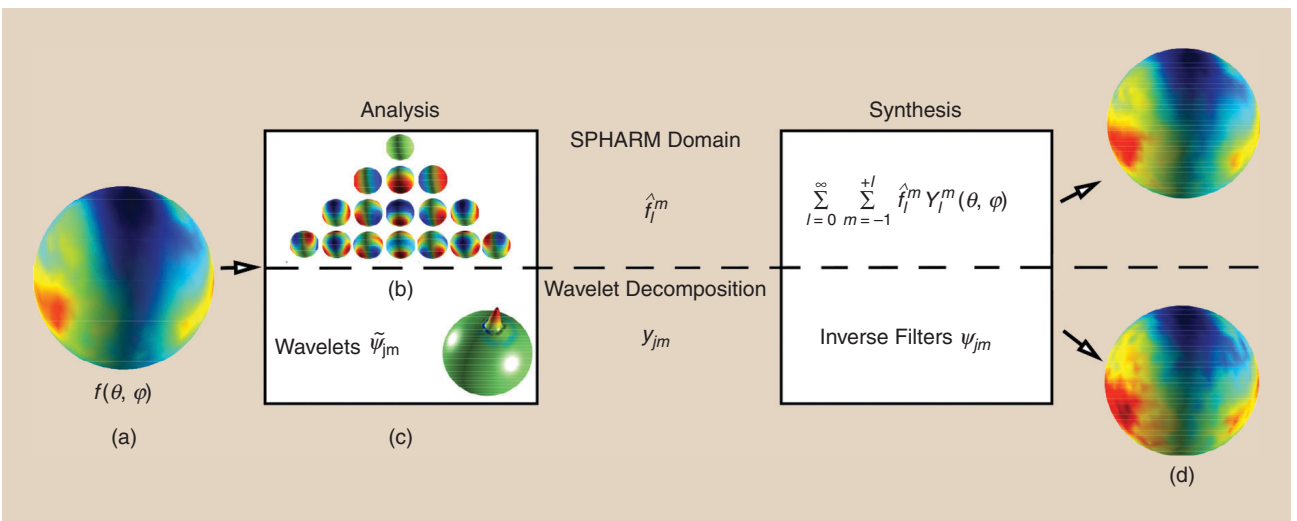
In the context of cell-shape description, such approaches are generally not applied on the raw image data. Instead, a binary mask or outline of the shape is first extracted and then projected onto an appropriate basis [14], [16], [19], [20], [46], [47]. These methods can further be decomposed into two categories. On the one hand, the 2-D cell outline is projected directly onto a chosen basis (e.g., Fourier [46] or splines [14]), and the process is repeated on each slice of the cell shape to obtain a 3-D set of descriptors [14]. On the other hand, the surface of interest is first mapped onto the sphere using appropriate spherical parameterization techniques [48] and then projected onto a reference function basis living on the sphere. Two popular and complementary candidates in this family are the spherical harmonics (SPHARM) [47] and spherical wavelets (SWAVE) [49], which significantly differ from the eponym graph-based approach, notably in the way they are constructed. In the former case, the spherical signal is projected onto a basis of Legendre polynomials, extending the classical Fourier analysis to signals on the two-sphere [cf. Figure 3(b)]. SPHARM therefore have global spatial support, and each coefficient describes the general conformation of the shape of interest at different spatial scales. Applications of SPHARM include molecular surface modeling [50], [51], medical-shape analysis [52], and cell-shape analysis [15], [16], [19], [53]. In the latter case, the function basis is formed of wavelets (hence its name), and are constructed by analogy to wavelets in the plane via appropriate spherical projections [17], [20], [54]. Here the local spatial support provided by SWAVE is of particular interest to localize specific features along the surface [cf. Figure 3(c)].

SHAPE EXTRACTION AND DESCRIPTION HAVE LED TO SIMPLIFIED REPRESENTATIONS OF THE RAW DATA FROM 3-D IMAGES DOWN TO A SMALLER SET OF DESCRIPTORS PER CELL.

While the function bases utilized here are not specific to cellular shapes, other approaches have been proposed to increase their specificity by locally adapting the basis to the data set at hand (e.g., using the Laplace–Beltrami operator, as in [19] and [55]). Nevertheless, the use of standard function bases preserves two major advantages: 1) they are ubiquitousness in signal processing applications that propels the creation of ever-more efficient computational implementations and 2) they permit an unbiased description and comparison of shapes across multiple experimental conditions, and also serve as a basis to perform shape synthesis (cf. Figure 3) or build so-called generative models of the cell [24].

RECOGNITION, CLASSIFICATION, AND INTERPRETATION

In terms of dimensionality, shape extraction and description have led to simplified representations of the raw data from 3-D images (on the order of 10^{8-10} voxels) down to a smaller set of descriptors (also called *features*) per cell (on the order of 10^{1-3}). Unfortunately, these feature sets are rarely translatable to a concrete, biologist-friendly interpretation of the biological experiment, rather appearing as large arrays of poorly informative numbers. This motivates the following questions: Which features really matter? What is the influence of the experimental conditions on these features? How do they translate into biological terms? Machine-learning approaches are particularly well suited to answer these questions, and choosing the appropriate technique depends on the application and how



[FIG3] The comparison of spherical harmonics and spherical wavelets for signal processing on the sphere. (a) The signal of interest $f(\theta, \varphi)$ can be projected on a basis of functions with global spatial support such as (b) spherical harmonics, or a basis of functions with local support such as (c) spherical wavelets. In both cases, the spherical signal can be synthesized back (d) from the coefficients.

easily the initial question can be cast into a machine-learning framework [23].

Machine learning consists of building computer models that accurately describe a given population of individuals with the ultimate goal to either characterize subgroups of individuals with similar properties, or to predict the properties of a new unknown (or simulated) individual [62]. Machine-learning techniques are generally split into two categories: 1) supervised techniques require that a subset of individuals in each subpopulations be manually annotated to classify the rest of the data set, while 2) unsupervised techniques learn the model directly from the inherent structure of the feature set without user intervention. Both families have their set of advantages and drawbacks, which we discuss next in the context of cell-shape analysis (we refer the reader to [23] for an applications-oriented review, and [63] for a more theoretical introduction).

> QUANTITATIVE MORPHOLOGY OF SINGLE CELLS IS ONLY THE VISIBLE PART OF THE DIGITAL BIOIMAGING "ICEBERG."

Supervised techniques require a subset of the data to be annotated, i.e., one or more examples describing all known subclasses of the population must be indicated beforehand. This so-called training set is then used to learn an optimal classifier. Popular approaches for cell classification from image data include k-nearest neighbors [64], decision trees [65], and support vector machines (SVMs) [57]. These approaches can be further combined, resulting in so-called ensemble learning techniques, e.g., boosting [22] and random forests [66]. Supervised techniques are widely popular due to their robustness to noise and apparent intuitiveness and versatility. Annotating a small finite number of examples is generally acceptable in many applications, while the classifier itself remains sufficiently generic to accept a wide range of applications. Unfortunately, some of these methods may suffer from overfitting when the dimension becomes much larger than the number of samples. Variable selection techniques provide a solution to this issue, while improving learning accuracy and often facilitating interpretation. Typical examples include forward selection, backward elimination and sparsity-constrained classifiers [67]. The major limitation of supervised learning is that it requires class labels to be available: nonannotated subpopulations will not be learned and, hence, not recognized, and by extension, novel unknown subpopulations (e.g., unpredicted cell phenotypes) cannot be discovered.

Unsupervised techniques do not require a training set and can be applied to unlabeled populations. They learn the inherent structure of the data set using a predefined metric (e.g., a similarity or distance measure between individuals), permitting homogeneous groups or dimensions to be distinguished. One usually distinguishes clustering techniques, which aim at extracting subpopulations sharing similar properties according to the considered metric, from dimensionality reduction techniques, which aim at selecting a subset of essential principal components that best represent a high-dimensional data set to facilitate user interpretation. Classical clustering techniques include *k*-means [68] and Gaussian mixture modeling (GMM) [69], while dimensionality reduction techniques include principal component analysis (PCA) [46] and independent component analysis (ICA) [70]. The major advantage of unsupervised learning is that the data labels need not be known in advance, alleviating the need for data annotation while allowing the discovery of unexpected subpopulations, giving them remarkable exploratory potential in biology [71]. Unfortunately, unsupervised techniques also have their drawbacks: they are more sensitive to noise, defining the appropriate metric for the data set at hand can be complex for high-dimensional data set, as is the interpretation of the results.

It is worth pointing out that most of these techniques can also be applied directly to the raw image data without necessarily needing a preliminary shape extraction and description step. Shape description becomes necessary as soon as both qualitative and quantitative characterization or modeling of the cell shape is required, notably when studying the effect of known experimental conditions on the cell phenotype.

AVAILABILITY AND REPRODUCIBILITY

Emerging interdisciplinary fields such as bioimage informatics foster interactions across an ever broader portfolio of scientific expertise (this article only mentions six of them: optics, signal processing, image segmentation, object tracking, shape description, and machine learning). Unfortunately, novel algorithmic developments in many of these fields are only rarely published in the form of ready-to-use software, while reimplementing the underlying method becomes increasingly challenging for the nonspecialist in

[TABLE 1] OPEN-SOURCE SOFTWARE SOLUTIONS WITH DEDICATED MODULES FOR CELL-SHAPE ANALYSIS.

NAME	REFERENCE	SUPPORTED LANGUAGES	2-D/3-D	SHAPE DESCRIPTORS	MACHINE LEARNING
CELLCLASSIFIER	[56]	MATLAB*	2-D	GEOMETRIC	SUPERVISED
CELLCOGNITION	[57]	PYTHON, C++	2-D	GEOMETRIC	SUPERVISED AND UNSUPERVISED
CELLORGANIZER	[24]	MATLAB*	2-D, 3-D	SPLINES	UNSUPERVISED (GENERATIVE)
CELLPROFILER	[22]	PYTHON, VISUAL PROGRAMMING	2-D	GEOMETRIC	SUPERVISED
EBIMAGE	[58]	R	2-D	GEOMETRIC	(VIA R)
ICY	[59]	JAVA, SCRIPTING, VISUAL PROGRAMMING	2-D, 3-D	GEOMETRIC, SPHARM	(VIA PLUGINS)
IMAGEJ/FIJI	[60]	JAVA, SCRIPTING, MACRO RECORDING	2-D, 3-D	GEOMETRIC	(VIA PLUGINS)
TANGO	[61]	IMAGEJ, R	3-D	GEOMETRIC	(VIA R)

*MATLAB is licensed by Mathworks.

SPHARM AS SHAPE DESCRIPTORS

SPHARM are defined as

$$Y_l^m(\theta,\varphi) = k_{l,m} P_l^m(\cos\theta) e^{im\varphi},$$

where θ and φ parameterize the spherical domain, l and m are respectively the degree and order of the harmonic, $k_{l,m}$ is the expansion coefficient, and P_l^m is the associated Legendre polynomial.

Spherical harmonic analysis is a natural extension of traditional Fourier analysis for signals defined on the unit sphere. Hence, any arbitrary function f defined on the sphere can be expanded using the SPHARM transform, given by

$$f(\theta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} C_l^m \cdot Y_l^m(\theta,\varphi),$$

where C_l^m are the generalized Fourier coefficients with respect to the SPHARM basis, or more simply, SPHARM coefficients.

To conduct a SPHARM expansion of the cell shape, its surface must be written as a spherical function, which is done via so-called spherical parameterization techniques [47]. While some surfaces may not be bijectively transposable to the sphere by a simple radial projection (also referred to as nonstar-shaped surfaces), a classical approach is to project each Cartesian component of the surface independently, yielding a vector of spherical functions $f = [f_x(\theta,\varphi)\, f_y(\theta,\varphi)\, f_z(\theta,\varphi)]$ [48]. Expanding f thus yields three sets of SPHARM coefficients $C_l^m = [(C_l^m)_x\, (C_l^m)_y\, (C_l^m)_z]$. Rotation invariant coefficients are subsequently obtained by considering their L2-norm:

$$\hat{C}_l^m = \| C_l^m \| = \sqrt{(C_l^m)_x^2 + (C_l^m)_y^2 + (C_l^m)_z^2}.$$

[TABLE 2] A COMPARISON OF SEVERAL SUPERVISED CLASSIFICATION APPROACHES FOR CELL SHAPE CLASSIFICATION.

CLASSIFICATION METHOD	SPHARM FEATURE SET	ERROR (%)
k-NEAREST NEIGHBORS [64]	STANDARD ($K = 1, Q = 1$)	43.96
	POPULATION	33.75
	TEMPORAL	26.25
	COMBINED	29.23
DECISION TREES [65]	STANDARD ($K = 1, Q = 1$)	45.52
	POPULATION	35.94
	TEMPORAL	31.67
	COMBINED	30.94
SVM [57]	STANDARD ($K = 1, Q = 1$)	42.69
	POPULATION	27.73
	TEMPORAL	18.08
	COMBINED	20.00
STRUCTURED SVM [74]	STANDARD ($K = 1, Q = 1$)	40.45
	POPULATION	31.36
	TEMPORAL	17.69
	COMBINED	19.23

SPHARM-BASED FEATURE DESIGN FOR CLASSIFICATION

Protocol

- SPHARM expansion of the cell surfaces (cf. "Spherical Harmonics as Shape Descriptors") is conducted with an empirical precision of $l = 5$, yielding an array of 21 rotationally invariant coefficients per cell (this value depends on the application, and defines the balance between shape and noise information).
- The data set is then divided into groups of $K \in \{1, 2, \ldots, 5\}$ randomly selected cells observed over $Q \in \{1, 5, 10, 15, 20, 25\}$ consecutive frames (the starting frame is random if Q is less than the entire length of the video). We avoid imbalance in the training samples by randomly subsampling the larger class so that the sizes of both classes are identical.
- Finally, the coefficients are indexed throughout the data set as $\hat{C}_l^m(i, \text{cell}_k, q)$, where $i \in \{1, 2, \ldots, n\}$ is the sample indicator, $k \in \{1, \ldots, K\}$ indicates the cell, and $q \in \{1, \ldots, Q\}$ indicates the time at which the frame is acquired. From this data set, one can design a structured combination of features. Here we illustrate two possible combinations:

Population features (μ) are obtained by averaging each coefficient over Q frames for each cell of a group, yielding $21K$ features:

$$\mu_l^m(i, \text{cell}_k) = \frac{1}{Q} \sum_{q=1}^{Q} \hat{C}_l^m(i, \text{cell}_k, q).$$

Temporal features ($\tilde{\mu}$) are obtained by averaging each coefficient over a group of K cells in each frame, yielding $21Q$ features:

$$\tilde{\mu}_l^m(i, q) = \frac{1}{K} \sum_{k=1}^{K} \hat{C}_l^m(i, \text{cell}_k, q).$$

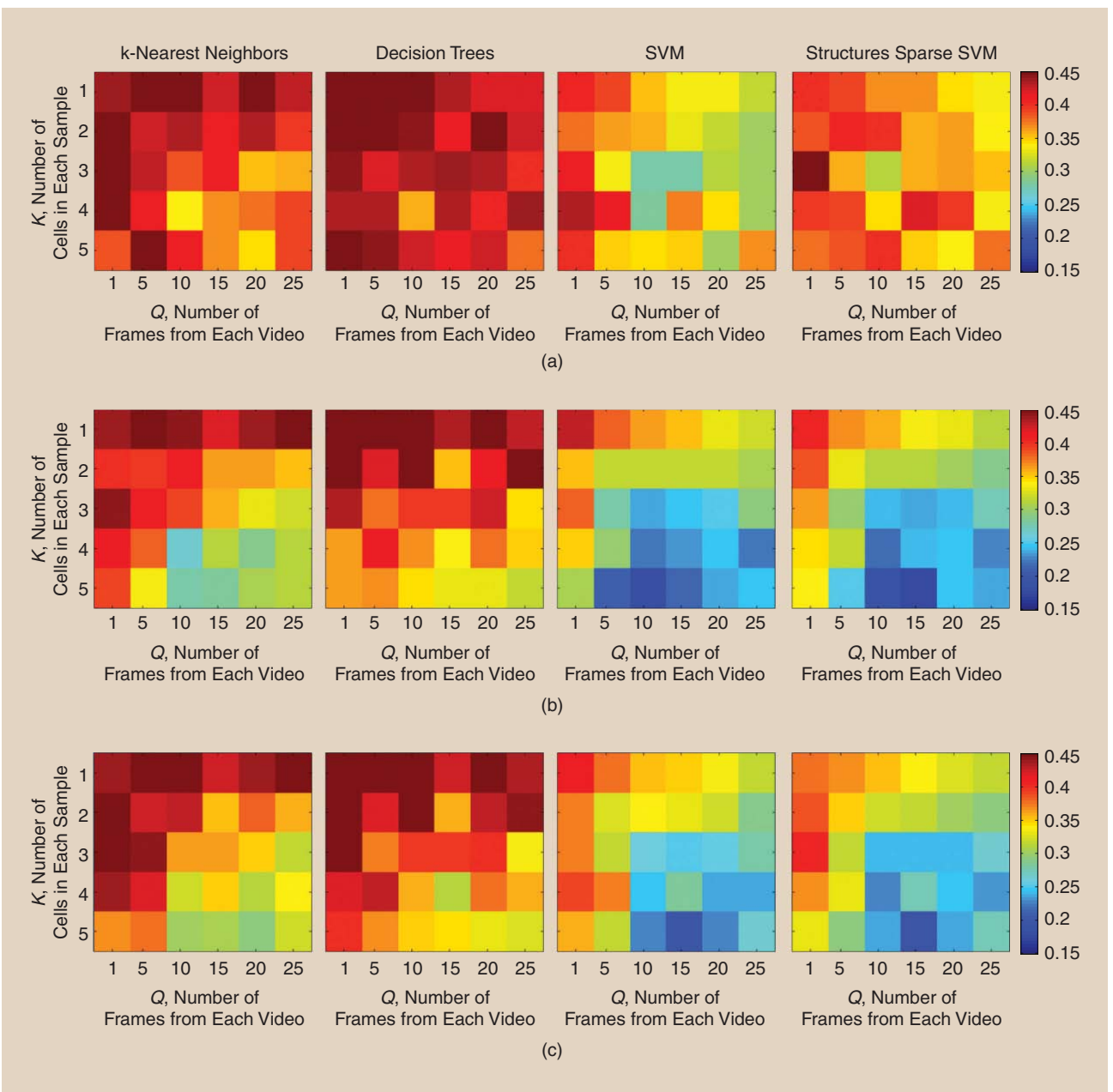
From these descriptors, a set of group structures can be derived:

$$I_{\mu l^*} = \{\mu_l^m(i, \text{cell}_k) : l = l^*, \forall i, k\}$$
$$I_{\tilde{\mu} l^*} = \{\tilde{\mu}_l^m(i, t) : l = l^*, \forall i, t\}.$$

addition to being time-consuming and error prone [25]. Fortunately, the bioimage informatics community has been a proactive driver of the Reproducible Research Initiative [26] with the appearance of multiple software solutions for quantification in bioimaging over the last decades (cf. [72] and Table 1). Some of these frameworks rely on a so-called plug-in architecture, allowing their enrichment via third-party contributions, removing much of the redundant work including data loading and visualization, and streamlining the publishing process. Each platform will tend to be specialized in a specific discipline (mostly driven by its core developers), and current efforts in the community are devoted to provide higher interoperability across software, so as to combine the best solutions available for to tackle a given problem. Table 1 provides a nonexhaustive list of such software platforms, focusing specifically on the solutions available for cell-shape analysis and

quantification. Such software can be classified in two categories: general-purpose image-based quantification software for which specific third-party modules for cellular analysis have been developed (this is notably the case of Fiji/ImageJ [60] and Icy [59]), and more streamlined software dedicated almost exclusively to cell-based analysis (such as CellClassifier [56], CellCognition [57], CellOrganizer [24], CellProfiler [22], EBImage [58], and Tango [61]). While some of these software packages limit their analysis to the 2-D case, the methods are in principle extensible to 3-D. This limitation is notably present in the field of high-throughput, high-content screening (HT-HCS), where 3-D imaging considerably increased acquisition and analysis times, alongside data management issues [22].

THE ROAD TO BUILDING A THE CELL IS PAVED WITH MAJOR DIFFICULTIES IN TERMS OF MATHEMATICAL MODELING AND SIGNAL PROCESSING.



[FIG4] The performance comparison of several classification methods on various combinations of SPHARM features. Each table cell indicates the classification error for a particular combination of the number of cells per group (K) and the number of time points per cell (Q), as defined in "SPHARM-Based Feature Design for Classification." (a) Population features (μ). (b) Temporal features ($\tilde{\mu}$). (c) Population and temporal features ($\mu, \tilde{\mu}$).

CLASSIFICATION METHODS

We consider an input feature set $x_i \in R^p$ and the class labels $y_i \in \{+1, -1\}$, where $i = 1, \ldots, N$. Under this notation, we can describe the classification methods used in this review as follows.

k-Nearest Neighbors

Let $N_k(x)$ be the neighborhood of x found by the k nearest samples, defined by some metric, e.g., Euclidean distance. The decision rule is defined by a majority vote on $\{y_i \,|\, x_i \in N_k(x)\}$.

Decision Trees

The decision tree is an greedy algorithm that adds splitting nodes to the tree by defining half planes $P_1 = \{x \,|\, x_j \leq s\}$ and $P_2 = \{x \,|\, x_j > s\}$, in which x_j is the splitting variable and s is the splitting point. At each candidate node, compute the impurity, e.g., the Gini index

$$I = p_{-1}(1 - p_{-1}) + p_{+1}(1 - p_{+1}),$$

in which p_k is the fraction of class k observed at that node. The splitting nodes are selected to improve the homogeneity sequentially, and the decision at each leaf node is by majority vote.

SVMs

Let $f(x) = w'x + b$ be a decision function that assigns observation x to $sign(f(x))$, then $f(x) = 0$ is a hyperplane in R^p. In the general case, where classes are not linearly separable, the SVM is written as

$$\min_f \sum_{i=1}^{n} V(f, x_i) + \lambda R(f)$$

with

$$V(f, x_i) = [1 - y_i f(x_i)]_+, R(f) = \frac{1}{2} \| w \|_2^2.$$

Structured SVMs

The structured sparse SVM is formulated as an SVM but with a different regularization:

$$V(f, x_i) = [1 - y_i f(x_i)]_+, R(f) = \sum_{g=1}^{G} \| w_{I_g} \|_2,$$

where I_g is the set indexing the variables that are in the gth group.

On the hardware side, recent efforts to make public blueprints and to increase the interaction between hardware and software has led to the development of so-called smart acquisition systems. Some of the key contributions on the hardware side include the Warwick Open Microscope System (http://wosmic.org) and the OpenSPIM project (http://openspim.org), while integrated hardware and software approaches have been developed in several projects including μManager (http://micro-manager.org) and MicroPilot [73].

CASE STUDY: AMOEBOID CELL DEFORMATION

In this section, we illustrate the use of shape description and classification in the context of studying morphology and motility of *Entamoeba histolytica,* a unicellular parasite responsible for the amoebiasis disease. Recent studies in vitro and ex vivo have suggested that parasites specifically modified to prevent their degrading of the extracellular matrix remained able to migrate at the same speed as unmodified parasites, possibly due to subtle shape changes that simple descriptors could not fully capture [32]. Illustrative slices of the 3-D data set are presented in Figure 2. Here cells are segmented using active contours [30], and described by SPHARM features (cf. [16] and "SPHARM as Shape Descriptors" and "SPHARM-Based Feature Design for Classification"). Although other combinations of methods and descriptors are possible, our aim here is to show how an adequate design of feature vectors may lead to discoveries of significant and interpretable differences between populations that otherwise seem visually identical.

Table 2 reports the comparative performance of several classification approaches (described in "Classification Methods") with various combinations of features. The performance is evaluated by leave-one-out cross-validation, while the entire leave-one-out process is repeated for ten trials to report the averaged performance for each method. Both the temporal features (averaged over groups of cells) and population features (averaged over a fixed number of frames) are tested against variable group sizes, as illustrated in the error rate heat-maps presented in Figure 4. As the group size increases, the classification performance systematically improves, suggesting that more robust classification can be achieved by group-based analysis. Table 2 further indicates that the best performance is obtained either using temporal features or a combination of temporal and population features. However, combining both feature sets does not necessarily yield the best performance, suggesting that temporal information is more discriminant. Finally, structured SVM imposes sparsity to select an optimal subset of features and has better classification performance. Also, by specifying the most predictive subset of features, the structured SVM has the advantage of providing feature interpretability. While each feature alone does not necessarily translate directly to biological knowledge, this example illustrates that a careful and systematic design of the features can highlight significant and unexpected discrepancies between cell populations, which in turn can potentially lead to new interpretations and hypotheses that enhance the design of the next experiment.

FUTURE PERSPECTIVES AND CHALLENGES

Numerous breakthroughs in imaging and computational techniques have had a considerable impact on the amount of quantitative data that describe the behavior of single cells evolving in their 3-D environment. After suitable standardization, such data become amenable to proper mathematical characterization, invariant description, and classification. Thus, these tools have exciting potential to reveal the complexity of biological mechanisms at all spatial scales: 1) at the single cell level, the mechanisms and signals responsible for the transition between modes of migration (notably in cancer development); 2) at the group level, the short- and long-distance signaling queues that induce cells to interact, differentiate,

or migrate to distant locations (e.g., for tissue repair and immune responses); and 3) at the tissue level, how information is propagated across a dense cell network, and how this information is locally interpreted to drive cell intercalation, differentiation, and renewal (e.g., during embryo- and morphogenesis). In this article, we have highlighted some of the latest developments in quantitative tools and associated software packages to study some of these processes and illustrated how group-based analysis of cell morphology provides a much more powerful and discriminant description of a cell population as compared to single-cell analysis, while temporal information carries a significant potential to improve the overall classification performance. Yet, quantitative morphology of single cells is arguably only the visible part of the digital bioimaging "iceberg."

THE NEXT MAJOR CHALLENGES IN THE BIOIMAGING AND BIOSIGNAL PROCESSING FIELD LIE IN STUDYING SPATIOTEMPORAL PROCESSES BEYOND SINGLE CELLS, FROM THE NANOSCOPIC TO THE MACROSCOPIC SCALE.

The next major challenges in the bioimaging and biosignal processing field lie in studying spatiotemporal processes beyond single cells, from the nanoscopic to the macroscopic scale. At the subcellular level, the mechanisms that underlie cell deformation and motility are still poorly understood, mostly due to the lack of proper visual insight into the various architectural components (down to individual proteins) forming the cell cytoskeleton. The road to building a computational model of the cell (not to mention the huge variety of cell types in plant or animal models) is paved with major difficulties in terms of mathematical modeling and signal processing and will require the development of novel biophysics-inspired algorithms to: understand how the cytoskeleton is formed, acts, and reacts to internal and environmental signals; and generates force and adhesion that ultimately lead to deformation, movement, and division. At the macroscopic level, studying biological processes at large spatial and temporal scales requires the integrating of single-cell analyses over millions of cells and hours of imaging data in multiple modalities and experimental conditions, raising major visualization and computational bottlenecks. One example of such a challenge is illustrated by the recent advances in selective plane illumination microscopy and its application to the reliable observation of embryonic development in numerous animal models, from the single cell up to tens of thousands of cells [10]. Such data sets have already initiated many developments in image denoising, cell segmentation, cell tracking, and data manipulation software, however, comprehensive modeling of the morphology and trajectory of these cells and their clustering into biologically relevant subpopulations remain open challenges.

AUTHORS

Alexandre Dufour (adufour@pasteur.fr) received the M.Sc. degree in artificial intelligence and pattern recognition from Paris Pierre and Marie Curie University, France, in 2004, and the Ph.D. degree in image processing from Paris Descartes University, France, in 2007. From 2005 to 2007, he was a junior scientist at the Institut Pasteur of Seoul, South Korea. Since 2008, he has been with the Institut Pasteur of Paris, France, where he focuses on three-dimensional deformable models and shape modeling approaches to study morphological and dynamic processes at the cellular and multicellular level. He is also actively involved in open-source software development for reproducible research in bioimage informatics. He is a member of the IEEE Signal Processing Society, the IEEE Bioimaging and Signal Processing Technical Committee, and served as area chair and associate editor in multiple conferences including the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing; the 2014 IEEE International Conference on Image Processing; the European Signal Processing Conference; and the International Symposium on Biomedical Imaging.

Tzu-Yu Liu (tyliu@eecs.berkeley.edu) received the B.S. degree from the National Taiwan University in 2007 and the Ph.D. degree from the University of Michigan, Ann Arbor, in 2013, both in electrical engineering. Her research interests include statistical learning from high-dimensional and small sample size problems, optimization, structured variable selection, and their applications to biomedical data. She is a postdoctoral researcher at the University of California, Berkeley, where she is focusing on computational biology and machine learning.

Christel Ducroz (christel.ducroz89@gmail.com) received the M.Sc. degree in automatic control, signal, and image in 2011 from Ecole Centrale de Nantes, France. She is currently in her fifth year of medical school. From 2010 to 2012, she was a research assistant with the Quantitative Image Analysis Unit at Institut Pasteur, France, working on three-dimensional cell-shape characterization.

Robin Tournemenne (rtournemenne@gmail.com) received the M.Sc. degree in biomedical engineering in 2013 from RWTH in Aachen, Germany, and graduated from the engineering school Ecole Centrale de Nantes, France, the same year. He is currently pursuing his Ph.D. degree at the Institut de Recherche en Communications et Cybernétique de Nantes, France, where he focuses on physical modeling applied to sound synthesis.

Beryl Cummings (bcummings@g.harvard.edu) received the B.Sc. degree in biology and a certificate in genome sciences and policy from Duke University in Durham, North Carolina. She is currently pursuing her Ph.D. degree in the Harvard Biological and Biomedical Sciences Program, Genetics and Genomics Division in Boston, Massachusetts. Her research is focused on utilizing three-dimensional microscopy to elucidate the relationship between phenotypic plasticity in parasites and their disease pathogenesis.

Roman Thibeaux (roman.thibeaux@pasteur.fr) received the B.Sc. degree in molecular and cellular biology in 2006 and the M.Sc. degree in genetics in 2008 from Paris Denis Diderot

University, France. He then pursued his Ph.D. studies at the Institut Pasteur in Paris, France. Since 2011, he has been a postdoctoral research fellow pursuing his research on tissue invasion and places a particular emphasis on determining the role of host inflammatory response during this invasive process. His research interests include host–pathogen interactions, parasitology, virology, microbiology, extracellular matrix biology, tissue engineering, live imaging, and two-photon microscopy.

Nancy Guillen (nguillen@pasteur.fr) is Research Director 1 at the National Center for Scientific Research, France. She is head of the Cell Biology of Parasitism Unit-INSERM U786, Department of Cell Biology and Infection, Institut Pasteur, Paris, France. She is a member of the Scientific Council, Institut Pasteur, Animal Health Division, Institut National de Recherche Agronomique, and is scientific coordinator of the Pasteur-Weizmman Research Council. Her research interests include the cellular and molecular biology studies of the pathogenic process in *Entamoeba histolytica*, the agent of human amoebiasis, parasitology, cytoskeleton, myosin, cell motility, transcriptome, and physiopathology.

Alfred Hero III (hero@eecs.umich.edu) received the B.S. degree (summa cum laude) from Boston University in 1980 and the Ph.D. degree from Princeton University in 1984, both in electrical engineering. Since 1984, he has been with the University of Michigan, Ann Arbor, where he is the R. Jamison and Betty Williams Professor of Engineering. He received the 2011 University of Michigan Distinguished Faculty Achievement Award. He has been plenary and keynote speaker at major workshops and conferences. He has received numerous Best Paper Awards. He received an IEEE Signal Processing Society (SPS) Meritorious Service Award (1998), the IEEE Third Millenium Medal (2000), an IEEE SPS Distinguished Lecturership (2002), and an IEEE SPS Technical Achievement Award (2014). He was president of the IEEE SPS from 2006 to 2007. He was on the Board of Directors of the IEEE (2009–2011). His recent research interests are in statistical signal processing, machine learning, and the modeling and analysis of high-dimensional spatiotemporal data. He is a Fellow of the IEEE.

Jean-Christophe Olivo-Marin (jcolivo@pasteur.fr) received the Ph.D. and Habilitation à Diriger des Recherchesdegrees in optics and signal processing from the Institut d'Optique Théorique et Appliquée, University of Paris-Orsay, France. He is the head of the Bioimage Analysis Unit and the director of the Center for Innovation and Technological Research at Institut Pasteur, Paris. He has chaired the Cell Biology and Infection Department and was a cofounder of the Institut Pasteur Korea, Seoul, South Korea. His research interests are in image analysis of multidimensional microscopy images, computer vision, and motion analysis for cellular dynamics, and in multidisciplinary approaches for biological imaging. He is a Fellow of the IEEE, an IEEE Signal Processing Society Distinguished Lecturer, chair of the IEEE International Symposium on Biomedical Imaging Steering Committee, a senior area editor of *IEEE Signal Processing Letters*, and a member of the editorial boards of *Medical Image Analysis* and *BMC Bioinformatics*. He was the general chair of the 2008 IEEE International Symposium on Biomedical Imaging.

REFERENCES

[1] A. D. Doyle, R. J. Petrie, M. L. Kutys, and K. M. Yamada, "Dimensions in cell migration," *Curr. Opin. Cell Biol.*, vol. 25, no. 5, pp. 642–649, 2013.

[2] M. M. Frigault, J. Lacoste, J. L. Swift, and C. M. Brown, "Live-cell microscopy—Tips and tools," *J. Cell Sci.*, vol. 122, pp. 753–767, Mar. 2009.

[3] Y. Le Montagner, E. Angelini, and J.-C. Olivo-Marin, "An unbiased risk estimator for image denoising in the presence of mixed Poisson-Gaussian noise," *IEEE Trans. Image Processing*, vol. 23, no. 3, pp. 1255–1268, 2014.

[4] M. a. Bruce and M. J. Butte, "Real-time GPU-based 3D deconvolution," *Opt. Express*, vol. 21, pp. 4766–4773, Feb. 2013.

[5] M. Arigovindan, J. C. Fung, D. Elnatan, V. Mennella, Y.-H. M. Chan, M. Pollard, E. Branlund, J. W. Sedat, and D. a. Agard, "High-resolution restoration of 3D structures from widefield images with extreme low signal-to-noise-ratio," in *Proc. Natl. Acad. Sci. USA*, vol. 110, Oct. 2013, pp. 17344–17349.

[6] M. Marim, E. Angelini, and J.-C. Olivo-Marin, "A compressed sensing approach for biological microscopic image processing," in *Proc. IEEE Int. Symp. Biomedical Imaging*, 2009, no. 1, pp. 1374–1377.

[7] P. Ye, J. L. Paredes, G. R. Arce, Y. Wu, C. Chen, and D. W. Prather, "Compressive confocal microscopy," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Apr. 2009, pp. 429–432.

[8] L. Shao, P. Kner, E. H. Rego, and M. G. L. Gustafsson, "Super-resolution 3D microscopy of live whole cells using structured illumination," *Nat. Methods*, vol. 8, pp. 1044–1046, Dec. 2011.

[9] M. Ting, R. Raich, and A. O. Hero, "Sparse image reconstruction for molecular imaging," *IEEE Trans. Image Processing*, vol. 18, no. 6, pp. 1215–1227, 2009.

[10] R. Tomer, K. Khairy, F. Amat, and P. J. Keller, "Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy," *Nat. Methods*, vol. 9, pp. 755–63, July 2012.

[11] E. Meijering, "Cell segmentation: 50 years down the road," *IEEE Signal Processing Mag.*, vol. 29, no. 5, pp. 140–145, 2012.

[12] M. Maska, V. Ulman, D. Svoboda, P. Matula, P. Matula, C. Ederra, A. Urbiola, T. España, S. Venkatesan, D. M. W. Balak, P. Karas, T. Bolcková, M. Streitová, C. Carthel, S. Coraluppi, N. Harder, K. Rohr, K. E. G. Magnusson, J. Jaldén, H. M. Blau, O. Dzyubachyk, P. Krízek, G. M. Hagen, D. Pastor-Escuredo, D. Jimenez-Carretero, M. J. Ledesma-Carbayo, A. Muñoz Barrutia, E. Meijering, M. Kozubek, and C. Ortiz-de Solorzano, "A benchmark for comparison of cell tracking algorithms," *Bioinformatics*, vol. 30, no. 11, pp. 1609-1617, June 2014.

[13] Z. Pincus and J. a. Theriot, "Comparison of quantitative methods for cell-shape analysis," *J. Microsc.*, vol. 227, pp. 140–156, Aug. 2007.

[14] T. Peng and R. Murphy, "Image-derived, three-dimensional generative models of cellular organization," *Cytom. Part A*, vol. 79, no. 5, pp. 383–391, 2011.

[15] K. Khairy, J. Foo, and J. Howard, "Shapes of red blood cells: Comparison of 3D confocal images with the bilayer-couple model," *Cell. Mol. Bioeng.*, vol. 1, pp. 173–181, Sept. 2010.

[16] C. Ducroz, J.-C. Olivo-Marin, and A. Dufour, "Characterization of cell shape and deformation in 3D using spehrical harmonics," in *Proc. IEEE Int. Symp. Biomedical Imaging*, Barcelona, Spain, 2012, pp. 1–4.

[17] C.-j. Du, J. G. Ferguson, P. T. Hawkins, L. R. Stephens, and T. Bretschneider, "Local shape representation in 3D: From weighted spherical harmonics to spherical wavelets," in *Proc. British Machine Vision Conf.*, 2012, pp. 1–12.

[18] C. Ducroz, J.-C. Olivo-Marin, and A. Dufour, "Automatic detection of 3D cell protrusions using spherical wavelets," in *Proc. IEEE Int. Conf. Image Processing*, Melbourne, Australia, 2013, pp. 3499–3502.

[19] C.-J. Du, P. T. Hawkins, L. R. Stephens, and T. Bretschneider, "3D time series analysis of cell shape using Laplacian approaches," *BMC Bioinform.*, vol. 14, p. 296, Jan. 2013.

[20] R. Tournemenne, C. Ducroz, J.-C. Olivo-Marin, and A. Dufour, "3D shape analysis using overcomplete spherical wavelets: Application to bleb detection in cell biology," in *Proc. IEEE Int. Symp. Biomedical Imaging*, 2014, pp. 1–4.

[21] S. Torcida, S. I. Perez, and P. N. Gonzalez, "An integrated approach for landmark-based resistant shape analysis in 3D," *Evol. Biol.*, vol. 41, no. 2, pp. 351–366, 2014.

[22] T. R. Jones, A. E. Carpenter, M. R. Lamprecht, J. Moffat, S. J. Silver, J. K. Grenier, A. B. Castoreno, U. S. Eggert, D. E. Root, P. Golland, and D. M. Sabatini, "Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning," *Proc. Natl. Acad. Sci. USA*, vol. 106, pp. 1826–1831, Feb. 2009.

[23] C. Sommer and D. W. Gerlich, "Machine learning in cell biology—Teaching computers to recognize phenotypes," *J. Cell Sci.*, vol. 126, pp. 5529–5539, Dec. 2013.

[24] T. E. Buck, J. Li, G. K. Rohde, and R. F. Murphy, "Toward the virtual cell: Automated approaches to building models of subcellular organization 'learned' from microscopy images," *BioEssays: News Rev. Mol. Cell. Dev. Biol.*, vol. 34, pp. 791–799, Sept. 2012.

[25] D. C. Ince, L. Hatton, and J. Graham-Cumming, "The case for open computer programs," *Nature*, vol. 482, pp. 485–488, Feb. 2012.

[26] P. Vandewalle, J. Kovacevic, and M. Vetterli, "Reproducible research in signal processing," *IEEE Signal Processing Mag.*, vol. 26, pp. 37–47, May 2009.

[27] N. Chenouard, I. Smal, F. de Chaumont, M. Maška, I. F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, A. R. Cohen, W. J. Godinez, K. Rohr, Y. Kalaidzidis, L. Liang, J. Duncan, H. Shen, Y. Xu, K. E. G. Magnusson, J. Jaldén, H. M. Blau, P. Paul-Gilloteaux, P. Roudot, C. Kervrann, F. Waharte, J.-Y. Tinevez, S. L. Shorte, J. Willemse, K. Celler, G. P. van Wezel, H.-W. Dan, Y.-S. Tsai, C. O. de Solórzano, J.-C. Olivo-Marin, and E. Meijering, "Objective comparison of particle tracking methods," *Nat. Methods*, vol. 11, pp. 281–289, Mar. 2014.

[28] T. Lämmermann and M. Sixt, "Mechanical modes of 'amoeboid' cell migration," *Curr. Opin. Cell Biol.*, vol. 21, pp. 636–644, Oct. 2009.

[29] M. Bergert, S. D. Chandradoss, R. a. Desai, and E. Paluch, "Cell mechanics control rapid transitions between blebs and lamellipodia during migration," *Proc. Natl. Acad. Sci. USA*, vol. 109, pp. 14434–14439, Sept. 2012.

[30] A. Dufour, R. Thibeaux, E. Labruyère, N. Guillén, and J.-C. Olivo-Marin, "3D active meshes: Fast discrete deformable models for cell tracking in 3D time-lapse microscopy," *IEEE Trans. Image Processing*, vol. 20, pp. 1925–1937, July 2011.

[31] R. Delgado-Gonzalo, N. Chenouard, and M. Unser, "Spline-based deforming ellipsoids for interactive 3D bioimage segmentation," *IEEE Trans. Image Processing*, vol. 22, pp. 3926–3940, Oct. 2013.

[32] R. Thibeaux, A. Dufour, P. Roux, M. Bernier, A.-C. Baglin, P. Frileux, J.-C. Olivo-Marin, N. Guillén, and E. Labruyère, "Newly visualized fibrillar collagen scaffolds dictate Entamoeba histolytica invasion route in the human colon," *Cell. Microbiol.*, vol. 14, pp. 609–621, Jan. 2012.

[33] F. L. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*, vol. 10. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[34] J. Ashburner and K. J. Friston, "Voxel-based morphometry—The methods," *NeuroImage*, vol. 11, pp. 805–821, June 2000.

[35] S. M. Weinberg, N. C. Andreasen, and P. Nopoulos, "Three-dimensional morphometric analysis of brain shape in nonsyndromic orofacial clefting," *J. Anatomy*, vol. 214, pp. 926–936, June 2009.

[36] P. Perakis, G. Passalis, T. Theoharis, and I. A. Kakadiaris, "3D facial landmark detection under large yaw and expression variations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1552–1564, 2013.

[37] I. L. Dryden and K. V. Mardia, *Statistical Shape Analysis*. New York: Wiley, 1998.

[38] J. Neustupa and L. Hodac, "Changes in shape of the coenobial cells of an experimental strain of Pediastrum duplex var. duplex (Chlorophyta) reared at different pHs," *Preslia*, vol. 77, pp. 439–452, 2005.

[39] Y. Xiong, C. Kabacoff, J. Franca-Koh, P. N. Devreotes, D. N. Robinson, and P. a. Iglesias, "Automated characterization of cell shape changes during amoeboid motility by skeletonization," *BMC Syst. Biol.*, vol. 4, p. 33, Jan. 2010.

[40] D. Tsygankov, C. G. Bilancia, E. a. Vitriol, K. M. Hahn, M. Peifer, and T. C. Elston, "CellGeo: A computational platform for the analysis of shape changes in cells with complex geometries," *J. Cell Biol.*, vol. 204, pp. 443–460, Feb. 2014.

[41] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[42] M. Lozano and F. Escolano, "Protein classification by matching and clustering surface graphs," *Pattern Recognit.*, vol. 39, no. 4, pp. 539–551, 2006.

[43] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inform. Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[44] P. Yap, R. Paramesran, and S. Ong, "Image analysis by Krawtchouk moments," *IEEE Trans. Image Processing*, vol. 12, pp. 1367–1377, Jan. 2003.

[45] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Mag.*, vol. 16, no. 6, pp. 22–38, 1999.

[46] L. Tweedy, B. Meier, J. Stephan, D. Heinrich, and R. G. Endres, "Distinct cell shapes determine accurate chemotaxis," *Sci. Rep.*, vol. 3, p. 2606, Jan. 2013.

[47] C. Brechbühler, G. Gerig, and O. Kübler, "Parametrization of closed surfaces for 3-D shape description," *Comput. Vis. Image Understand.*, vol. 61, no. 2, pp. 154–170, 1995.

[48] L. Shen and F. Makedon, "Spherical mapping for processing of 3D closed surfaces," *Image Vis. Comput.*, vol. 24, pp. 743–761, July 2006.

[49] J.-P. Antoine and P. Vandergheynst, "Wavelets on the 2-sphere: A group-theoretical approach," *Appl. Computat. Harm. Anal.*, vol. 7, pp. 262–291, Nov. 1999.

[50] D. W. Ritchie and G. J. L. Kemp, "Fast computation, rotation, and comparison of low resolution spherical harmonic molecular surfaces," *J. Computat. Chem.*, vol. 20, no. 4, pp. 383–395, 1999.

[51] R. J. Morris, R. J. Najmanovich, A. Kahraman, and J. M. Thornton, "Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons," *Bioinformatics (Oxford, England)*, vol. 21, pp. 2347–2355, May 2005.

[52] M. Styner, I. Oguz, S. Xu, C. Brechbuhler, D. Pantazis, J. Levitt, M. Shenton, and G. Gerig, "Framework for the statistical shape analysis of brain structures using SPHARM-PDM," *Insight J.*, vol. 1071, pp. 242–250, 2006.

[53] S. Singh, F. Janoos, T. Pecot, E. Caserta, K. Huang, J. Rittscher, G. Leone, and R. Machiraju, "Non-parametric analysis of cell phenotypes," in *Proc. MICCAI*, part II, pp. 343–351, 2011.

[54] B. T. T. Yeo, W. Ou, and P. Golland, "On the construction of invertible filter banks on the 2-sphere," *IEEE Trans. Image Processing*, vol. 17, pp. 283–300, Mar. 2008.

[55] M. Niethammer, M. Reuter, F. Wolter, S. Bouix, N. Peinecke, M.-S. Koo, and M. Shenton, "Global medical shape analysis using the Laplace-Beltrami spectrum," in *Proc. MICCAI*, 2007, vol. 10, pp. 850–857.

[56] P. Rämö, R. Sacher, B. Snijder, B. Begemann, and L. Pelkmans, "CellClassifier: Supervised learning of cellular phenotypes," *Bioinformatics*, vol. 25, pp. 3028–3030, Nov. 2009.

[57] M. Held, M. H. A. Schmitz, B. Fischer, T. Walter, B. Neumann, M. H. Olma, M. Peter, J. Ellenberg, and D. W. Gerlich, "CellCognition: Time-resolved phenotype annotation in high-throughput live cell imaging," *Nat. Methods*, vol. 7, no. 9, pp. 747–754, 2010.

[58] G. Pau, F. Fuchs, O. Sklyar, M. Boutros, and W. Huber, "EBImage—An R package for image processing with applications to cellular phenotypes," *Bioinformatics*, vol. 26, pp. 979–981, Apr. 2010.

[59] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, and J.-C. Olivo-Marin, "Icy: An open bioimage informatics platform for extended reproducible research," *Nat. Methods*, vol. 9, no. 7, pp. 690–696, 2012.

[60] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: An open-source platform for biological-image analysis," *Nat. Methods*, vol. 9, pp. 676–682, July 2012.

[61] J. Ollion, J. Cochennec, F. Loll, C. Escudé, and T. Boudier, "TANGO: A generic tool for high-throughput 3D image analysis for studying nuclear organization," *Bioinformatics (Oxford, England)*, vol. 29, pp. 1840–1841, July 2013.

[62] C. M. Bishop, *Pattern Recognition and Machine Learning*, vol. 4. New York: Springer, 2006.

[63] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. New York: Springer, 2009.

[64] X. Chen, X. Zhou, and S. T. C. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 762–766, Apr. 2006.

[65] E. Segal, C. B. Sirlin, C. Ooi, A. S. Adler, J. Gollub, X. Chen, B. K. Chan, G. R. Matcuk, C. T. Barry, H. Y. Chang, and M. D. Kuo, "Decoding global gene expression programs in liver cancer by noninvasive imaging," *Nat. Biotechnol.*, vol. 25, pp. 675–680, June 2007.

[66] C. Sommer, C. Straehle, U. Kothe, and F. A. Hamprecht, "Ilastik: Interactive learning and segmentation toolkit," in *Proc. IEEE Int. Symp. Biomedical Imaging*, 2011, pp. 230–233.

[67] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.

[68] C. Ducroz, J.-C. Olivo-Marin, and A. Dufour, "Spherical harmonics based extraction and annotation of cell shape in 3D time-lapse microscopy sequences," in *Proc. IEEE Engineering in Medicine and Biology Conf.*, Boston, MA, 2011, pp. 6619–6622.

[69] Q. Zhong, A. G. Busetto, J. P. Fededa, J. M. Buhmann, and D. W. Gerlich, "Unsupervised modeling of cell morphology dynamics for time-lapse microscopy," *Nat. Methods*, vol. 9, pp. 711–713, July 2012.

[70] K. Masood, N. Rajpoot, K. Rajpoot, and H. Qureshi, "Hyperspectral colon tissue classification using morphological analysis," in *Proc. Int. Conf. Emerging Technologies*, 2006, pp. 735–741.

[71] B. Neumann, T. Walter, J.-K. Hériché, J. Bulkescher, H. Erfle, C. Conrad, P. Rogers, I. Poser, M. Held, U. Liebel, C. Cetin, F. Sieckmann, G. Pau, R. Kabbe, A. Wünsche, V. Satagopam, M. H. A. Schmitz, C. Chapuis, D. W. Gerlich, R. Schneider, R. Eils, W. Huber, J.-M. Peters, A. A. Hyman, R. Durbin, R. Pepperkok, and J. Ellenberg, "Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes," *Nature*, vol. 464, no. 7289, pp. 721–727, 2010.

[72] K. W. Eliceiri, M. R. Berthold, I. G. Goldberg, L. Ibáñez, B. S. Manjunath, M. E. Martone, R. F. Murphy, H. Peng, A. L. Plant, B. Roysam, N. Stuurman, N. Stuurmann, J. R. Swedlow, P. Tomancak, and A. E. Carpenter, "Biological imaging software tools," *Nat. Methods*, vol. 9, pp. 697–710, July 2012.

[73] C. Conrad, A. Wünsche, T. H. Tan, J. Bulkescher, F. Sieckmann, F. Verissimo, A. Edelstein, T. Walter, U. Liebel, R. Pepperkok, and J. Ellenberg, "Micropilot: Automation of fluorescence microscopy-based imaging for systems biology," *Nat. Methods*, vol. 8, pp. 246–249, Mar. 2011.

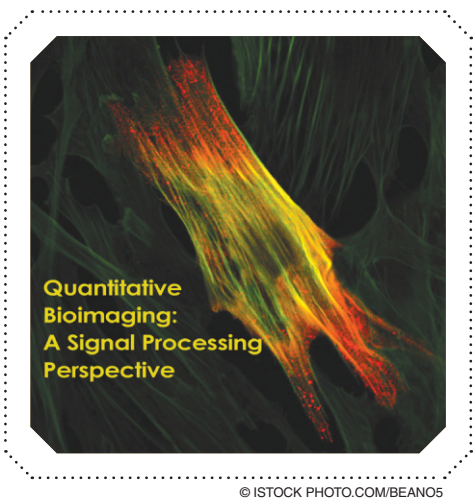
[74] T. Y. Liu, A. Wiesel, and A. O. Hero, "A sparse multiclass classifier for biomarker screening," in *Proc. IEEE Global Conf. Signal and Information Processing (GlobalSIP)*, 2013, pp. 77–83.

[SP]

[Ricard Delgado-Gonzalo, Virginie Uhlmann, Daniel Schmitter, and Michael Unser]

Snakes on a Plane

[A perfect snap for bioimage analysis]



© ISTOCK PHOTO.COM/BEANO5

In recent years, there has been an increasing interest in getting a proper quantitative understanding of cellular and molecular processes [1], [2]. One of the major challenges of current biomedical research is to characterize not only the spatial organization of these complex systems but also their spatiotemporal relationships [3], [4]. Microscopy has matured to the point that it enables sensitive time-lapse imaging of cells in vivo and even of single molecules [5], [6]. Making microscopy more quantitative brings important scientific benefits in the form of improved performance and reproducibility. This has been fostered by the development of technological achievements such as high-throughput microscopy. A direct consequence is that the size and complexity of image data are increasing. Time-lapse experiments commonly generate hundreds to thousands of images, each containing hundreds of objects to be analyzed [7]. These data often cannot be analyzed manually because the manpower required would be too extensive, which calls for automated methods for the analysis of biomedical images. Such computerized extraction of quantitative information out of the rapidly expanding amount of acquired data remains a major challenge. The development of the related algorithms is nontrivial and is one of the most active fronts in the new field of bioimage informatics [8]–[11]. Segmenting thousands of individual biological objects and tracking them over time is remarkably difficult. A typical algorithm will need to be tuned to the imaging modality and will have to cope with the fact that cells can be tightly packed and may appear in various configurations, making them difficult to segregate.

Digital Object Identifier 10.1109/MSP.2014.2344552
Date of publication: 5 December 2014

SNAKES (IN A NUTSHELL)

Deformable models have gained popularity in segmentation and tracking applications [12] since they provide an excellent tradeoff between flexibility and efficiency. Within this category, active contours, also called *snakes*, are very popular tools for image segmentation. It can indeed be observed that, among the roughly 12,000 publications about active contours indexed by Thomson Reuters' "Web of Science" (formerly known as the "Web of Knowledge"), about 2,000 have appeared within the past three years.

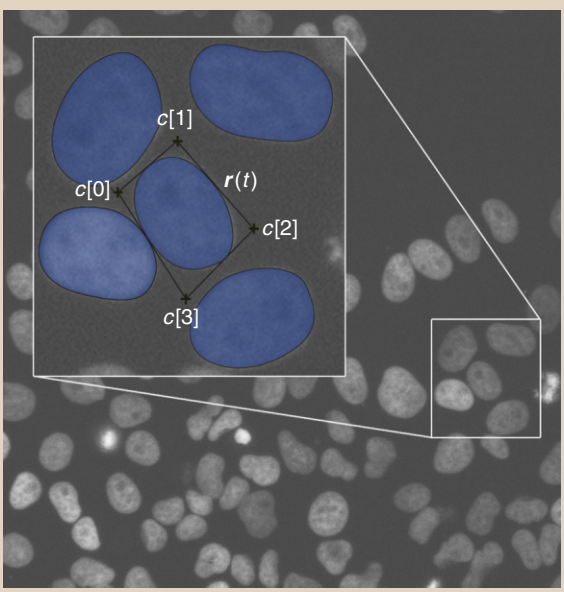
In essence, an active contour is a curve that evolves within an image from some initial position toward the boundary of the object of interest, in our case, a biological target. In Figure 1, we show our favorite snakes segmenting a group of Helacyton gartleri (HeLa) nuclei within a microscopic image. The initial position of the snake is usually specified by the user or is otherwise provided by an auxiliary rough detection algorithm. The evolution of the snake is formulated as a minimization problem. The associated cost function is usually referred to as *snake energy*.

Snakes have become popular because they facilitate user interaction, not only when specifying initial positions but also during the segmentation process. The user is typically provided with tools to easily control the shape of the snake. Research on active contours has been fruitful and resulted in many variants. They usually differ in the type of representation and in the choice of the energy term. The choice and design of both the representation and energy of the snake highly depends on the application and image modality.

CURRENT CHALLENGES

In the particular case of active contours for bioimage segmentation, several issues can be identified [13], [14]. In the following, we describe ongoing challenges for snake algorithms.

- *Robustness and stability*: Active contours are intended to be practical. It is, therefore, of utmost importance to ensure robustness under real-life imaging conditions where sources of image degradation include noise, bias field, and low contrast. Such image artifacts often appear in biological data. Guaranteeing stability in their presence is a major challenge to ensure the usability of algorithms.
- *Multitarget interaction*: Bioimages rarely feature only one object of interest. In most situations, several targets, as well as undesired structures such as water bubbles or dust particles, are present in the image. Hence, a major requirement for efficient segmentation is to avoid confusion between nearby targets. Active contours must be designed to be able to discriminate between close or even overlapping objects, which is difficult to achieve.
- *Flexibility and shape priors*: Another requirement that snake algorithms find difficult to satisfy is to balance flexibility (to accommodate a wide range of shapes) against sufficient constraints (to secure convergence of the optimization procedure). For instance, integrating prior knowledge about the shape to segment can be used efficiently to constrain the snake, but it often remains desirable to retain the capability to adapt to variations in size, orientation, and moderate deformations.
- *Initialization*: Snakes are semiautomated algorithms. The segmentation part is mostly automated through the optimization of the snake energy, but the initialization often relies on user input. Developing automated methods to conveniently initialize active contours is an area of open research.
- *Snake–human interaction*: Despite good initialization and efficient optimization, even the best segmentation algorithm will occasionally fail to produce satisfying results. In the case of active contours, the natural abilities of snakes to accommodate user interaction will allow one to suitably modify results in a user-friendly way.
- *Computational efficiency*: It is crucial to provide reasonably fast implementations that run on standard computers, thereby enabling biologists to use them online for the analysis of bioimages.
- *Cell tracking*: The last major image analysis challenge is to reliably segment thousands of individual biological objects and to track them over time. This is far from trivial due to the dependence on the imaging modality and the fact that the cells can be tightly packed in the growth chamber and may appear in various configurations making them difficult to segregate.



[FIG1] A representation of the capabilities of parametric active contours at outlining HeLa nuclei in a fluorescence microscopic image. On the central cell, two contours are depicted: the parametric contour and the control polygon. The parametric contour $\mathbf{r}(t)$ is shown as a solid line enclosing a shaded region and corresponds to the actual segmentation result. The "+" elements represent the location of the free parameters of the model $\mathbf{c}[k]$. The latter define the control polygon through which the user interacts: this polygon can be adjusted to change the shape of the enclosed parametric curve.

SNAKE TAXONOMY

Several types of snakes have been proposed over the past few decades. They can be broadly classified in three categories: point, geodesic, and parametric snakes. In the following, we describe the strong and weak points of each representation while putting emphasis on parametric representations, which is the main focus of this article.

The first category is made of point snakes, which are constructed on the most elementary representation of discrete curves. They simply consist of an ordered collection of neighboring points within a grid [15]. In Figure 2(a), we show an example of a two-dimensional point snake overlaid on the grid associated to a discrete image model. The discrete curve is displayed as shaded pixels and here satisfies an eight-neighbor connectivity. Due to the discrete nature of the representation, the concept of smoothness makes no sense. However, an approximate measure of smoothness can still be defined and deployed in the energy functional of the snake. The principal drawback of this representation is that it relies on many parameters, even to encode simple shapes. The robustness of the overall segmentation algorithm hence suffers and results in a high computational complexity.

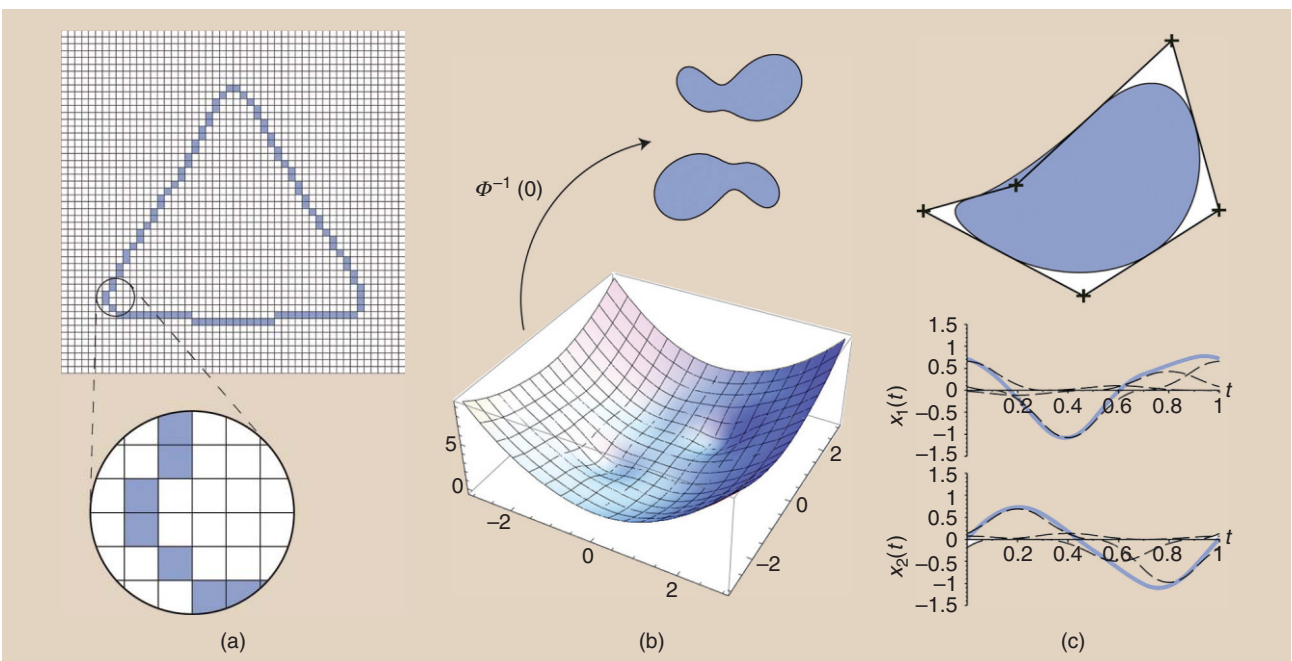
The second category is made of geodesic snakes, which have been used extensively during the past decade [16]–[22]. These snakes are implicit, being described as the zero level set of a higher-dimensional manifold. In the usual formalism, the snake contour is given by $\Phi^{-1}(0) = \{\mathbf{p} \in \mathbb{R}^n \,|\, \Phi(\mathbf{p}) = 0\}$, where Φ is a scalar function defined over the image domain. A unique characteristic of geodesic methods is that they can be extended to any number of dimensions. In Figure 2(b), we represent a set of curves generated as the result of computing $\Phi^{-1}(0)$. This approach originates from the work of Osher and Sethian, who aimed at modeling the propagation of solid–liquid interfaces with curvature-dependent speeds [23]. In this framework, the interface (or front) is represented as a closed, nonintersecting hypersurface flowing along its gradient field with either constant speed or a speed that depends on the curvature. Applying motion to it then

amounts to solve a Hamilton–Jacob-type equation written for a function in which the interface is a particular level set. This type of active contour has the interesting property of being particularly flexible in terms of topology. A single geodesic snake evolving under the appropriate energy functional is indeed able to split freely to segment disconnected objects within an image. This plasticity is especially convenient when segmenting complex shapes (perhaps involving significant protrusions) and when no prior assumption about the topology of the object is available. But useful geodesic models require many degrees of freedom. This makes it difficult to constrain shapes and can lead to overfitting in real-world imaging conditions. Another drawback of geodesic approaches is their very expensive computational needs, mainly due to the fact that they evolve a manifold with a higher number of dimensions than the actual contour to segment. In summary, geodesic snakes, based on level sets, are convenient when the shapes of the objects to segment exhibit high variability. However, geodesic snakes are suboptimal when segmenting known shapes.

The last category is made of parametric snakes. In this representation, the snake is described by some discrete set of coefficients and a continuous parameter [25], [26]. Parametric snakes are usually built in a way that ensures continuity and smoothness. Many different techniques for representing continuous curves have been published; for a review, see [27]. The continuous definition of parametric snakes suggests that the segmentation task can be conducted at arbitrary resolution, hence enabling subpixel accuracy. Another advantage is that these contours require much fewer coefficients and result in faster optimization schemes than point snakes or geodesic snakes. In the particular case of B-spline parameterizations, the computational complexity of the energy of the snake and, therefore, the speed of the optimization algorithms, is related to the degree of overlap of the basis functions [24]. This overlap is therefore a critical parameter to consider while designing parametric snakes. User interaction is usually achieved by allowing the user to specify anchor points for the curve to go through [15]. Smoothness and shape constraints can easily be introduced [28]. They are particularly suitable when the objects to segment have a reproducible shape that can be naturally encoded within the parameterization. The most relevant example is the segmentation of circles and ellipses. In medical imaging in particular, it is desirable to segment arteries and veins within tomographic slices. Because those physiological objects are tubes, their sections show up as ellipses in the image. Ellipse-like objects are also present at microscopic scales. For instance, cell nuclei are known to be nearly circular. Similarly, water drops are spherical. However, these elements deform and become elliptical or ellipsoidal when they are subject to stress forces. For approaching these problems, ultrafast parametric snakes were designed exploiting the properties of the parameterization, which led to high-throughput applications [29], [30]. The downside of parametric snakes is that the topology of the curve is imposed by the parameterization. Parametric snakes are thus less suitable than geodesic ones for accommodating changes of topology during optimization, though solutions have been proposed for specific cases [31].

SPLINE SNAKES

Curve and surface parameterizations based on Fourier descriptors [32] and uniform B-spline functions [24], [26] are popular in



[FIG2] A depiction of the three main families of active contours. (a) The discrete curve defined by a point snake over the grid associated to a discrete image model. (b) Continuous curves defined as the zero level set $\Phi^{-1}(0)$ of a scalar function Φ. (c) The continuous curve defined by a B-spline basis. The snake contour is shown as a solid line enclosing a shaded region, while the "+" elements are the control points of the spline. The parametric coordinate functions $x_1(t)$ and $x_2(t)$ are displayed in solid lines, and the dashed lines indicate the weighted basis functions.

image processing due to the existence of efficient signal processing algorithms and to their invariance to affine transformations. Among these, the B-spline curves have the unique advantage of featuring locality of control, hence, favoring user-friendly interactions with the snake. We show in Figure 2(c) an example of a curve parameterized with a B-spline basis, its spline control points, as well as its corresponding coordinate functions.

Formally, the parameterization of these snakes is expressed as a curve $\mathbf{r}(t)$ on the plane. This curve corresponds to a pair of Cartesian coordinate functions $x_1(t)$ and $x_2(t)$, where $t \in \mathbb{R}$ is a continuous parameter. The one-dimensional functions $x_1(t)$ and $x_2(t)$ are efficiently parameterized by linear combinations of suitable basis functions ϕ_k. The parametric representation of the active contour can be expressed as the vectorial equation

$$\mathbf{r}(t) = \sum_{k=0}^{M-1} \mathbf{c}[k]\,\phi_k(t), \tag{1}$$

with $\{\mathbf{c}[k]\}_{k\in\mathbb{Z}}$ a sequence of control points. The number M of control points determines the degrees of freedom in the model. Small numbers lead to constrained shapes; large numbers lead to additional flexibility and more general shapes.

In the case of spline snakes, typical bases are those derived from a compactly supported generator ϕ and its integer shifts $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$. Hence, $\phi_k(t) = \phi(Mt - k)$ in (1). In this setting, fast and stable interpolation algorithms can be used [33] to compute the curve. As an illustrative example, three commonly used functions to generate spline snakes are represented in Figure 3. Along with the plot of the functions, the curves generated by the same set of control points and each of the depicted functions following (1) are also illustrated.

A noticeable technical feature of spline snakes is their computational cost. Computing points on a spline curve is extremely efficient [34]. It can be shown that the expense for computing any point $\mathbf{r}(t_0)$ on the spline curve is proportional to the support of the basis function ϕ [35]. Moreover, the length of the support of a spline basis function is directly related to the degree of regularity of the function, and therefore to its approximation power [36]. Therefore, there is a tradeoff between computational complexity and the approximation properties of the model. The choice of the basis function ϕ is hence influenced by two aspects: the computational complexity of the resulting segmentation algorithm and the ability of the snake to adopt specific shapes and to retain smoothness. This has led to research toward obtaining minimum-support basis functions that preserve good reproduction properties [24].

It is desirable for the parametric curve to satisfy several properties. First, the representation should be unique and stable. The snake should indeed be defined by its coefficients in such a way that the unicity of representation of $x_1(t)$ and $x_2(t)$ is guaranteed and that the interpolation procedure is ensured to be numerically stable. This requirement translates into the so-called Riesz-basis condition on the generating function ϕ, for which space- and time-domain formulations exist. Second, as it is preferable for the model to be able to represent shapes irrespective of their position and orientation, invariance to affine transformations has to be enforced. This constraint yields the partition-of-unity condition, well known in approximation theory. It implies that ϕ is capable of reproducing constants; consequently, the snake gains the ability to approximate any curve when the number of control points increases. Finally, the last common prerequisite is that the curvature of the parametric snake should be a bounded function with respect to t. This imposes for the basis ϕ, or, equivalently, for each coordinate function, to be at least $C^1(\mathbb{R})$ with bounded second derivative. Starting from these three basic requirements, specific functions can be designed by imposing additional characteristics on the generator ϕ, for instance, the ability to reproduce ellipses [37], or to control the tangents [38].

Under the parametric representation, snakes made of closed curves can be described with a periodic sequence of control points. In this case, both $x_1(t)$ and $x_2(t)$ are periodic with the same period. When normalized to unity such that $\mathbf{r}(t) = \mathbf{r}(t+1)$ for all $t \in \mathbb{R}$ and divided into M segments, $\mathbf{r}(t)$ involves weighted periodized basis functions. Under some mild refinability conditions, this model naturally leads to a stationary subdivision scheme. Closed parametric snakes hence appear to be particularly convenient for the segmentation of cells and cellular components.

Alternatively, open-ended active contours can also be generated with suitable basis functions. Natural conditions on the extremities of the curve typically involve two complementary interpolation functions that provide both point-wise and tangential control [38]. Such snakes have been applied to the segmentation of ridgelike objects such as chromosomes and thin rod-shaped bacteria.

ENERGY-GUIDED SEGMENTATION

The segmentation process with snakes is formulated as an energy-minimization problem. The quality of segmentation is determined by the choice of the energy terms; it is generally agreed that specific image energies need to be defined for each particular imaging device. Kass et al. [15] originally formulated the snake energy as a linear combination of three terms:

- the *image energy*, E_{image}, which is purely data driven
- the *internal energy*, E_{int}, which ensures that the segmented region has smooth boundaries
- the *constraint energy*, E_{c}, which provides a means for the user to interact with the snake.

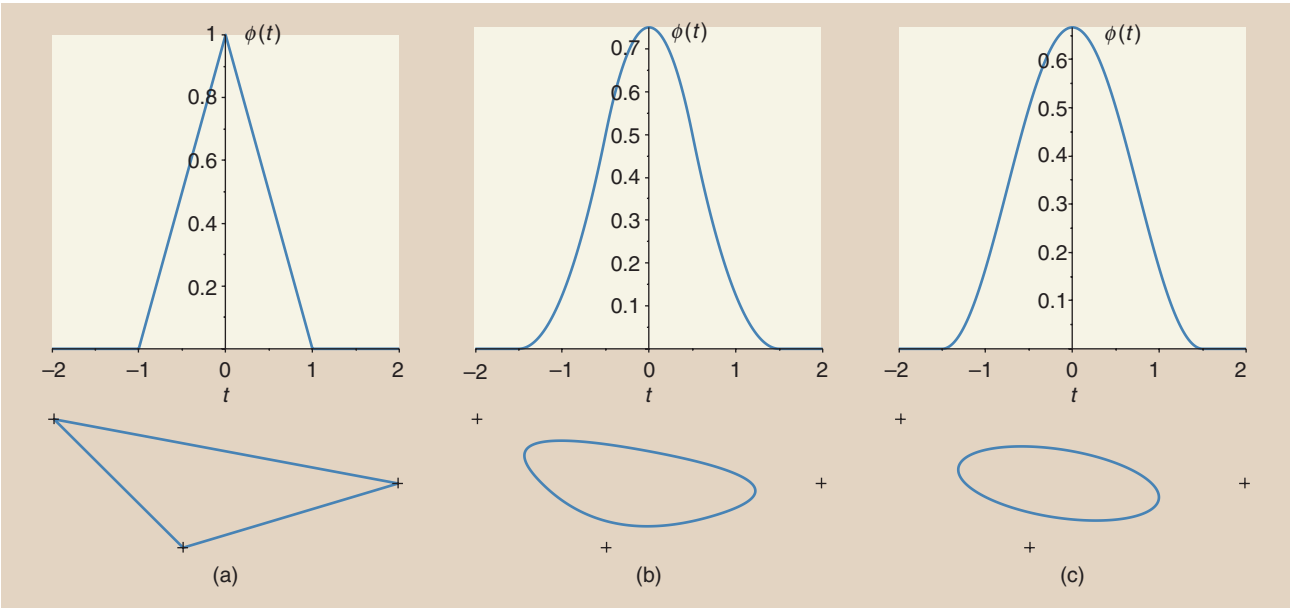
The total energy of the snake is written as

$$E_{\text{snake}}(\Theta) = E_{\text{image}}(\Theta) + E_{\text{int}}(\Theta) + E_{\text{c}}(\Theta), \tag{2}$$

where Θ encodes the snake representation (snake points, parameters, or manifolds). Then, the optimal Θ is formally obtained as

$$\Theta_{\text{opt}} = \arg\min_{\Theta} E_{\text{snake}}(\Theta).$$

The energy-minimization process is nothing but an optimization procedure, where the snake representation is iteratively updated so as to reach the minimum of the energy function from a starting position. This starting position is usually specified by the user. However, many application-dependent techniques exist capable of providing a first estimate of the position of the target (e.g., thresholding, difference of Gaussians (DoG) filtering, match



[FIG3] A comparison between the curves generated by three different basis functions $\phi(t)$ using the same set of control points. (a) The linear B-spline, (b) quadratic B-spline, and (c) minimum-support ellipse-reproducing spline [24].

filtering, watershed transform, and Hough transform). In Figure 4, we illustrate the convergence of our favorite snake segmenting a cell. Many methods exist to minimize the energy functional (e.g., gradient descent, partial differential equations, and dynamic programming), and each optimization scheme is usually linked to a particular representation of the snake. In the case of spline snakes, the optimization refers to iteratively updating the position of the control points, and the fastest approaches usually are gradient-based methods. These techniques require a smooth energy functional as the optimizing process relies on the computation of the partial derivative of the energy with respect to the control points, $(\partial E_{\text{snake}})/(\partial c[k])$. When splines basis functions are involved, the cost of this calculation is proportional to the degree of overlap of ϕ.

The image energy is the most important of the three terms in (2) since it is the one that guides the snake to the object of interest. Traditional snakes rely on edge maps derived from the image [15], [28]. These edge-based energies can provide a good localization of the contour of the object to segment. However, they have a narrow basin of attraction, making their success depend on the quality of their initialization. Several authors have developed alternative solutions to this lack of robustness. The most important ones are the introduction of balloon forces [41], the introduction of gradient vector-fields defined everywhere on the image domain [42], or multiresolution approaches [26], [43]. Image energies can also use statistical information to distinguish different homogeneous regions [25], [30]. The region-based energies have a larger basin of attraction and can converge even if explicit edges are not present [44]. However, they provide worse localization than edge-based image energies. Image energies often require the computation of surface integrals, which are computationally expensive. Applying Green's theorem to convert them into line integrals provides a way to drastically decrease the computational load [35].

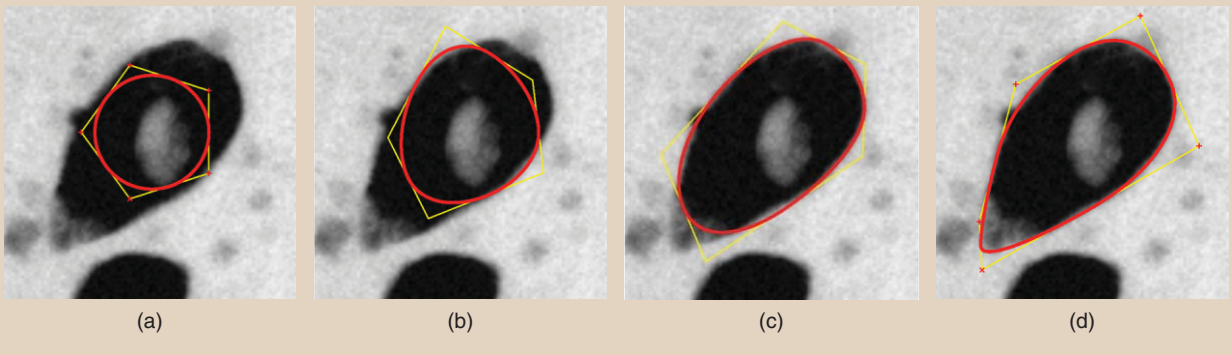
The internal energy is responsible for ensuring the smoothness of the snake. It combines the length of the contour and the curvature of the snake [15] and still corresponds to the most widely used [45]. Some authors also incorporate prior knowledge as shape constraints in this energy [59].

The constraint energy provides a means for the user to interact with the snake. Usually, this is obtained by introducing an energy functional that behaves as virtual springs that pull the snake toward the desired points [35]. Some implementations obviate the constraint energy while accommodating user interaction as hard constraints [37].

The parameters regulating the specifics of each energy and the tradeoff between them are usually specified by the user. In automatic pipelines, the full segmentation algorithm is trained on real images for a particular application. This usually provides a range of acceptable values that exchange robustness and accuracy in the overall segmentation algorithm.

Instead of (2), an alternative minimization framework is the multipurpose Mumford–Shah functional. In this framework, the image is modeled as a piecewise-smooth function. The functional penalizes the distance between the model and the input image, the lack of smoothness of the model within the subregions, and the length of the boundaries of the subregions. This approach is quite popular in the context of geodesic snakes [44].

The optimization process can sometimes lead to self-intersecting snakes. This phenomenon may arise when the image energy forces some control points to move faster than others. This compromises the approach based on Green's theorem, which assumes non-self-intersecting structures and makes some snake variants unsuitable for segmenting the edges of filament-like structures (e.g., axons). An extensive body of research can be found on the intersection problem, with numerous articles presenting different approaches for the intersection of free-form curves and



[FIG4] The typical evolution of a parametric active contour during optimization. (a) A five-control point snake is first initialized as a circle roughly in the center of the target cell. Upon optimization, (b) and (c), the contour is automatically deformed and monotonically converges from its initial position to a minimizer of the snake energy. (d) At convergence, the snake precisely outlines the target cell.

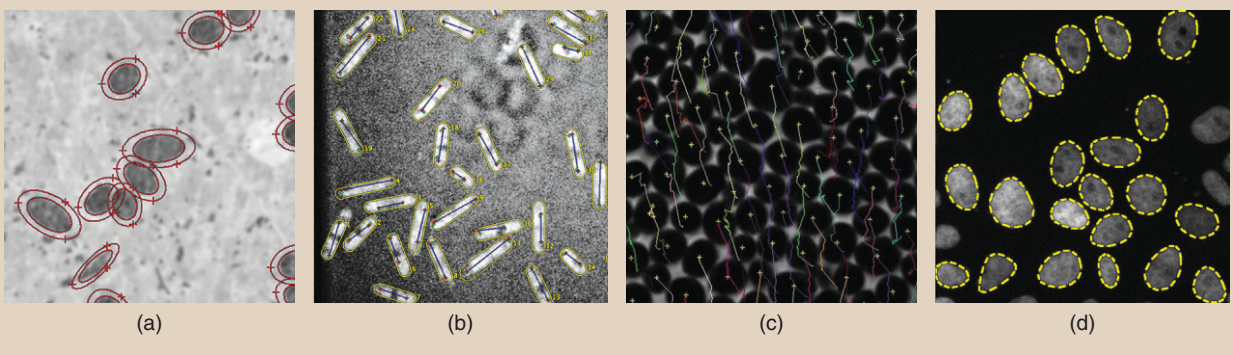
surfaces [46]. When a self-intersection is detected, there are several ways to proceed. Some authors split their shape descriptor in a way that new smaller snakes are born [31]. Others preserve the topology by introducing self-repulsive forces [47] or opting to stop the optimization routine and ask for user assistance [37].

OPEN-SOURCE FRAMEWORK

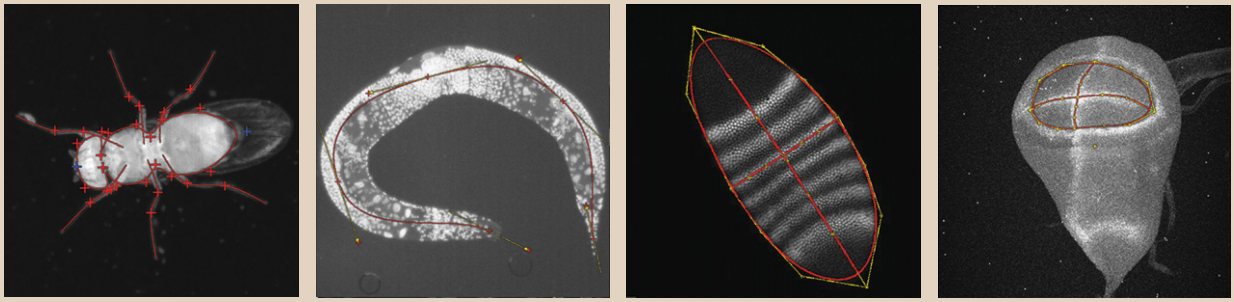
In the context of bioimaging, an interactive and user-friendly implementation of the framework is crucial. Since parametric snakes are completely defined by their control points and the generator basis function, the intuitive manipulation of the snakes can be achieved by letting the user directly move the control points with mouse-dragging gestures [29], [37]. Various ready-to-use implementations of parametric snakes are freely available as plug-ins for most popular bioimage platforms such as Fiji [10], Icy [11], and ImageJ [48]. More recently, some also have been made available for tablets. Other solutions consist of providing libraries that have active contour models integrated, such as the Insight Toolkit that can be used to develop custom applications [49].

BIOMEDICAL APPLICATIONS

The advantageous characteristics of snakes have attracted the attention of many researchers working with biological imaging data. For example, a biologist might require the segmentation of structures at a large scale (see Figure 5), meaning that an automatic application based on high-throughput analysis is required. Meanwhile, if a few but large objects need to be segmented where small details are important, an interactive scheme might be more suitable (see Figure 6).



[FIG5] Applications of parametric active contours to the segmentation and tracking. (a) The egmentation of steel needles within a cross-section of a block of concrete. The needles are cylindrical rods of identical diameter and length. Since the intersection of the plane of the cut with a cylindrical needle takes the shape of an ellipse, ovuscules are particularly appropriate tools for the task. The outer red contour represents an enclosing shell around the snake, whose volume is used to compute a region-based energy. This energy discriminates an object from its background by maximizing the contrast between the intensity of the data averaged within the volume enclosed by the snake (i.e., inner contour) and the intensity of the data averaged within the volume enclosed by the shell [30]. (b) The segmentation of rod-shaped yeast cells of the type *Saccharomyces Cervisiae* at a large scale. The segmentation problem is coupled with the task of detecting fluorescent proteins inside the cells (red/green dots). First, the red fluorescent proteins are detected at the poles of each cell. The blue line connecting them indicates the orientation of the cell, which is used for the initialization of the snake. Because the cells all have the same shape (but different sizes), prior knowledge was integrated in the design of the snake [39]. (c) The segmentation and tracking of a dense set of yeast cells within phase-contrast microscopic images. In this application, the tracking was performed using a graph-based algorithm that matched snakes across time maintaining temporal consistency (i.e., smooth trajectories). The snake-defining nodes were used as low-dimensionality descriptor for the inner states of the Bayesian tracker [4], [40]. (d) The segmentation of HeLa nuclei that express fluorescent core histone 2B on an RNAi live cell array [37].



[FIG6] Applications of parametric active contours to the segmentation of structures in organisms. The "+" elements represent the control points, and their location define the effective contour of the snakes. (a) The segmentation of the body and legs of a *Drosophila* fly. (b) The detection of the anteroposterior axis of a *Caenorhabditis elegans* with nuclear staining [38]. (c) The detection of the structure of the embryo of a *Drosophila* fly within a stack of even-skipped confocal images [58]. (d) The detection of the structure of the wing pouch of a *Drosophila* fly within a stack of Wg-Ptc labeled confocal images [58].

A vast literature exists on the use of snakes to segment biological structures such as biological tissues (nerve fibers [50], [51]), cell structures (mitochondria [52]), protein-based structures (actin filaments [53]), or model organisms such as zebra fish embryos [54] or *C. elegans* (see Figure 5). The versatile nature of snakes makes them suitable for problems that combine segmentation and tracking (Leukocyte tracking [55], motility analysis [56]), organelle tracking (microtubule tracking), or even the reconstruction of cell lineages [57].

ACKNOWLEDGMENTS

This work was funded by the Swiss SystemsX.ch initiative under grant 2008/005 and the Swiss National Science Foundation under grants 200020-121763 and 200020-144355.

AUTHORS

Ricard Delgado-Gonzalo (ricard.delgado@epfl.ch) received diplomas in telecommunications engineering and in mathematics from the Universitat Politècnica de Catalunya in 2006 and 2007, respectively. In 2008, he joined the Biomedical Imaging Group at the École polytechnique fédérale de Lausanne, Switzerland, where he obtained his Ph.D. degree in biomedical image processing in 2013. He currently works on the design of highly efficient image-analysis open software capable of running on portable devices with an intuitive interface oriented to final users. He received the 2013 Swiss Society for Biomedical Engineering Research Award for the best Ph.D. thesis for his contributions to the field of bioimage informatics.

Virginie Uhlmann (virginie.uhlmann@epfl.ch) was awarded the competitive Excellence Fellowship at the master's level from the École polytechnique fédérale de Lausanne (EPFL), Switzerland, from September 2011 to August 2012. She pursued her master's thesis as a visiting student in the Imaging Platform, Broad Institute, Cambridge, Massachusetts, under the supervision of Anne Carpenter. She obtained her M.S. degree in bioengineering from EPFL in September 2012 and was awarded four prizes, including third best grade point average among EPFL's M.S. graduates of 2012. She started a Ph.D. degree in 2012 with the Biomedical Imaging Group at EPFL under the direction of Michael Unser. She is currently working on an applied problem related to image segmentation and tracking, as well as on wavelet and spline theory. She is also involved in several teaching assistant duties. Her research interests include computer vision, machine learning, and life sciences.

Daniel Schmitter (daniel.schmitter@epfl.ch) received his master's degree in bioengineering and biomedical technologies from the École polytechnique fédérale de Lausanne (EPFL), Switzerland, in 2013. He was with the Advanced Clinical Imaging Technology Group, Siemens, at the Center for Biomedical Imaging, Switzerland, where he was one of the main contributors working on brain-imaging software and related image-processing algorithms. Currently, he is a Ph.D. student in the Biomedical Imaging Group at EPFL, where he is working on spline-related segmentation problems. He has developed several segmentaion and tracking software in the field of biomedical imaging.

Michael Unser (michael.unser@epfl.ch) is a professor and the director of the Biomedical Imaging Group at École polytechnique fédérale de Lausanne, Switzerland. His main research area is biomedical image processing. His interest is in sampling theories, multiresolution algorithms, wavelets, the use of splines for image processing, and, more recently, stochastic processes. He has published approximately 250 journal papers on those topics. He has held various editorial positions for *IEEE Transactions on Medical Imaging* and *IEEE Transactions on Image Processing*. He is on the editorial board of *Proceedings of the IEEE*. He co-organized the first IEEE International Symposium on Biomedical Imaging and was the founding chair of the technical committee of the IEEE Signal Processing Society (SPS) on Bioimaging and Signal Processing. He is a Fellow of the IEEE and EURASIP and a member of the Swiss Academy of Engineering Sciences. He received several international prizes including three IEEE SPS Best Paper Awards and two Technical Achievement Awards from the IEEE (2008 SPS and IEEE Engineering in Medicine and Biology Society 2010).

REFERENCES

[1] Z. N. Demou, "Time-lapse analysis and microdissection of living 3D melanoma cell cultures for genomics and proteomics," *Biotechnol. Bioeng.*, vol. 101, no. 2, pp. 307–316, Oct. 2008.

[2] K. Kirkegaard, I. E. Agerholm, and H. J. Ingerslev, "Time-lapse monitoring as a tool for clinical embryo assessment," *Hum. Reprod.*, vol. 27, no. 5, pp. 1277–1285, May 2012.

[3] C. Zimmer, B. Zhang, A. Dufour, A. Thébaud, S. Berlemont, V. Meas-Yedid, and J.-C. Olivo-Marin, "On the digital trail of mobile cells," *IEEE Signal Processing Mag.*, vol. 23, no. 3, pp. 54–62, Jan. 2006.

[4] N. Dénervaud, J. Becker, R. Delgado-Gonzalo, P. Damay, A. S. Rajkumar, M. Unser, D. Shore, F. Naef, and S. J. Maerki, "A chemostat array enables the spatio-temporal analysis of the yeast proteome," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 110, no. 39, pp. 15842–15847, Sept. 24, 2013.

[5] A. Miyawaki, "Visualization of the spatial and temporal dynamics of intracellular signaling," *Dev. Cell*, vol. 4, no. 3, pp. 295–305, Mar. 2003.

[6] D. Muzzey and A. Van Oudenaarden, "Quantitative time-lapse fluorescence microscopy in single cells," *Annu. Rev. Cell Dev. Biol.*, vol. 25, pp. 301–327, Nov. 2009.

[7] M. Oheim, "Advances and challenges in high-throughput microscopy for live-cell subcellular imaging," *Expert Opin. Drug Disc.*, vol. 6, no. 12, pp. 1299–1315, Dec. 2011.

[8] G. Myers, "Why bioimage informatics matters," *Nat. Meth.*, vol. 6, no. 7, pp. 659–660, July 2012.

[9] A. Cardona and P. Tomancak, "Current challenges in open-source bioimage informatics," *Nat. Meth.*, vol. 9, no. 7, pp. 661–665, July 2012.

[10] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: An open-source platform for biological-image analysis," *Nat. Meth.*, vol. 9, no. 7, pp. 676–682, July 2012.

[11] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, and J.-C. Olivo-Marin, "Icy: An open bioimage informatics platform for extended reproducible research," *Nat. Meth.*, vol. 9, no. 7, pp. 690–696, July 2012.

[12] A. P. Britto and G. Ravindran, "Review of deformable curves—A retro analysis," *Inf. Technol. J.*, vol. 6, no. 1, pp. 26–36, Jan. 2007.

[13] G. Hamarneh and X. Li, "Watershed segmentation using prior shape and appearance knowledge," *Image Vision Comput.*, vol. 27, no. 1, pp. 59–68, Jan. 2009.

[14] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape," *Int. J. Comput. Vision*, vol. 72, no. 2, pp. 195–215, Apr. 2007.

[15] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vision*, vol. 1, no. 4, pp. 321–331, Jan. 1987.

[16] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 2, pp. 158–175, Feb. 1995.

[17] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comput. Vision*, vol. 22, no. 1, pp. 61–79, Feb. 1997.

[18] H. Zhang, Z. Bian, Y. Guo, B. Fei, and M. Ye, "An efficient multiscale approach to level set evolution," in *Proc. 25th Annu. Int. Conf. IEEE Engineering in Medicine and Biological Society*, Sept. 17–21, 2003, pp. 694–697.

[19] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, "Image segmentation using active contours: Calculus of variations or shape gradients?" *SIAM J. Appl. Math.*, vol. 63, no. 6, pp. 2128–2154, Aug./Sept. 2003.

[20] B. Appleton and H. Talbot, "Globally optimal geodesic active contours," *J. Math. Imaging Vis.*, vol. 23, no. 1, pp. 67–86, July 2005.

[21] X. Bresson, P. Vandergheynst, and J.-P. Thiran, "Multiscale active contours," *Int. J. Comput. Vision*, vol. 70, no. 3, pp. 197–211, Dec. 2006.

[22] K. Zhang, L. Zhang, H. Song, and W. Zhou, "Active contours with selective local or global segmentation: A new formulation and level set method," *Image Vision Comput.*, vol. 28, no. 4, pp. 668–676, Apr. 2010.

[23] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, vol. 79, no. 1, pp. 12–49, Nov. 1988.

[24] R. Delgado-Gonzalo, P. Thévenaz, and M. Unser, "Exponential splines and minimal-support bases for curve representation," *Comput. Aided Geom. Des.*, vol. 29, no. 2, pp. 109–128, Feb. 2012.

[25] M. A. T. Figueiredo, J. M. N. Leitão, and A. K. Jain, "Unsupervised contour representation and estimation using B-splines and a minimum description length criterion," *IEEE Trans. Image Processing*, vol. 9, no. 6, pp. 1075–1087, June 2000.

[26] P. Brigger, J. Hoeg, and M. Unser, "B-spline snakes: A flexible tool for parametric contour detection," *IEEE Trans. Image Processing*, vol. 9, no. 9, pp. 1484–1496, Sept. 2000.

[27] V. V. Kindratenko, "On using functions to describe the shape," *J. Math. Imaging Vis.*, vol. 18, no. 3, pp. 225–245, May 2003.

[28] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 11, pp. 1061–1075, Nov. 1992.

[29] P. Thévenaz and M. Unser, "Snakuscules," *IEEE Trans. Image Processing*, vol. 17, no. 4, pp. 585–593, Apr. 2008.

[30] P. Thévenaz, R. Delgado-Gonzalo, and M. Unser, "The ovuscule," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 382–393, Feb. 2011.

[31] T. McInerney and D. Terzopoulos, "T-snakes: Topology adaptive snakes," *Med. Image Anal.*, vol. 4, no. 2, pp. 73–91, June 2000.

[32] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Comput.*, vol. C-21, no. 3, pp. 269–281, Mar. 1972.

[33] M. Unser, "Sampling—50 Years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.

[34] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[35] M. Jacob, T. Blu, and M. Unser, "Efficient energies and algorithms for parametric snakes," *IEEE Trans. Image Processing*, vol. 13, no. 9, pp. 1231–1244, Sept. 2004.

[36] R. Delgado-Gonzalo and M. Unser, "Spline-based framework for interactive segmentation in biomedical imaging," *IRBM*, vol. 34, no. 3, pp. 235–243, June 2013.

[37] R. Delgado-Gonzalo, P. Thévenaz, C. S. Seelamantula, and M. Unser, "Snakes with an ellipse-reproducing property," *IEEE Trans. Image Processing*, vol. 21, no. 3, pp. 1258–1271, Mar. 2012.

[38] V. Uhlmann, R. Delgado-Gonzalo, and M. Unser, "Snakes with tangent-based control and energies for bioimage analysis," in *Proc. 11th IEEE Int. Symp. Biomedical Imaging*, Apr. 29–May 2, 2014, pp. 806–809.

[39] D. Schmitter, P. Wachowicz, D. Sage, A. Chasapi, I. Xenarios, V. Simanis, and M. Unser, "A 2D/3D image analysis system to track fluorescently labeled structures in rod-shaped cells: Application to measure spindle pole asymmetry during mitosis," *Cell Div.*, vol. 8, no. 6, pp. 1–13, Apr. 27, 2013.

[40] R. Delgado-Gonzalo, "Segmentation and tracking in high-throughput bioimaging," Ph.D. dissertation, Biomedical Imaging Group, EPFL, Switzerland, thesis no. 5657, 2013.

[41] L. D. Cohen, "On active contour models and balloons," *CVGIP: Imag. Understand.*, vol. 53, no. 2, pp. 211–218, Mar. 1991.

[42] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Processing*, vol. 7, no. 3, pp. 359–369, Mar. 1998.

[43] J. Tang and S. T. Acton, "Vessel boundary tracking for intravital microscopy via multiscale gradient vector flow snakes," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 2, pp. 316–324, Feb. 2004.

[44] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[45] A. K. Jain, Y. Zhong, and M.-P. Dubuisson-Jolly, "Deformable template models: A review," *Signal Process.*, vol. 71, no. 2, pp. 109–129, 1998.

[46] C. M. Hoffmann, *Geometric and Solid Modeling*, 1st ed. San Francisco, CA: Morgan Kaufmann, 1989.

[47] C. Zimmer and J.-C. Olivo-Marin, "Coupled parametric active contours," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1838–1842, Nov. 2005.

[48] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nat. Meth.*, vol. 9, no. 7, pp. 671–675, July 2012.

[49] L. Ibáñez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*, 1st ed. Clifton Park, NY: Kitware, Inc., 2003.

[50] Y.-L. Fok, J. C. K. Chan, and R. T. Chin, "Automated analysis of nerve-cell images using active contour models," *IEEE Trans. Med. Imag.*, vol. 15, no. 3, pp. 353–368, June 1996.

[51] Y. Wang, Y. Sun, C. K. Lin, and M. S. Ju, "Segmentation of nerve fibers using multi-level gradient watershed and fuzzy systems," *Artif. Intell. Med.*, vol. 54, no. 3, pp. 189–200, Mar. 2012.

[52] E. U. Mumcuoglu, R. Hassanpour, S. F. Tasel, G. Perkings, M. E. Martone, and M. N. Gurcan, "Computerized detection and segmentation of mitochondria on electron microscope images," *J. Microsc.*, vol. 246, no. 3, pp. 248–265, June 2012.

[53] H. Li, T. Shen, D. Vavylonis, and X. Huang, "Actin filament segmentation using spatiotemporal active-surface and active-contour models," in *Proc. 13th Int. Conf. Med. Image Computing Computer Assisted Intervention*, Sept. 20–24, 2010, vol. 6361, pp. 86–94.

[54] T. Wu, J. Lu, J., Y. Lu, T. Liu, and J. Yang, "Embryo zebrafish segmentation using an improved hybrid method," *J. Microsc.*, vol. 250, no. 1, pp. 68–75, Feb. 2013.

[55] N. Ray, S. T. Acton, and K. Ley, "Tracking leukocytes in vivo with shape and size constrained active contours," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1222–1235, Oct. 2002.

[56] X. Wang, H. Weijun, D. Metaxas, R. Mathew, and E. White, "Cell segmentation and tracking using texture-adaptive snakes," in *Proc. 4th IEEE Int. Symp. Biomedical Imaging*, Apr. 12–16, 2007, pp. 101–104.

[57] B. Zhang, C. Zimmer, and J.-C. Olivo-Marin, "Tracking fluorescent cells with coupled geometric active contours," in *Proc. 2004 IEEE Int. Symp. Biomedical Imaging*, Apr. 15–18, 2004, pp. 476–479.

[58] T. Schaffter, "From genes to organisms: Bioinformatics system models and software," Ph.D. dissertation, Laboratory of Intelligent Systems, EPFL, Switzerland, thesis no. 6081, 2014.

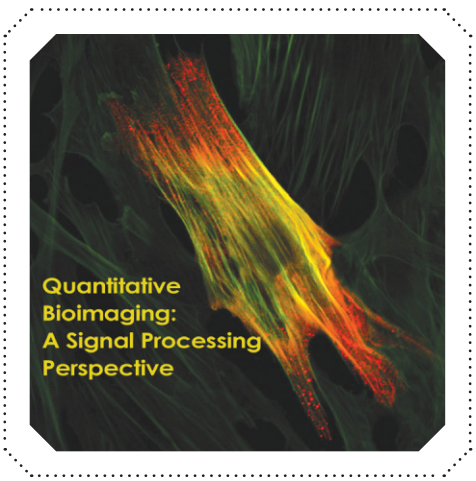
[59] S. Ali and A. Madabhushi, "An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery," *IEEE Trans. Med. Imaging*, vol. 31, no. 7, pp. 1448–1460, July 2012.

[SP]

[Bernd Rieger, Robert P.J. Nieuwenhuizen, and Sjoerd Stallinga]

Image Processing and Analysis for Single-Molecule Localization Microscopy

[Computation for nanoscale imaging]



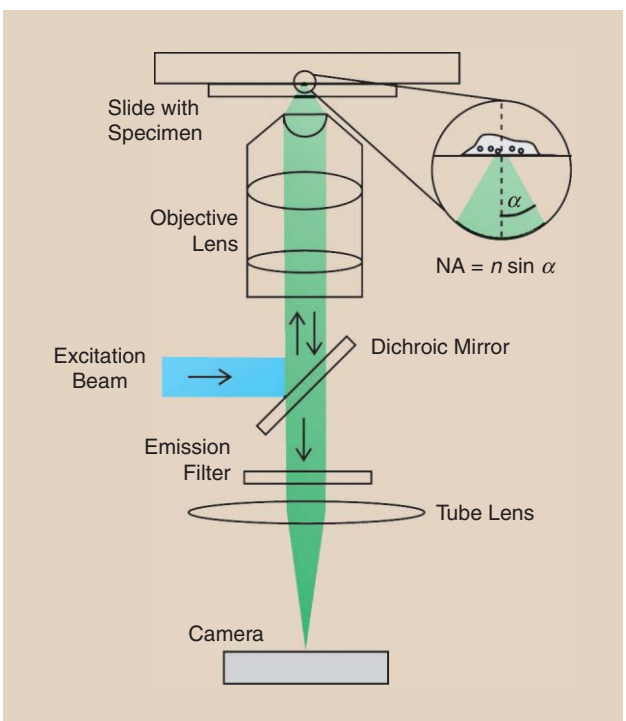
Fluorescence microscopy is currently the most important tool for visualizing biological structures at the subcellular scale. The combination of fluorescence, which enables a high imaging contrast, and the possibility to apply molecular labeling, which allows for a high imaging specificity, makes it a powerful imaging modality. The use of fluorescence microscopy has risen tremendously, in particular since the introduction of the green fluorescent protein (GFP) in the mid-1990s and the possibility to genetically engineer cells to express these proteins. Figure 1 shows the basic layout of a fluorescence microscope. Excitation light of a certain wavelength is reflected via a dichroic beamsplitter and projected onto the specimen via the objective lens of the microscope. The light is absorbed by the fluorescent labels and re-emitted, slightly Stokes-shifted by ~10–100 nm, at a larger wavelength, typically a few nanoseconds later. The emission light is captured by the objective lens and directed toward the camera via the dichroic beamsplitter.

The resolution of a state-of-the-art microscope is limited by diffraction to a length scale $\lambda/2\text{NA}$, where λ is the emission wavelength, and $\text{NA} = n \sin(\alpha)$ is the so-called numerical aperture (NA) of the microscope, where n is the refractive index of the immersion medium n and α is the marginal ray angle of the collected beam (see Figure 1). For visible light and high-NA immersion objectives, this gives resolutions ~200 nm. While this is sufficient for imaging many subcellular structures, it is insufficient for providing an image of the molecular machinery that underlies the functioning of the cell. Electron microscopy, however, can reveal image detail on the order of nanometers but does not allow live-cell imaging nor efficient specific labeling.

Over the last decade, a number of optical nanoscopy techniques have been proposed to bridge the resolution gap between electron and conventional light microscopy. Localization microscopy is one of these superresolution techniques [1]–[4]. These techniques rely on the localization of single fluorescent molecules, which was

Digital Object Identifier 10.1109/MSP.2014.2354094
Date of publication: 5 December 2014

already commonly done, e.g., in the field of single particle tracking before the advent of localization microscopy [5]. In localization microscopy, the fluorescent labels are photochemically manipulated to switch on and off stochastically, such that at each instant in time only a sparse subset of all molecules is in the on-state in which they can fluoresce. By now there is a whole plethora of stochastic switching mechanisms and suitable fluorescent labels [4]. The required ratio of on/off times to see only single emitters in a region of size λ/NA depends on the labeling density, camera frame time, etc., but is typically fewer than 1/100. Recording many frames of blinking emitting molecules thus provides a sequence of images of different random subsets of all molecules. The active molecules appear as well-separated spots that can be identified and processed to provide the position of the molecules. The localization precision is on the order of $\lambda/\mathrm{NA}/\sqrt{N_{\mathrm{ph}}} \approx 10$ nm with N_{ph} the number of detected photons (typically a few hundred to a few thousand). Assembling the localization data obtained from all frames into one visualization of the final superresolution image reveals details on the length scale of 10–100 nm; this is about one order below the diffraction limit of conventional light microscopy.



[FIG1] The schematics of an epi-fluorescence light microscope. The excitation light is focused onto the sample and the emission light is captured by the same lens and recorded on a camera. The dichroic mirror is chosen such that it reflects the excitation light but transmits the fluorescent emission light, which is of a slightly larger wavelength. The objective is characterized by the NA, which combines the refractive index of the immersion medium n and the maximum angle α at which light is captured.

THE NECESSARY TECHNOLOGY FOR LOCALIZATION MICROSCOPY IS NOT PROHIBITIVE: A STATE-OF-THE-ART SETUP ONLY REQUIRES A FLUORESCENCE MICROSCOPE, POWERFUL LIGHT SOURCES, AND A CAMERA WITH HIGH QUANTUM EFFICIENCY AND LOW READOUT NOISE.

The necessary technology for localization microscopy is not prohibitive: a state-of-the-art setup only requires a fluorescence microscope, powerful light sources, and a camera with high quantum efficiency and low readout noise. Next to this hardware, software for image processing and analysis is essential for extracting the desired molecular locations in a robust, optimal, and fast way. In this review article, we detail the image processing and workflow from raw camera frames to the visualization and quantitative analysis of the superresolution image. Figure 2 shows an overview of this workflow.

THE IMAGE PROCESSING PIPELINE

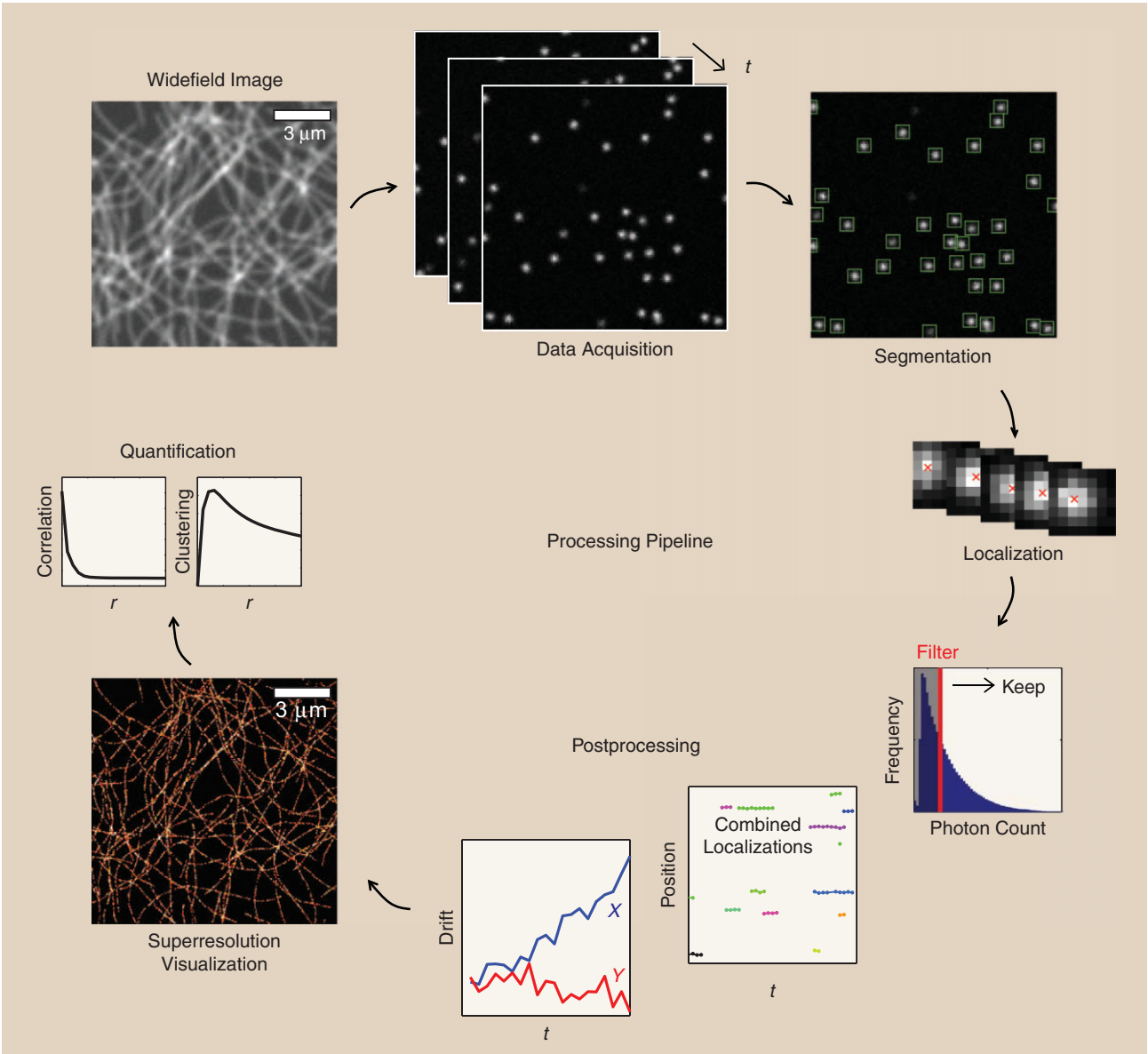
SEGMENTATION

The first step in processing the raw frames consists of identifying and segmenting regions of interest (ROIs) that contain the emissions of single fluorescent emitters. Usually this is done by thresholding the raw frames based on the pixel intensity relative to the (local) background noise level [1], [2]. Pixels in which the value is larger than a fixed threshold value or larger than a multiple of the background intensity b are taken as the center of ROIs that are used for localization of possible fluorophore positions in the next processing step.

Besides this basic thresholding approach, more advanced segmentation algorithms have also been proposed. In one proposed method, the raw images are first decomposed into wavelet maps to separate the fluorescence signal from bloblike sources from the background intensity and noise [6], [7]. Subsequently, ROIs are identified using a watershed segmentation algorithm.

Another approach to identifying ROIs makes explicit use of local hypothesis testing against the null hypothesis that a pixel belongs to the local background. This is achieved by computing the P-value for each pixel under the assumption that it is drawn from a normal distribution with the local mean and standard deviation of pixel values as parameters [8]. A related method that was proposed for single particle tracking employs a likelihood ratio test in each pixel. In this test, the ratio is computed between the likelihoods of the null hypothesis and that of the hypothesis of having a single emission from a fluorophore in the center of the pixel, assuming that the noise per pixel is Gaussian [9]. Under the null hypothesis, this ratio follows a chi-squared distribution. Pixels are thus thresholded based on the P-value of the chi-squared distribution for the likelihood ratio value of that pixel.

Segmentation algorithms typically assume a locally uniform background intensity. This is reasonable if the ROI is only a few pixels wide, unless there is a high degree of autofluorescence and the fluorophores themselves are relatively dim. For such cases, temporal median filtering has been proposed as a method for estimating the local background intensity [10].



[FIG2] The complete pipeline for generating a 2-D superresolution image based on raw frames of sparsely activated fluorophores. The consecutive steps in this pipeline are acquisition of raw data, segmentation of ROIs, localization of potential fluorophores in the ROIs, postprocessing of the localizations (e.g., filtering, frame connection, drift correction), and visualization of the localizations.

SINGLE-MOLECULE LOCALIZATION

Once ROIs in the raw data have been segmented, the next step is to estimate the positions of the emitting fluorescent molecules in these regions. The most common approaches for this are the center of mass (CM) algorithm and algorithms that fit a point spread function (PSF) model to the data with a (weighted) least-squares (LS) estimator or a maximum likelihood estimator (MLE).

The CM algorithm computes the center of the intensity distribution. In the absence of any background intensity, this estimate corresponds well to the emitter's true location. However, for nonnegligible background intensities, this leads to a bias toward the CM of the background intensity, which is usually in the center of the ROI. Therefore the local background intensity needs to be estimated and subtracted before the CM can be computed.

LS and MLE algorithms attempt to fit a PSF model to the pixel intensities in a ROI. Typically the PSF model consists of a circularly symmetric Gaussian function for two-dimensional (2-D) localization microscopy

$$\mathrm{PSF}(x,y)=\frac{1}{2\pi\sigma_g^2}e^{-\frac{(x-x_c)^2+(y-y_c)^2}{2\sigma_g^2}}. \tag{1}$$

Here the parameters x_c and y_c denote the position of the emitter in the x- and y-direction and σ_g specifies the width of the PSF. The Gaussian PSF model is not derived from optical theory but is instead chosen for its conceptual simplicity and computational

efficiency. However, for typical imaging conditions, the Gaussian PSF approximates the theoretical PSF sufficiently well for accurate and precise localization [11], [12].

From the PSF model follows the expected intensity μ_k per pixel k that is fitted to the data

$$\mu_k = I_0 \int_{A_k} \mathrm{PSF}(u, v)\, du dv + b, \tag{2}$$

where I_0 denotes the sum intensity of the fluorophore, b the expected background photon count, and the integration runs over the area A_k of the kth pixel. The parameters that are to be estimated are thus x_c, y_c, I_0, b, and possibly σ_g.

In addition to an optical model for μ_k, fitting the PSF model to the data also requires a noise model for the imaging system. LS algorithms implicitly assume a Gaussian noise model, whereas the slower but more precise MLE algorithms assume a Poissonian noise model. The latter algorithms can be implemented on a graphical processing unit (GPU) to estimate the positions of many emitters in parallel and so achieve real-time computation [13].

An important issue in localization microscopy is the precision with which single fluorophores can be localized [14]–[16]. This is often analyzed using the concept of the Cramér–Rao lower bound (CRLB), which expresses the lowest variance of any unbiased estimator of a fluorophore's position for a given noise model [17]. For a Poissonian noise model, a good analytical approximation for this bound is given by [18]

$$\Delta x_{\mathrm{loc}}{}^2 = \frac{\sigma_e^2}{N}\left(1 + 4\tau + \sqrt{\frac{2\tau}{1+4\tau}}\right). \tag{3}$$

Here N is the number of signal photons, $\sigma_e^2 = \sigma_g^2 + a^2/12$ with a^2 is the pixel area, and τ is a normalized dimensionless background parameter $\tau = 2\pi\sigma_g^2 b/(Na^2)$ with b the number of background photons per pixel.

The noise in the commonly used scientific CMOS (sCMOS) and electron multiplying charge-coupled device (EMCCD) cameras deviates from the Poisson noise model in two important ways. sCMOS cameras suffer from a small amount of (pixel-dependent) Gaussian readout noise, which effectively acts as if b is increased with the variance of the readout noise [19]. EMCCD cameras suffer much less from readout noise due to the electron multiplication process. However, the stochasticity of this process also introduces so-called excess noise, which typically deteriorates the localization variance $\Delta x_{\mathrm{loc}}{}^2$ by a factor of two [20]. Balancing the effects of readout noise and excess noise implies that sCMOS cameras are preferred over EMCCD cameras, except in extremely low light conditions that are not typically encountered in localization microscopy [19]. Other considerations in choosing between cameras are that EMCCD cameras have a better photosensitivity, and that sCMOS cameras typically have a smaller physical pixel size and faster frame rate. Finally, sCMOS cameras require a calibration of the gain and readout noise of each pixel for accurate localization because they often vary substantially among different pixels on the same camera. For a more extensive introduction into the choice of localization algorithm, we refer to the review in [21].

POSTPROCESSING

After all the segmented ROIs have been processed by the localization algorithm, postprocessing of the raw localizations is needed. In the first postprocessing step, raw localizations are usually filtered. The goal of this filtering is to remove localizations that do not represent accurate position estimates of single fluorescent molecules, e.g., because they are due to overlapping emissions of multiple fluorophores or due to autofluorescence or residual sample contaminations. The filtering is usually done based on information that is returned by the localization algorithm, such as the estimated intensity of the fluorophore, the localization precision, the width of the PSF, as well as on the goodness of fit of the model to the data [1], [22]. The latter can be expressed as the (weighted) sum of squared errors between the fitted model and the data or as a ratio between the likelihoods of a fluorophore being present or absent.

In the second postprocessing step, localizations originating from the same fluorophore in consecutive frames of the raw image sequence are combined. This is attempted by searching for localizations in subsequent frames that are also spatially proximate, typically within a few times the estimated localization precision. The rationale for this operation is that fluorophores are often visible in multiple consecutive frames before transitioning into a stable dark state or photobleached state, whereas it is unlikely that a nearby fluorophore starts emitting during this time. In practice, fluorophores will not always be localized in all frames before going into a stable dark state, either due to failures of the localization algorithm or due to short blinking events during which the fluorophores briefly stop emitting light. Therefore, spatially proximate localizations are usually still combined if they are only a few frames apart in time [23].

A third common postprocessing operation is to correct for drift during the acquisition. Since localization microscopy experiments can last anywhere from a few minutes up to several hours, the sample often moves relative to the detector over distances larger than the localization precision of about 10 nm. This movement can be reduced with hardware solutions, e.g., by mechanically fixing the objective lens to the stage or by using a control system that actively controls the position of the sample in the image plane [24], [25]. Axial drift, causing the sample to drift out of focus, must be suppressed or controlled just as well as the lateral drift in the image plane.

One option is to add fiducial markers such as fluorescent microbeads to the sample that are visible during the entire acquisition [1]. These fiducial beads can then be localized and used to determine the position of the sample at each moment in time. Another option for drift correction is to estimate the shifts between images of the sample at different time points. This can be achieved by determining the maximum of the cross-correlation [24], [26], [27] between these images, which can either be raw camera images or superresolution images that visualize the localizations from these frames. The latter, however, is preferred for precision due to the larger high-frequency content of the superresolution image. The shift estimation should not be done between subsequent images only, as this leads to compounding of registration errors, but between image pairs further

apart in time. The main benefit of this approach is that it does not require any changes on the experimental side.

VISUALIZATION

The final step in the processing pipeline from raw data to superresolution image is the actual visualization of the data. For standard fluorescence microscopy acquisitions, this sampling occurs in the camera where the pixel positions along with the magnification determine the sampling of the image. In addition, the values per pixel are determined by the number of recorded photons per pixel bin that are translated into analog-to-digital units (ADUs) with a linear amplification factor. Unlike these standard fluorescence microscopy techniques, localization microscopy does not sample an image at pixel locations but produces a list of coordinates that represent the estimated fluorophore locations.

Several methods have been proposed for visualizing localizations in pixelated images that can be shown on a display device [28]. First, a scatterplot can be made of the localizations where the coordinates are plotted as a symbol in a Cartesian coordinate system [2]. Second, a histogram image can be made where the field of view is divided into square pixel bins and the number of localizations that fall in each bin is counted and used to assign intensity values to pixels. The resulting images often have a low signal-to-noise-ratio (SNR) per pixel and may cause aliasing problems if the resolution of the display device is too low. Therefore, these images are often blurred with a Gaussian kernel with a size on the order of the average estimated localization precision.

Third, localizations may be rendered as Gaussian blobs with a variable width proportional to their estimated localization precisions [1]. Alternatively, a fast method for obtaining a similar image is to sum several histogram images where the localizations are jittered for each image with a zero-mean normal probability distribution with a standard deviation proportional to the localization precision per localization [22]. These visualization methods have the benefit of conveying information about the precision of each localization in the images. However, additional factors such as uncorrected stage drift also lead to additional blurring. Therefore the blobs in these images cannot always be taken to represent the likelihood functions for the positions of the localized fluorophores.

Figure 3 shows an example where these visualization methods have been applied. In general, Gaussian rendering is the preferred method of visualization: it is best at conveying the information present in the data and it does not suffer much from aliasing with low-resolution displays.

ONE IMPORTANT EXTENSION OF 2-D LOCALIZATION MICROSCOPY IMAGING IS THE LOCALIZATION OF FLUOROPHORES IN THREE DIMENSIONS. THIS REQUIRES INFORMATION ABOUT THE AXIAL POSITION OF THE FLUOROPHORE IS PRESENT AND CAN BE EXTRACTED FROM THE RECORDINGS.

EXTENSIONS

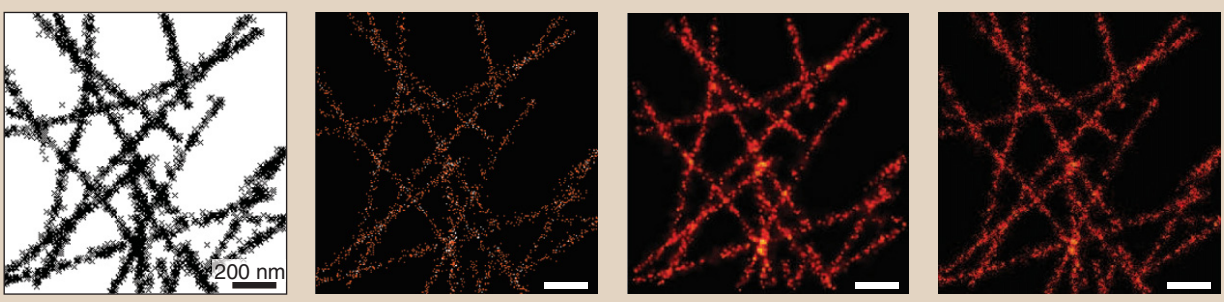
Until now, the discussion focused in detail on the complete pipeline for generating a 2-D superresolution image based on raw frames of sparsely activated fluorophores. Here we will address several extensions of this pipeline involving localization in three dimensions, multicolor localization, and imaging with overlapping spots.

THREE-DIMENSIONAL LOCALIZATION

One important extension of 2-D localization microscopy imaging is the localization of fluorophores in three dimensions. This requires information about the axial position of the fluorophore is present and can be extracted from the recordings.

A first approach to this problem is to modify the optical setup such that the shape and/or size of the PSF can be uniquely related to the axial position of the fluorophore. The most common method to achieve this is to introduce astigmatism into the optical system [24]. This causes the minimum width of the PSF in the x- and y-direction to occur at different axial positions. The position can then be determined based on the ellipticity of the PSF.

A second approach to obtain the axial position is to modify the setup such that multiple images of the fluorophores with different



[FIG3] An illustration of commonly used visualization methods. (a) The scatterplot method, (b) histogram binning method, (c) Gaussian rendering method, and (d) jittered binning method. The scale bar is 200 nm.

defocus are simultaneously acquired. This is usually accomplished with a beam splitter that splits the emission light into two channels with different optical path lengths to the camera, such that the two images of the fluorophores are defocused with respect to each other [29].

For both these approaches to three-dimensional (3-D) localization, the PSF model that is used in the basic 2-D localization algorithm needs to be modified. The modified PSF model must provide a specification of the appearance of the fluorophore for the full range of axial positions under consideration and for all image channels on which it is observed. The PSF shapes for 3-D localization techniques may be difficult to describe in an analytic formula such as the Gaussian PSF model. An example of this is the double-helix PSF, where a spot doublet rotates with the axial focus position [30]. In such cases, the PSF can also be determined numerically or empirically. The latter approach then requires subsequent interpolation between the measured axial positions to provide a full specification of a fluorophore's appearance.

MULTICOLOR LOCALIZATION

Another important extension of the basic pipeline is the imaging of different labeled molecules in an experiment. A common method for doing this is to label these molecules with fluorophores with different emission spectra [31]. Wavelength dependent beam splitters are then inserted in the emission light path such that the light at different wavelengths ends up at different parts of the camera or at different cameras. The observed fluorophores can subsequently be classified into the different used species based on the fraction of the photons of each fluorophore ending up in the different color channels. Usually, though, the beam splitters are optimally selected such that each color channel only shows a single fluorescent species.

An important problem that arises when imaging fluorophores in different color channels is the registration of the various channels with respect to each other. This needs to be done with an accuracy comparable to the localization precision, which is typically 10% or less of the camera pixel size. A common solution employs fiducial markers that are visible in all color channels. These markers are first imaged and localized, and subsequently, a nonaffine mapping function is computed that maps the positions of the markers in one color channel to their positions in the other channels [32].

An alternative approach to multicolor imaging is to use photoswitchable dye pairs with different activator dyes but identical reporter dyes [31]. In this way, the wavelength of the illumination can be used to determine which dye pairs are activated and therefore which labeled molecules are imaged. The emitted light of all reporter dyes can then be imaged in a single image on the camera, thus circumventing chromatic aberration problems and obviating the need for a registration procedure between different images.

HIGH-DENSITY METHODS

A common problem when localizing fluorophores is that segmented ROIs contain overlapping spots of multiple active fluorophores. This issue is particularly important when the density of active fluorophores is high. Several solutions have been proposed that attempt to fit a PSF model to each of the spots in the ROI, either by fitting spots one by one [33] or by finding the model with the number of PSFs that best matches the data [34].

Several other methods for dealing with overlapping spots have been proposed that do not estimate fluorophore positions but rather estimate the density of fluorophores instead. One such approach is to deconvolve the entire raw data set [35]. This means that for each frame, a fluorophore density is estimated, which has the highest likelihood of producing the experimentally recorded data after convolution with the PSF. To achieve sub-diffraction resolution, this density is sampled with a smaller pixel size than the experimental data. The estimation also incorporates a prior probability for the density per frame that promotes sparsity: because relatively few emitters are active in each frame, the solution should also have few pixels with nonzero density. A related approach to estimating the density is provided by compressive sensing [36], [37]. Unlike the deconvolution approach, an estimate $\rho(x, y)$ is made for each frame independently, which minimizes the balanced sum between a data misfit term and sparsity promoting "L1-norm" of the form $\sum_{x,y} |\rho(x, y)|$. A subtlety in these approaches is that, in principle, the final estimated density is a relative rather than an absolute estimation of the molecular density, as fluorophores can reappear in the on-state multiple times during the data acquisition.

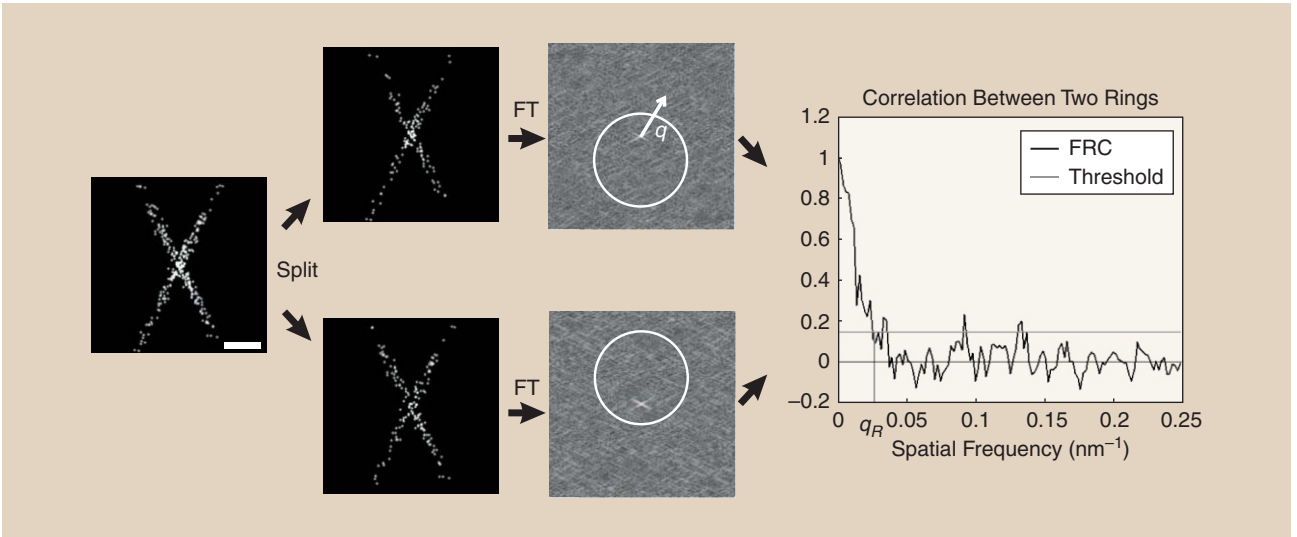
The final approach to be mentioned here is called the *Bayesian analysis* of the blinking and bleaching (3B) method [38]. In this method, the on- and off-switching and bleaching behavior of each fluorophore is modeled as a Markov process. Using this model, many different estimates are made of the number of fluorophores, their positions, and their activity in each frame. These estimates are then all used to create a probability map of the positions of the fluorophores. A major drawback of this method is its high computational cost.

QUANTITATIVE IMAGE ANALYSIS

RESOLUTION QUANTIFICATION

Localization microscopy produces images that have a resolution below the diffraction limit. However, so far we have not answered this question: Exactly how far below the diffraction limit is the resolution in these images?

The apparent width of structures is clearly blurred on the scale of the localization precision. Additional blurring is introduced by sample drift and the nonzero size of the fluorescently labeled markers (e.g., antibodies). Another factor limiting resolution is that the molecules of interest are labeled with fluorophores at a finite density. In addition to this, usually not all of the target molecules are labeled and not all labels result in actual localizations. So, both the overall density of recorded labels and the different blurring factors influence what detail can be reliably discerned. It is noteworthy that the sampling theorem does not apply here: the localizations that are



[FIG4] A schematic illustration of FRC resolution computation. The localizations are divided into two halves, and their Fourier transforms are correlated over the perimeters of circles in Fourier space of radius q. The resulting FRC curve decays with spatial frequency, and the image resolution is taken to be the inverse of the spatial frequency q_R where the FRC curve drops below the threshold 1/7.

obtained are positions rather than samples of a bandwidth limited density function.

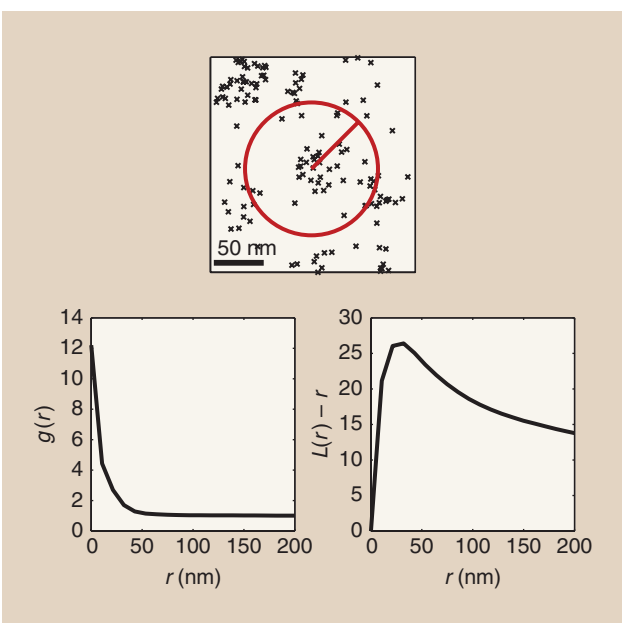
Fourier ring correlation (FRC) provides a practical image-resolution measure that incorporates all of the aforementioned factors that influence the resolution [39]. To compute the FRC resolution, the full set of estimated fluorophore positions is divided into two independent subsets. An image is then made for each subset, yielding two subimages $f_1(x, y)$ and $f_2(x, y)$. Statistical correlation of their Fourier transforms $\hat{f}_1(\vec{q})$ and $\hat{f}_2(\vec{q})$ over the pixels on the perimeter of circles of constant spatial frequency magnitude $q = |\vec{q}|$ then gives the FRC

$$\mathrm{FRC}(q) = \frac{\sum\limits_{\vec{q}\in\text{circle}} \hat{f}_1(\vec{q})\hat{f}_2(\vec{q})^*}{\sqrt{\sum\limits_{\vec{q}\in\text{circle}} \hat{f}_1(\vec{q})^2}\sqrt{\sum\limits_{\vec{q}\in\text{circle}} \hat{f}_2(\vec{q})^2}}. \tag{4}$$

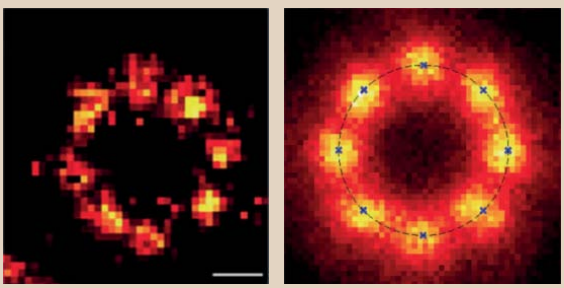
For low spatial frequencies, the FRC curve is close to unity, and for high spatial frequencies, noise dominates the data and the FRC decays to zero. The image resolution is defined as the inverse of the spatial frequency $R = 1/q_R$ for which the FRC curve drops below a given threshold. Various threshold criteria are in use in the field of cryo-EM, but the fixed threshold of $\mathrm{FRC}(q_R) = 1/7 \approx 0.143$ was empirically found to be the most appropriate for localization microscopy images. Thus the FRC resolution describes the length scale below which the image lacks signal content; smaller details are not resolved in the image. The steps needed to compute the FRC resolution are illustrated in Figure 4.

CLUSTER ANALYSIS

Another common quantification task for localization microscopy is cluster analysis. In this analysis, the objective is to assess the degree to which the imaged molecules cluster together. In other words, the question is: To what extent does the distance distribution between the molecules differ from that of randomly distributed molecules? The two most common analysis techniques used for this problem are the pair correlation function and techniques based on Ripley's K function. See Figure 5 for an illustrative example.



[FIG5] A schematic illustration of cluster analysis. The cross-correlation function $g(r)$ indicates the average density of localizations around each localization at a distance r. Ripley's K analysis considers the number of localizations within a distance r of each localization. The linearized version $L(r)$ is equal to r for randomly distributed localizations and thus allows for easy visualization of clustering of the localizations by plotting $L(r) - r$.



[FIG6] (a) A single superresolution image of a completely labeled nuclear pore complex (NPC) and (b) the sum of 426 of these superresolved and registered NPC images. The scale bar is 50 nm. (Image used with permission from [42].)

The pair-correlation function $g(r)$ describes the average number of localizations at a distance r from each other localization, normalized by the average density of localizations [40]

$$g(x, y) = \langle \rho(x', y')\rho(x' - x, y' - y)\rangle / \rho^2. \quad (5)$$

Here $\rho(x, y)$ is the density of localizations at position (x, y). When using the pair-correlation function to assess clustering, there are two important caveats that have to be taken into account. First, if fluorophores are localized multiple times then this will result in excess clustering on the length scale of the localization uncertainty which has to be accounted for. This will almost always be the case unless fluorescent proteins are used that are permanently photobleached after their first photoactivation. Second, the pair-correlation between localizations does not exactly reflect that of the actual labeled molecules due to the localization uncertainty.

A closely related alternative to pair-correlation analysis is provided by the use of Ripley's K function [41]. Ripley's K function measures the average number of localizations within a distance r from each other localization

$$K(r) = \frac{1}{N\rho} \sum_{i=1}^{N} \sum_{j \neq i} \delta_{i,j}, \quad (6)$$

where N is the number of localizations and $\delta_{i,j} = 1$ if the distance between localizations i and j is smaller than r and zero otherwise. Often, the linearized functions $L(r) = \sqrt{K(r)/\pi}$ or $H(r) = L(r) - r$ are used for data analysis, since the latter has an expected value of zero for a random distribution of localizations.

IDENTICAL STRUCTURE AVERAGING

The resolution of localization microscopy can even be further increased by image processing in case the field-of-view contains many copies of identical structures or macromolecules. All superresolution reconstructions from these structures can then be segmented and subsequently registered with respect to each other and added together. This results in one compound superresolution image of the structure with a very high effective localization density. This, in turn, increases the SNR and thus the resolution as detailed in the section "Resolution Quantification." In addition, this registration can deal with partially unoccupied labeling sites on the structure as long as the empty labeling sites are random and even these incomplete structures contribute to an increased SNR. An example of identical structure averaging is given by superresolution imaging of the NPC, which is the gateway between the cell nucleus and the cytoplasm; compare Figure 6(a) and (b) [42], [43]. The same idea of averaging biological identical structures, but with low SNR of each individual image, has been employed by Engelenburg et al. [44] to gain insight into the protein distribution within HIV.

OUTLOOK

As image processing and analysis is such an essential ingredient of superresolution microscopy, substantial efforts have already been dedicated to it over the years. The standard 2-D processing pipeline is now well established, but there are still significant opportunities for improvement. In particular, the detection of all suitable ROIs that contain single molecule events and the filtering of localization events are suitable topics, as these have been treated largely ad hoc rather than fundamentally. The reason that these tasks have been somewhat neglected is, in part, because a reasonably good superresolution reconstruction can be obtained without such a very fundamental treatment. The methods using deconvolution on the raw data instead circumvent this problem; however, they suffer from other drawbacks (see the section "High-Density Methods"). We see a need for fast and reliable identifications of ROIs especially for very weak signals (fewer than a few hundred photons) on substantial background. Such signals are typically encountered when imaging fluorescent proteins, while imaging at high frame rates for live cell observation or when imaging deeper inside a cell or tissue ($> 1\,\mu$m). We also foresee a rise for methods that deal with overlapping spots of proximate emitters, as higher active emitter densities are beneficial for faster and higher resolution imaging. Finally, quantitative image analysis methods in general and the use of prior knowledge in particular (as in the section "Identical Structure Averaging") may further open up the nanoscopic world for image-based exploration.

AUTHORS

Bernd Rieger (b.rieger@tudelft.nl) received the M.Sc. degree in physics from the Technische Universität München, Germany, in 1999. He received his Ph.D. degree in 2004 in image processing and analysis from Delft University of Technology, The Netherlands. After a postdoctoral position at the Max Planck Institute for Biophysical Chemistry, Göttingen, and a research position at FEI Electron Optics, Eindhoven, in 2006 he rejoined the Department of Imaging Physics at Delft University of Technology, where he is now an associate professor. His current research interests are the application of image analysis to optical nanoscopy and to electron microscopy.

Robert P.J. Nieuwenhuizen (r.p.j.nieuwenhuizen@tudelft.nl) received a B.Sc. degree in applied physics in 2008 (cum laude), an M.Sc. degree in applied physics in 2011 (cum laude), and an M.Sc. degree in management of technology in 2011 (cum laude), all

from Delft University of Technology, The Netherlands. He is currently a Ph.D. candidate in the Department of Imaging Physics at Delft University of Technology, under the supervision of Bernd Rieger and Sjoerd Stallinga, doing research on quantitation of localization microscopy.

Sjoerd Stallinga (s.stallinga@tudelft.nl) obtained his Ph.D. degree in 1995 from the University of Nijmegen after graduating from that university in 1993 on a theoretical study in liquid crystal physics. He worked from 1995 to 2008 as a senior scientist and project leader at Philips Research in Eindhoven on various optical topics, and, in 2009, he joined the Delft University of Technology, The Netherlands, as an associate professor in the Department of Imaging Physics. His current research interests are the analysis, design, and realization of computational optical imaging systems, with an emphasis on optical nanoscopy and digital pathology.

REFERENCES

[1] E. Betzig, G. Patterson, R. Sougrat, O. Lindwasser, S. Olenych, J. Bonifacino, M. Davidson, J. Lippincott-Schwartz, and H. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, pp. 1643–1645, Sept. 2006.

[2] M. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods*, vol. 3, no. 10, pp. 793–795, 2006.

[3] S. T. Hess, T. P. K. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophys. J.*, vol. 91, no. 11, pp. 4258–4272, 2006.

[4] T. Klein, S. Proppert, and M. Sauer, "Eight years of single-molecule localization microscopy," *Histochem. Cell Biol.*, vol. 141, no. 6, pp. 561-575, 2014.

[5] M. J. Saxton and K. Jacobson, "Single-particle tracking: Applications to membrane dynamics," *Annu. Rev. Biophys. Biomolec. Struct.*, vol. 26, pp. 373–399, June 1997.

[6] J.-C. Olivo-Marin, "Extraction of spots in biological images using multiscale products," *Pattern Recogn.*, vol. 35, no. 9, pp. 1989–1996, 2002.

[7] I. Izeddin, J. Boulanger, V. Racine, C. Specht, A. Kechkar, D. Nair, A. Triller, D. Choquet, M. Dahan, and J. Sibarita, "Wavelet analysis for single molecule localization microscopy," *Opt. Express*, vol. 20, no. 3, pp. 2081–2095, 2012.

[8] Y. Li, Y. Ishitsuka, P. Hedde, and G. Nienhaus, "Fast and efficient molecule detection in localization-based super-resolution microscopy by parallel adaptive histogram equalization," *ACS Nano*, vol. 7, no. 6, pp. 5207–5214, 2013.

[9] A. Serge, N. Bertaux, H. Rigneault, and D. Marguet, "Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes," *Nat. Methods*, vol. 5, no. 8, pp. 687–694, 2008.

[10] E. Hoogendoorn, K. Crosby, D. Leyton-Puig, R. Breedijk, K. Jalink, T. Gadella, and M. Postma, "The fidelity of stochastic single-molecule super-resolution reconstructions critically depends upon robust background estimation," *Sci. Rep.*, vol. 4, p. 3854, Jan. 2014.

[11] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Appl. Opt.*, vol. 46, no. 10, pp. 1819–1829, 2007.

[12] S. Stallinga and B. Rieger, "Accuracy of the Gaussian point spread function model in 2D localization microscopy," *Opt. Express*, vol. 18, no. 24, pp. 24,461–24,476, 2010.

[13] C. Smith, N. Joseph, B. Rieger, and K. Lidke, "Fast, single-molecule localization that achieves theoretically minimum uncertainty," *Nat. Methods*, vol. 7, no. 5, pp. 373–375, 2010.

[14] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophys. J.*, vol. 82, no. 5, pp. 2775–2783, 2002.

[15] M. P. Gordon, T. Ha, and P. R. Selvin, "Single-molecule high-resolution imaging with photobleaching," *Proc. Natl. Acad. Sci. USA*, vol. 101, no. 17, pp. 6462–64654, 2004.

[16] H. Deschout, F. C. Zanacchi, M. Mlodzianoski, A. Diaspro, J. Bewersdorf, S. T. Hess, and K. Braeckmans, "Precisely and accurately localizing single emitters in fluorescence microscopy," *Nat. Methods*, vol. 11, no. 3, pp. 253–266, 2014.

[17] R. Ober, S. Ram, and S. Ward, "Localization accuracy in single-molecule microscopy," *Biophys. J.*, vol. 86, no. 2, pp. 1185–1200, 2004.

[18] B. Rieger and S. Stallinga, "The lateral and axial localization uncertainty in super-resolution light microscopy," *ChemPhysChem*, vol. 15, no. 4, pp. 664–670, 2014.

[19] F. Huang, T. Hartwich, F. Rivera-Molina, Y. Lin, C. Whitney, J. Long, P. Uchil, J. Myers, M. Baird, W. Mothes, M. Davidson, D. Toomre, and J. Bewersdorf, "Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms," *Nat. Methods*, vol. 10, no. 7, pp. 653–658, 2013.

[20] J. Chao, S. Ram, E. Ward, and R. Ober, "Ultrahigh accuracy imaging modality for super-localization microscopy," *Nat. Methods*, vol. 10, no. 4, pp. 335–338, 2013.

[21] A. Small and S. Stahlheber, "Fluorophore localization algorithms for super-resolution microscopy," *Nat. Methods*, vol. 11, no. 3, pp. 267–279, 2014.

[22] P. Krizek, I. Raska, and G. Hagen, "Minimizing detection errors in single molecule localization microscopy," *Opt. Express*, vol. 19, no. 4, pp. 3226–3235, 2011.

[23] P. Annibale, S. Vanni, M. Scarselli, U. Rothlisberger, and A. Radenovic, "Quantitative photo activated localization microscopy: Unraveling the effects of photoblinking," *PLoS ONE*, vol. 6, no. 7, p. e22678, 2011.

[24] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science*, vol. 319, no. 5864, pp. 810–813, 2008.

[25] A. Pertsinidis, Y. Zhang, and S. Chu, "Subnanometre single-molecule localization, registration and distance measurements," *Nature*, vol. 466, pp. 647–651, July 2010.

[26] M. Mlodzianoski, J. Schreiner, S. Callahan, K. Smolkova, A. Dlaskova, J. Santorova, P. Jezek, and J. Bewersdorf, "Sample drift correction in 3D fluorescence photoactivation localization microscopy," *Opt. Express*, vol. 19, no. 16, pp. 15,009–15,019, 2011.

[27] C. Geisler, T. Hotz, A. Schönle, S. W. Hell, A. Munk, and A. Egner, "Drift estimation for single marker switching based imaging schemes," *Opt. Express*, vol. 20, no. 7, pp. 7274–7289, 2012.

[28] R. Nieuwenhuizen, S. Stallinga, and B. Rieger, "Visualization and resolution in localization microscopy," in *Cell Membrane Nanodomains: From Biochemistry to Nanoscopy*. New York: Taylor & Francis, 2014, pp. 409–430.

[29] M. Juette, T. Gould, M. Lessard, M. Mlodzianoski, B. Nagpure, B. Bennett, S. Hess, and J. Bewersdorf, "Three-dimensional sub 100 nm resolution fluorescence microscopy of thick samples," *Nat. Methods*, vol. 5, no. 6, pp. 527–529, 2008.

[30] S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. E. Moerner, "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function," *Proc. Natl. Acad. Sci. USA*, vol. 106, no. 9, pp. 2995–2999, 2009.

[31] M. Bates, B. Huang, G. T. Dempsey, and X. Zhuang, "Multicolor super-resolution imaging with photo-switchable fluorescent probes," *Science*, vol. 317, no. 5845, pp. 1749–1753, 2007.

[32] L. S. Churchman, Z. Okten, R. S. Rock, J. F. Dawson, and J. A. Spudich, "Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 5, pp. 1419–1423, 2005.

[33] S. Holden, S. Uphoff, and A. Kapanidis, "Daostorm: An algorithm for high-density super-resolution microscopy," *Nat. Methods*, vol. 8, no. 4, pp. 279–280, 2011.

[34] F. Huang, S. Schwartz, J. Byars, and K. Lidke, "Simultaneous multiple-emitter fitting for single molecule super-resolution imaging," *Biomed. Opt. Express*, vol. 2, no. 5, pp. 1377–1393, 2011.

[35] E. Mukamel, H. Babcock, and X. Zhuang, "Statistical deconvolution for superresolution fluorescence microscopy," *Biophys. J.*, vol. 102, no. 10, pp. 2391–2400, 2012.

[36] L. Zhu, W. Zhang, D. Elnatan, and B. Huang, "Faster STORM using compressed sensing," *Nat. Methods*, vol. 9, no. 7, pp. 721–726, 2012.

[37] E. J. Candes and C. Fernandez-Grandam, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.* vol. 67, no. 6, pp. 906–956, 2014.

[38] S. Cox, E. Rosten, J. Monypenny, T. Jovanoivc-Talisman, D. Burnette, J. Lippincott-Schwartz, G. Jones, and R. Heintzmann, "Bayesian localization microscopy reveals nanoscale podosome dynamics," *Nat. Methods*, vol. 9, no. 2, pp. 195–200, 2012.

[39] R. Nieuwenhuizen, K. Lidke, M. Bates, D. Leyton Puig, D. Grünwald, S. Stallinga, and B. Rieger, "Measuring image resolution in optical nanoscopy," *Nat. Methods*, vol. 10, no. 6, pp. 557–562, June 2013.

[40] P. Sengupta, T. Jovanovic-Talisman, D. Skoko, M. Renz, S. Veatch, and J. Lippincott-Schwartz, "Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis," *Nat. Methods*, vol. 8, no. 11, pp. 969–975, 2011.

[41] D. Owen, C. Rentero, J. Rossy, A. Magenau, D. Williamson, M. Rodriguez, and K. Gaus, "PALM imaging and cluster analysis of protein heterogeneity at the cell surface," *J. Biophoton.*, vol. 3, no. 7, pp. 446–454, 2010.

[42] A. Löschberger, S. van de Linde, M. Dabauvalle, B. Rieger, M. Heilemann, G. Krohne, and M. Sauer, "Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution," *J. Cell Sci.*, vol. 125, no. 3, pp. 570–575, 2012.

[43] A. Szymborska, N. Marco, A. de Daigle, V. Cordes, J. Briggs, and J. Elllenberg, "Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging," *Sci. Express*, vol. 341, no. 6146, pp. 655–658, Aug. 2013.

[44] S. B. van Engelenburg, G. Shtengel, P. Sengupta, K. Waki, M. Jarnik, S. D. Ablan, E. O. Freed, H. F. Hess, and J. Lippincott-Schwartz, "Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits," *Science*, vol. 343, no 6171, pp. 653–656, 2014.

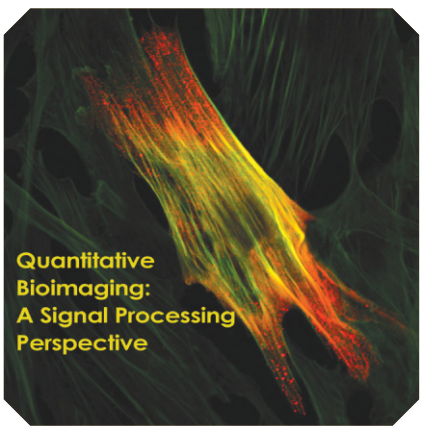
[SP]

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

Raimund J. Ober, Amir Tahmasbi, Sripad Ram, Zhiping Lin, and Elizabeth Sally Ward

Quantitative Aspects of Single-Molecule Microscopy

Information-theoretic analysis of single-molecule data



Single-molecule microscopy is a relatively new optical microscopy technique that allows the detection of individual molecules such as proteins in a cellular context. This technique has generated significant interest among biologists, biophysicists, and biochemists, as it holds the promise to provide novel insights into subcellular processes and structures that otherwise cannot be gained through traditional experimental approaches. Single-molecule experiments place stringent demands on experimental and algorithmic tools due to the low signal levels and the presence of significant extraneous noise sources. Consequently, this has necessitated the use of advanced statistical signal- and image-processing techniques for the design and analysis of single-molecule experiments. In this tutorial article, we provide an overview of single-molecule microscopy from early works to current applications and challenges. Specific emphasis will be on the quantitative aspects of this imaging modality, in particular single-molecule localization and resolvability, which will be discussed from an information-theoretic perspective. We review the stochastic framework for image formation, different types of estimation techniques, and expressions for the Fisher information matrix. We also discuss several open problems in the field that demand highly nontrivial signal processing algorithms.

Digital Object Identifier 10.1109/MSP.2014.2353664
Date of publication: 5 December 2014

INTRODUCTION

Optical microscopy has a long history going back several centuries during which it was a key technique for the discovery of biological processes. The basic optical principles have not changed, but what has changed in the instrumentation in recent decades is the availability of highly sensitive detectors, computer control, and powerful laser-based light sources [1], [2]. With these improvements in instrumentation came the possibility to analyze the acquired microscopy data using advanced signal and image processing techniques (see, e.g., [3] and [4]). Equally important, however, are the major advances in molecular biology and physical chemistry that have drastically improved the available technology for the labeling of cellular specimens [5]–[7].

These technological developments coincided with a time when the revolution in molecular biology has demanded powerful exploratory tools for the investigation of molecular processes in cells [1], [7]. For example, through genomic analyses, biologists have identified a large array of proteins, such as growth factor receptors, that are known to play a role in cancer. Standard techniques in molecular biology and biophysics, e.g., X-ray crystallography, allow the study of these proteins to a very high level of detail. However, to investigate their biological functions, it is important that these proteins are studied in their cellular context.

Fluorescence microscopy is the imaging technique of choice for the study of molecular processes within cells due to its ability

1053-5888/15©2015IEEE

to detect specifically labeled proteins, receptors, molecules, or structures [2], [7], [8]. There are, however, two aspects of fluorescence microscopy that limit its power. The first aspect is the spatial resolution of optical microscopy, which is a measure of the ability to distinguish two closely spaced pointlike objects [9]. While molecular interactions occur on the low nanometer scale, classical resolution criteria predict a resolution limit in the range of several hundred nanometers [9]–[11]. The second aspect is the sensitivity of the technique. A fluorescent molecule emits only a limited number of photons [1], [12]. This fact, together with the limited resolution of an optical microscope, implies that in classical fluorescence microscopy only relatively large accumulations of fluorescent molecules are detected. These detection limitations of classical fluorescence microscopy, and in particular their associated averaging effects, stand in the way of examining the molecular processes and structures at the level of individual molecules, i.e., precisely at the level that is required to study these phenomena in full detail.

SINGLE-MOLECULE MICROSCOPY IS A RELATIVELY NEW OPTICAL MICROSCOPY TECHNIQUE THAT ALLOWS THE DETECTION OF INDIVIDUAL MOLECULES SUCH AS PROTEINS IN A CELLULAR CONTEXT

Single-molecule microscopy is a technique that promises to overcome the deficiencies of classical fluorescence microscopy by allowing the detection of individual molecules rather than larger accumulations of molecules [1], [12]. Single-molecule microscopy goes back to the work by W.E. Moerner and L. Kador published in 1989 [13], followed by that of M. Orrit and J. Bernard published in 1990 [14]. Among the many stages of development, we mention just a few. In 1991, the image of a single-molecule was recorded for the first time [15]. In 2003, single-molecule microscopy played a crucial role in the measurement of the step size that the molecular motor myosin V takes in moving along an actin filament in an in vitro model [16]. This was based on being able to estimate the location of the myosin V molecule within 1.5 nm [16]. The green fluorescent protein (GFP) brought about a major breakthrough in fluorescent microscopy of proteins in living cells as the protein of interest can be genetically tagged by the GFP gene [5], [6]. The first single-molecule experiments in live cells using a GFP tag were reported in [17] and [18]. In a series of papers, it was recognized that the classical resolution criteria do not apply and distances well below those criteria can be measured using single-molecule microscopy [10], [11], [19]. One of the key observations was that resolution is significantly improved if the molecules to be imaged are not excited at the same time [20]. Various photophysical processes were investigated such as blinking [19], photobleaching [11], and photoswitching [21]. This knowledge was exploited in [21]–[23] when it was recognized that various fluorophores can be stochastically excited, which allows only a small number of the total fluorophores present in a sample to be imaged at any time point. This led to the development of localization-based superresolution microscopy techniques [21]–[23]. The development of techniques continues at a significant rate with the introduction of new approaches and refinements of existing ones.

IMAGE FORMATION

Figure 1(a) shows the schematic of an optical microscope. Excitation light from the light source is reflected off a dichroic mirror and passes through an objective lens to illuminate a fluorescent object (e.g., a point source) that is located in the object space. The fluorescence signal from the object is collected by the same objective lens, then passes through the dichroic mirror and an emission filter, and is focused on a detector by a tube lens. Image formation in an optical microscope can be described by optical diffraction theory [9]. A fluorophore, i.e., the fluorescent label of a single molecule, is typically modeled as a point source (i.e., a Dirac delta function) and as such its image is given by the point spread function (PSF), i.e., the impulse response, of the microscope [9]. For an in-focus single molecule, classical diffraction theory predicts that the image can be described by an Airy profile [see Figure 1(b)] whose analytical expression is given by [9]

$$f(\boldsymbol{r}) = \frac{J_1^2(\alpha \| \boldsymbol{r} \|)}{\pi \| \boldsymbol{r} \|^2}, \quad \boldsymbol{r} := (x, y) \in \mathbb{R}^2, \tag{1}$$

where α characterizes the width of the profile, J_1 denotes the first-order Bessel function of the first kind, and $\| \cdot \|$ denotes the Euclidean norm. It is important to note that the Airy profile may not be an accurate model in practice and more advanced PSF models are available (see the section "Stochastic Description of Single-Molecule Data") [24]–[26]. In addition, as will be discussed in the section "Imaging in Three Dimensions," the image of an out-of-focus single molecule depends strongly on the distance from the plane of focus and is distinct from the Airy profile [9], [24]. A fluorescent object can be described as a collection of closely spaced single molecules. As an optical microscope can be modeled as a linear shift-invariant system [9], the image of a fluorescent object is the superposition of the images of point sources at the locations of the single molecules, i.e., the superposition of PSFs, translated according to the locations of the corresponding single molecules.

Most important from our perspective is that the image of a point source is not a point itself but has a nonzero width. Therefore, if there are too many single molecules in close proximity, their images will overlap and the individual single molecules can no longer be differentiated in the image. As a result, in many situations, information about the locations of the single molecules is lost in a fluorescence microscopy image. Therefore, one of the approaches in single-molecule microscopy is to overcome this crowding problem, i.e., to arrange the imaging experiment in such a way that the images of the single molecules are placed sparsely enough so that they can be properly separated. This crowding problem is of course closely related to the notion of resolution that will be the topic of the section "Every Photon Counts: A Fisher Information Approach to Resolution and Localization Accuracy."

LOCALIZATION AND TRACKING EXPERIMENTS

In this section, we discuss the principles behind two of the most important single-molecule experiments. The first one, a single-molecule tracking experiment, aims at obtaining the trajectories of individual molecules as they move in a cell [3], [4], [27]–[29]. The second one, a localization-based superresolution experiment, aims to provide an image with a resolution well beyond what is achievable by classical methods [19], [21–[23].

TRACKING SINGLE MOLECULES

The movement of molecules such as receptors and proteins in cells is crucial for the functioning of the cells [16], [27]. Despite the importance of these processes much remains unknown. Therefore, tracking experiments, i.e., experiments that record such dynamic behavior over time, are of particular importance [3]. To obtain the most detailed analysis, it is essential to carry out these experiments in live cells at the single-molecule level (see Figure 2).

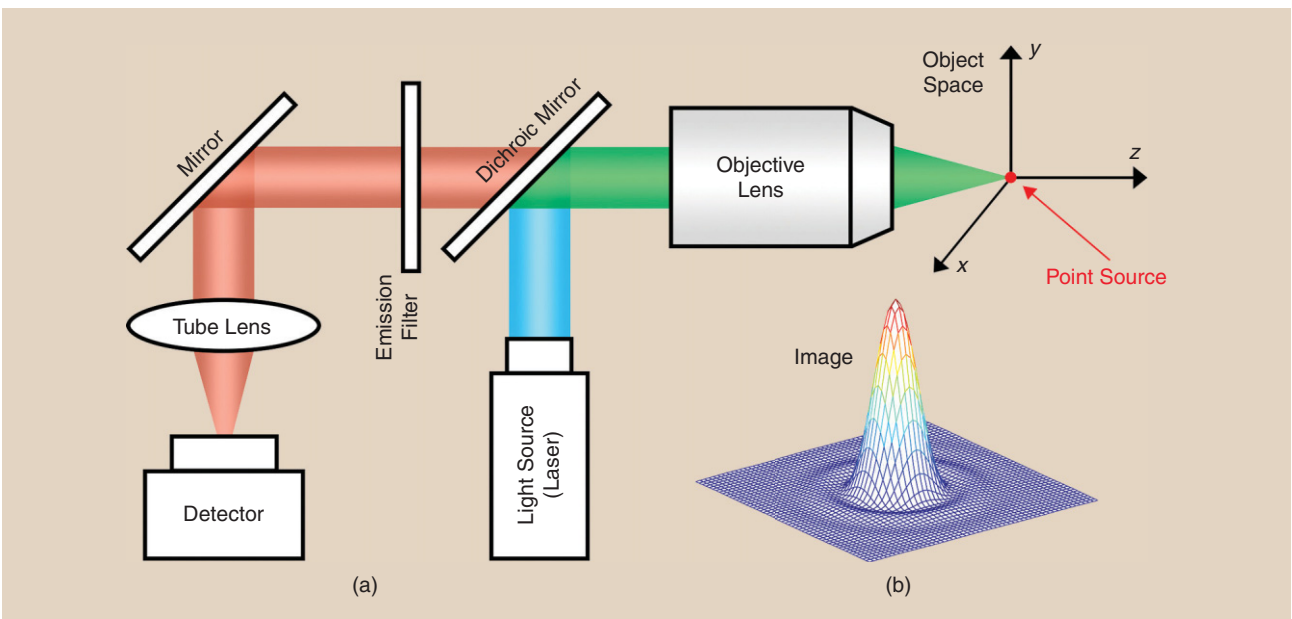
Such single-molecule tracking experiments, however, are not without significant challenges. Foremost among them is the need to be able to image isolated single molecules [3], [27] [see Figure 2(a)]. This can often be achieved with sparse labeling. Another significant problem is the photobleaching of many of the conventional fluorescent labels, which means that a fluorophore will only emit a certain, typically randomly distributed, number of photons before it ceases to emit photons [5], [6]. The phenomenon in effect limits the length of time for which the track of a single molecule can be followed.

MOST IMPORTANT FROM OUR PERSPECTIVE IS THAT THE IMAGE OF A POINT SOURCE IS NOT A POINT ITSELF BUT HAS A NONZERO WIDTH.

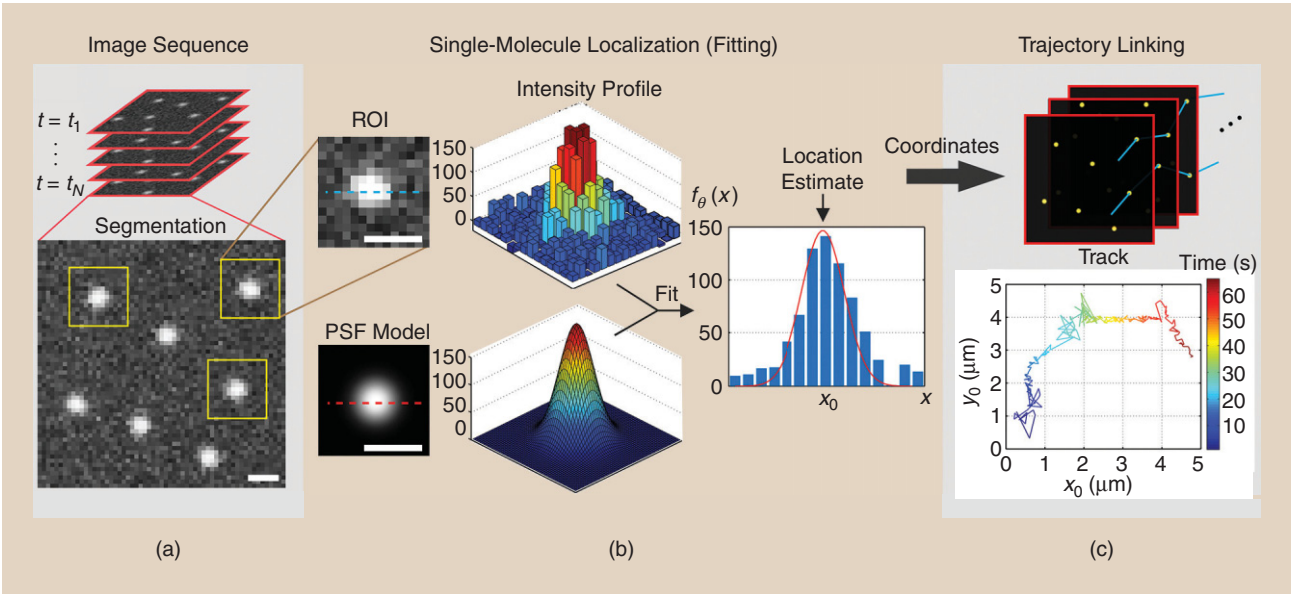
In designing a single-molecule tracking experiment, a number of important tradeoffs need to be made, in particular, regarding the frame rate of the acquisition and the associated exposure time for each of the images. High frame rates and corresponding short exposure times allow for better sampling of the dynamics of the single molecule. Reducing the exposure time, however, decreases the number of photons that are detected during the exposure interval and thereby, as will be shown later, will reduce the accuracy with which the parameters can be estimated that are associated with the trajectory [12], [27], [29]. Increasing the excitation light power could be used to increase the number of emitted photons per exposure. However, this will reduce the lengths of trajectories that can be imaged due to photobleaching. In addition, subjecting a cellular sample to excitation light that is too powerful might damage the living cell that is being imaged.

LOCALIZATION-BASED SUPERRESOLUTION MICROSCOPY

The second prototype experiment involves the imaging of fixed, i.e., dead, cells to obtain very high-resolution information concerning subcellular structures. In a classical fluorescence microscopy experiment, all fluorophores are simultaneously excited and imaged with one single exposure. As explained earlier, with densely spaced fluorophores, the result is that the individual fluorophores cannot be distinguished in the acquired image [see, e.g., Figure 3(a) and (b)]. The idea that underlies localization-based superresolution microscopy is to image the sample a large number of times, but in each of the images that make up the full



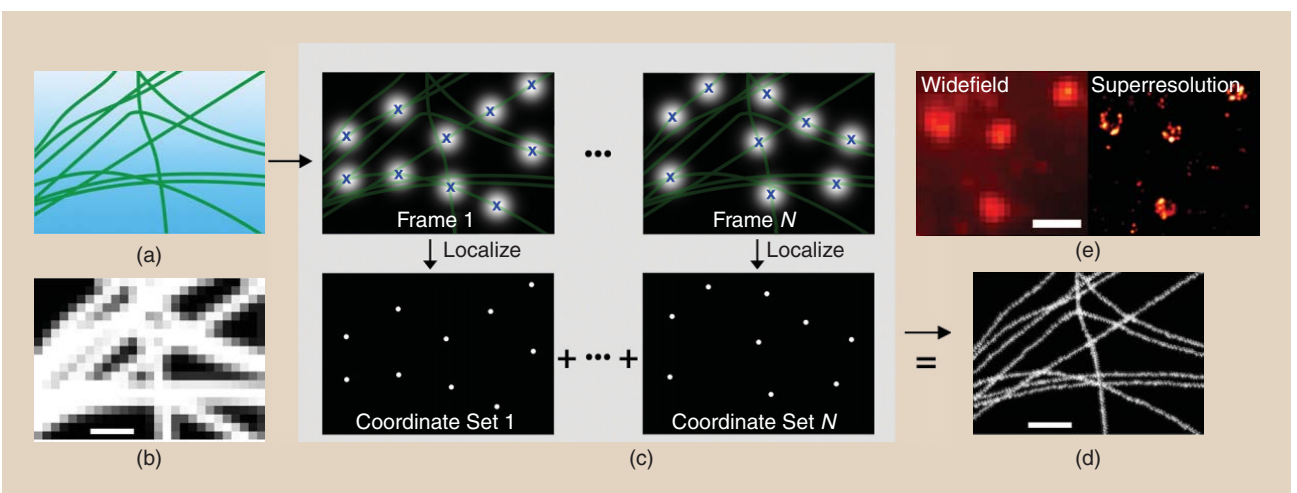
[FIG1] The schematic diagram of a basic fluorescence microscopy setup. (a) The excitation light, which is typically generated by a laser, passes through the objective lens to excite the fluorescent molecules in the object space. The fluorescent molecules emit photons at a specific wavelength that pass through the objective lens, the dichroic mirror, and the emission filter and are then collected by a detector. (b) The mesh plot of the image of an in-focus point source as seen on the detector plane.



[FIG2] Single-molecule tracking. (a) A sequence of images acquired at different time points are first segmented into multiple regions of interest (ROIs) each containing an isolated single molecule. (b) In the single-molecule localization step, a PSF model such as the Airy profile or a bivariate Gaussian distribution is fitted to each ROI to estimate the location of the single molecule with subpixel precision. This provides a set of coordinates of single molecules. (c) The set of coordinates together with its corresponding time points are then analyzed by a trajectory linking algorithm. In this way, the trajectory of each single molecule can be determined (a sample trajectory is shown). Size bars are 1 μm.

acquisition set, only a small and sparse subset of the fluorophores is imaged [see Figure 3(c)] [21], [22]. Through a particular choice of fluorescent labels, appropriate sample preparation and laser excitation, such sparse, random activation can in fact be achieved. The resulting images each are designed such that the positions of the sparsely located single molecules can be accurately determined. For each of the typically thousands of images, the locations of the single molecules are estimated [22], [23]. The final image is then assembled from the location estimates of the single molecules in each of the images [see Figure 3(d)].

Different techniques are available to produce these sparse subsets of fluorophores. These are primarily based on the exploitation



[FIG3] Localization-based superresolution microscopy. (a) The schematic shows a subcellular structure (a microtubule network) that is uniformly labeled with specific fluorophores. (b) In conventional imaging, all of the fluorophores in the sample are simultaneously excited. Due to the resolution limit of a fluorescence microscope, the resulting widefield image is poorly resolved and fails to reveal the underlying structure in the sample. (c) In localization-based superresolution microscopy, the imaging conditions facilitate activation of random subsets of fluorophores that are typically spatially well separated. These fluorophores are then localized with subpixel precision and their coordinates are used to create a superresolution image of the sample. (d) The resulting superresolution image provides fine structural information of the sample that is not accessible through a widefield image. (e) The comparison of a practical widefield image and a superresolution image. In (e), the size bar is 2 μm. In all other panels, size bars are 300 nm.

of new insights into the photophysics of fluorophores [11], [19], [20], whereby powerful excitation light sources can be used to stochastically excite subsets of fluorophores, put them in nonemitting states, or photobleach them. Depending on the specific mechanisms and fluorophores, these techniques are known as photoactivated localization microscopy (PALM), stochastic optical reconstruction microscopy (STORM), direct STORM, etc. [21]–[23].

BOTH THE SINGLE-MOLECULE TRACKING AND THE LOCALIZATION-BASED SUPERRESOLUTION EXPERIMENTS DEPEND ON THE ACCURATE DETERMINATION OF THE LOCATIONS OF THE IMAGED SINGLE MOLECULES.

STOCHASTIC DESCRIPTION OF SINGLE-MOLECULE DATA

Both the single-molecule tracking and the localization-based superresolution experiments depend on the accurate determination of the locations of the imaged single molecules [3], [29], [30]. To analyze the algorithmic aspects of the location estimation, it is necessary to carefully describe the data generation process that underlies fluorescence microscopy and, in particular, a single-molecule experiment. Before introducing a data model for the practical situation in which an image is acquired by a pixelated camera, it is useful to consider an idealized model. In this idealized model, termed the *fundamental data model*, we assume that the object being imaged emits photons as a Poisson process that are detected with a rate $\Lambda_\theta(\tau)$, $\tau \geq \tau_0$, on an infinitely large unpixelated detector [12], [31]. In this formulation, $\boldsymbol{\theta} \in \Theta$ denotes the parameter-vector of interest that contains the attributes of the object such as its position, where $\Theta \subseteq \mathbb{R}^n$ is an open parameter space. Making these assumptions allows us to ignore, for the time being, the deteriorating effects due to finite detector size, pixelation, and readout noise in the camera [32]. We assume that each photon is detected on the detector at a certain position that is distributed according to a two-dimensional (2-D) probability distribution $f_{\theta,\tau}(\boldsymbol{r})$, $\boldsymbol{r} = (x, y) \in \mathbb{R}^2$, where $\tau \geq \tau_0$ is the time of detection of the photon [12], [31]. This probability distribution is, in fact, the (continuous) image of the object at the particular time point, normalized such that $\int_{\mathbb{R}^2} f_{\theta,\tau}(\boldsymbol{r})\, d\boldsymbol{r} = 1$. For instance, this probability distribution can be the Airy profile [see (1)] or a bivariate Gaussian distribution [9], [12], [31], [33].

As an optical microscope is typically modeled as a linear shift-invariant system [9], the probability distribution function $f_{\theta,\tau}$ can be expressed in terms of an image function q:

$$f_{\theta,\tau}(x, y) = \frac{1}{M^2} q\Big(\frac{x}{M} - x_{0,\tau}, \frac{y}{M} - y_{0,\tau}\Big), \tag{2}$$

where $(x, y) \in \mathbb{R}^2$, $M > 0$ denotes the lateral magnification, and $(x_{0,\tau}, y_{0,\tau})$ is the position of the object at time $\tau \geq \tau_0$. The image function q describes the image of a stationary object that is located on the optical axis in the object space and is imaged at unit lateral magnification [12], [31]. In the case that the object is a point source, the image function is the same as the PSF of the microscope system.

In practice, the acquired data is corrupted by extraneous noise sources and by the pixelation that is introduced during the capture of the image by an imaging detector. In single-molecule experiments, the imaging detector is typically either a charge-coupled device (CCD) camera, complementary metal–oxide–semiconductor (CMOS) camera, or an electron multiplying CCD (EMCCD) camera [32], [34]. For the time being, we will concentrate on CCD or CMOS cameras and defer to the end of this section for the discussion regarding EMCCD cameras. We represent a pixelated detector with K_{pix} pixels as $\{C_1, \ldots, C_{K_{\text{pix}}}\}$, where $C_k \subseteq \mathbb{R}^2$ denotes the area occupied by the kth pixel of the detector. The acquired data at the kth pixel is given by $\mathcal{I}_{\theta,k} = S_{\theta,k} + B_k + W_k$, $k = 1, \ldots, K_{\text{pix}}$.

In the equation immediately above, $S_{\theta,k}$ denotes an independent Poisson random variable with mean $\mu_\theta(k)$ that describes the detected photon count from the object of interest [12], [35]; B_k denotes an independent Poisson random variable with mean b_k that describes the photon count due to background and scattering [31]; and W_k denotes an independent Gaussian random variable with mean η_k and variance σ_k^2 that describes the measurement noise that is introduced during the readout step in the detector [32]. The mean $\mu_\theta(k)$ of the random variable $S_{\theta,k}$ can be expressed in terms of Λ_θ and $f_{\theta,\tau}$, which describe the fundamental data model, and is given by [12], [31]

$$\begin{aligned}\mu_\theta(k) &= \int_{t_1}^{t_2} \int_{C_k} \Lambda_\theta(\tau) f_{\theta,\tau}(\boldsymbol{r})\, d\boldsymbol{r} d\tau \\ &= \frac{1}{M^2} \int_{t_1}^{t_2} \int_{C_k} \Lambda_\theta(\tau) q\Big(\frac{x}{M} - x_{0,\tau}, \frac{y}{M} - y_{0,\tau}\Big) d\boldsymbol{r} d\tau,\end{aligned}$$

for $k = 1, 2, \ldots, K_{\text{pix}}$, where $\boldsymbol{r} = (x, y) \in \mathbb{R}^2$, $[t_1, t_2]$ denotes the exposure time interval, and we have made use of (2). When the single molecule is stationary, the expression $\mu_\theta(k)$ reduces to

$$\mu_\theta(k) = \frac{N}{M^2} \int_{C_k} q\Big(\frac{x}{M} - x_0, \frac{y}{M} - y_0\Big) d\boldsymbol{r}, \quad k = 1, \ldots, K_{\text{pix}}, \tag{3}$$

where

$$N := \int_{t_1}^{t_2} \Lambda_\theta(\tau)\, d\tau$$

denotes the expected number of detected photons on an infinite detector plane [12], [36].

As we will see in the section "Every Photon Counts: A Fisher Information Approach to Resolution and Localization Accuracy," the readout noise in a CCD/CMOS detector can severely impair the quality of the acquired data, especially in the context of low signal levels, i.e., low photon counts. Therefore, over many decades significant efforts have been made to develop image intensifiers that amplify the signal before the readout process, with the expectation that this will minimize the detrimental effects of the readout noise on the measured signal. This is also the idea behind the EMCCD camera [37] that is widely used in single-molecule experiments.

The difficulty in analyzing the suitability of this and other amplification-based approaches lies in the fact that the amplification process is stochastic, which itself may imply a deterioration of the information content of the signal.

For an EMCCD camera, various probabilistic models have been proposed for the amplification process, which is in fact a branching process [34]. In [37], using a number of approximations, a binomial model was suggested for every stage of the amplification process. For the full process, approximate expressions were also derived for high photon counts in [37]. A comprehensive analysis of the modeling of the EMCCD amplification process was carried out in [34], where several approximate models were also investigated for their accuracy.

SINGLE-MOLECULE PARAMETER ESTIMATION

The benefits of single-molecule microscopy arise from being able to localize single molecules to very high precision [1], [12], 16]. The effective pixel size (i.e., the actual pixel size of the camera chip adjusted for the microscope magnification) in a standard microscope is typically in the range of 65×65 nm^2 to 400×400 nm^2. Localizing a single molecule up to a pixel would not bring any significant advantages, since the localization precision would be of the same order as that of the native resolution of the image [9]–[11] and, more importantly, biomolecular interactions typically occur at much lower distance scales. Therefore, it is necessary to localize single molecules with subpixel precision. This task is far from straightforward due to the often very low signal levels in the presence of significant noise sources, as discussed previously [32].

The first attempts were based on elementary approaches such as the center of gravity estimator [28], [38], while current algorithms are primarily based on fitting of a PSF model to the acquired data [see Figure 2(b)] [28], [33], [36]. The most frequently used fitting criterion is the least squares criterion [28], although the maximum likelihood estimator is better justified considering the probabilistic model of the acquired data [36]. Specifically, given the measured data $z_1, z_2, \ldots, z_K$ in the pixels that make up the ROI, which includes the image of the single molecule, the least squares criterion is given by [28]

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmin}_{\boldsymbol{\theta}\in\Theta} \sum_{k=1}^{K_{\text{pix}}} \| z_k - \nu_\theta(k) \|^2,$$

where $\nu_\theta(k) := \mu_\theta(k) + b_k$, and the maximum likelihood criterion is given by [12], [31], [35]

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmax}_{\boldsymbol{\theta}\in\Theta} \ln p_{\mathcal{I}_{\theta,1},\ldots,\mathcal{I}_{\theta,K_{\text{pix}}}}(z_1, \ldots, z_{K_{\text{pix}}} \,|\, \boldsymbol{\theta}),$$

where $p_{\mathcal{I}_{\theta,1},\ldots,\mathcal{I}_{\theta,K_{\text{pix}}}}(z_1, \ldots, z_{K_{\text{pix}}} \,|\, \boldsymbol{\theta}) := \prod_{k=1}^{K_{\text{pix}}} p_{\mathcal{I}_{\theta,k}}(z_k \,|\, \boldsymbol{\theta})$ denotes the joint probability distribution function (pdf) of the observed data. Considering the stochastic framework described in the previous section, for a CCD/CMOS detector, the pdf of the observed data at each pixel, for $k = 1, 2, \ldots, K_{\text{pix}}$, is given by [12], [31]

$$p_{\mathcal{I}_{\theta,k}}(z_k \,|\, \boldsymbol{\theta}) = \sum_{j=0}^{\infty} \frac{\nu_\theta(k)^j e^{-\nu_\theta(k)}}{j!} \frac{1}{\sqrt{2\pi}\,\sigma_k} e^{\frac{-(z_k - j - \eta_k)^2}{2\sigma_k^2}}.$$

The above expression shows that the observed data at each pixel of the detector has a Poisson–Gaussian mixture distribution where, as mentioned earlier, the Poisson and Gaussian parts model the photon detection and readout processes, respectively [31]. See [34] for the pdf for an EMCCD detector.

For the purpose of obtaining a localization-based superresolution image or for the purpose of single-molecule tracking, the main information that is necessary from this analysis is the location of the single molecule, i.e., the (x_0, y_0) coordinate. However, often other parameters also need to be estimated to be able to obtain the coordinate estimates. Examples include determining the width parameter of the image profile and the number of detected photons during the acquisition period.

The choice of image profile q in the estimation algorithms raises important questions. As discussed earlier, classical diffraction theory predicts a profile such as the Airy profile. However, very complex PSF models have been advocated to describe optical phenomena such as aberrations [9], [24] and the dipole nature of a single molecule [25]–[26], [39], or to deal with out-of-focus situations [9]. On the other hand, it has been argued that in many situations images of single molecules are adequately approximated by 2-D Gaussian functions [33], [38], [40] and, therefore, can be used for estimation purposes. It also needs to be recognized that, especially in the context of biological samples, even if there is a correct model, it is not likely that such a model can be identified with ultimate certainty due to the inherent variability of biological samples. There is also a tradeoff between computational complexity and the accuracy of the model of the resulting estimates. For instance, in localization-based superresolution microscopy, typically many tens of thousands of estimates have to be carried out to obtain one image [21], [22] and complex models are typically much more expensive to compute than simpler ones [41].

EVERY PHOTON COUNTS: A FISHER INFORMATION APPROACH TO RESOLUTION AND LOCALIZATION ACCURACY

An important topic in single-molecule microscopy has been the question of how well the different single-molecule estimation techniques perform in quantitative terms. This is a critical aspect in an experimenter's decision on whether the technique is appropriate for the scientific task, for experiment design and for the evaluation of algorithms. In general terms, there are two aspects that have received significant attention. One is the localization accuracy [12], i.e., the accuracy with which a single molecule can be localized. The second is resolution, which is (loosely speaking) the capability of the technique to distinguish different features in the sample [11]. When assessing the performance of a localization algorithm, its mean and standard deviation are most critical. Accuracy of the measurement is paramount even in the context of small data samples. Therefore, ideally unbiased estimators are sought with the lowest possible standard deviation [35]. While for general estimation problems, it is not always possible to obtain suitable unbiased estimators, many of the estimators that are currently applied in single-molecule microscopy have at

least numerically been shown to be unbiased [25], [36]. According to the Cramér–Rao lower bound, the (co) variance (matrix) of any unbiased estimator $\hat{\boldsymbol{\theta}}$ of a parameter (-vector) $\boldsymbol{\theta}$, such as the location parameters, is bounded from below by the inverse of the Fisher information matrix $\boldsymbol{I}(\boldsymbol{\theta})$ [12], [35], i.e.,

$$\operatorname{cov}(\hat{\boldsymbol{\theta}}) \geq \boldsymbol{I}^{-1}(\boldsymbol{\theta}).$$

The task of assessing the best accuracy with which the various parameters can be estimated therefore reduces to calculating the Fisher information matrix for the specific estimation problem and data model. In [31], a very general expression for the Fisher information matrix was derived for the fundamental data model, i.e., for the ideal case of an infinite detector without pixelation and in the absence of extraneous noise sources. Exploiting the nature of a spatiotemporal marked Poisson process [35], for a general image profile $f_{\theta,\tau}$ and photon detection rate $\Lambda_\theta(\tau)$, $t_1 \leq \tau \leq t_2$, we have [31]

$$\boldsymbol{I}(\boldsymbol{\theta}) = \int_{t_1}^{t_2} \int_{\mathbb{R}^2} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(\boldsymbol{r})} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(\boldsymbol{r})]}{\partial \boldsymbol{\theta}} \right)^T \times \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(\boldsymbol{r})]}{\partial \boldsymbol{\theta}} \right) d\boldsymbol{r} d\tau, \qquad \boldsymbol{\theta} \in \Theta.$$

Specializing this expression to the case of a constant photon detection rate, i.e., $\Lambda_\theta(\tau) = \Lambda$, $t_1 \leq \tau \leq t_2$, we immediately obtain that the Fisher information depends linearly on the number of photons detected [12], [31], i.e.,

$$\boldsymbol{I}(\boldsymbol{\theta}) = N \int_{\mathbb{R}^2} \frac{1}{f_{\theta,\tau}(\boldsymbol{r})} \left(\frac{\partial f_{\theta,\tau}(\boldsymbol{r})}{\partial \boldsymbol{\theta}} \right)^T \left(\frac{\partial f_{\theta,\tau}(\boldsymbol{r})}{\partial \boldsymbol{\theta}} \right) d\boldsymbol{r}, \qquad \boldsymbol{\theta} \in \Theta,$$

where $N := (t_2 - t_1)\Lambda$ is the expected number of photons during the exposure interval.

This implies that a lower bound on the standard deviation of the estimate of any parameter (-vector) has the form $(1/\sqrt{N})C$, where C is a constant (matrix) related to the specific parameter estimation problem. This is an important aspect of single-molecule microscopy. It shows that for algorithms that attain this bound, the accuracy of the parameter estimate depends reciprocally on the square root of the number of collected photons [12], [31].

For the case where the image function is the Airy profile and the single molecule can be assumed to be stationary, it can be shown that this expression implies the following limit on the standard deviation with which the x and y coordinates of the single molecule can be estimated [12]

$$\frac{1}{\sqrt{N}} \frac{\lambda}{2\pi n_a},$$

where λ is the wavelength of the emitted light and n_a is the numerical aperture of the microscope [9]. We refer to this lower bound as the *fundamental localization accuracy measure* (*FLAM*) [36].

THE ACCURACY OF THE PARAMETER ESTIMATE DEPENDS RECIPROCALLY ON THE SQUARE ROOT OF THE NUMBER OF COLLECTED PHOTONS.

The prior expressions are derived assuming the fundamental data model. For the practical data model, where we allow for a finite pixelated detector, background, and readout noise, an expression for the Fisher information matrix can also be derived as [34], [36]:

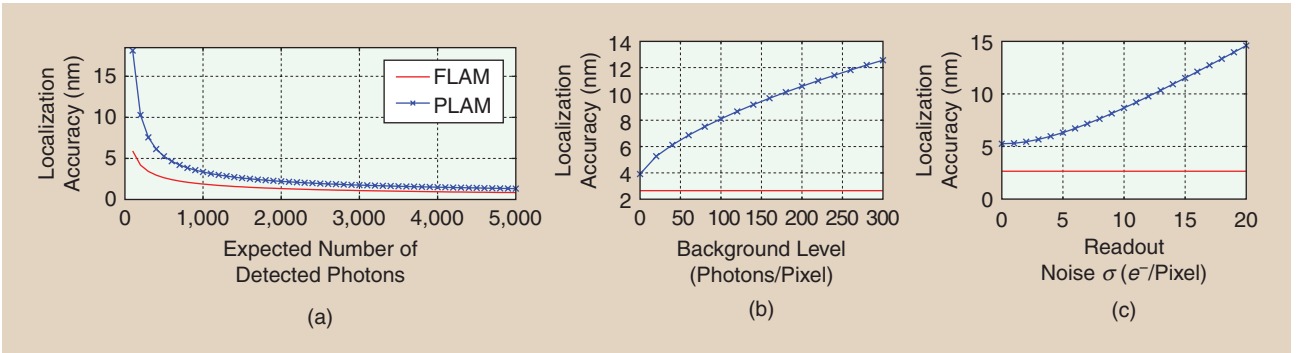
$$\boldsymbol{I}(\boldsymbol{\theta}) = \sum_{k=1}^{K_{\text{pix}}} \frac{\psi(k)}{\nu_\theta(k)} \left(\frac{\partial \mu_\theta(k)}{\partial \boldsymbol{\theta}} \right)^T \frac{\partial \mu_\theta(k)}{\partial \boldsymbol{\theta}}, \qquad \boldsymbol{\theta} \in \Theta, \tag{3}$$

where $\nu_\theta(k) = \mu_\theta(k) + b_k$ with $b_k, k = 1, \ldots, K_{\text{pix}}$, denoting the photon count due to the background noise at pixel C_k. The term $\psi(k), k = 1, \ldots, K_{\text{pix}}$, is referred to as the *noise coefficient* that depends on the type of detector [34]. In the absence of readout noise, $\psi(k) = 1$ for all $k = 1, \ldots, K_{\text{pix}}$ [12]. In the presence of readout noise and when using CCD and CMOS detectors, the noise coefficient is given by [31]

$$\psi(k) := \nu_\theta(k) \times \left(\frac{e^{-\nu_\theta(k)}}{\sqrt{2\pi}\,\sigma_k} \int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{\nu_\theta^{l-1}(k)}{(l-1)!} e^{\frac{-(z-l-\eta_k)^2}{2\sigma_k^2}} \right)^2}{\sum_{l=0}^{\infty} \frac{\nu_\theta^{l}(k)}{l!} e^{\frac{-(z-l-\eta_k)^2}{2\sigma_k^2}}} dz - 1 \right),$$

where η_k and σ_k^2 denote the mean and the variance of the readout noise at pixel $C_k, k = 1, 2, \ldots, K_{\text{pix}}$, respectively. The expression of the noise coefficient for an EMCCD camera is omitted for brevity but can be found in [34]. Using these expressions, a lower bound can be obtained on the standard deviation with which the x and y coordinates of the single molecule can be estimated in a practical situation. We refer to this lower bound as the *practical localization accuracy measure* (*PLAM*) (see [36]).

These expressions can be used to not only analyze the influence of pixelation but the various noise sources on the accuracy of the estimates of the location and other parameters. Importantly, these results can also be compared to those based on the fundamental expressions, which give us the theoretically best possible results and thereby let us understand how far a particular experimental configuration is away from the theoretically best possible one. For example, Figure 4 compares the behavior of the FLAM and PLAM versus the mean photon count and extraneous noise sources for a specific set of imaging conditions. The results can be reproduced using a free software package, the FandPLimitTool, available online at http://www.wardoberlab.com/software/. For small photon counts, the PLAM is significantly larger than FLAM implying that pixelation and extraneous noise worsen the localization accuracy whereas for large photon counts the difference is not appreciable [Figure 4(a)]. In addition, given a certain photon count, increasing the background noise [Figure 4(b)] and the readout noise [Figure 4(c)] considerably deteriorate the PLAM (when compared with the FLAM).



[FIG4] The single molecule localization measure. The behavior of the localization accuracy using the fundamental data model (i.e., the FLAM) and the practical data model (i.e., the PLAM) for the *x* coordinate of the single molecule as a function of (a) the expected number of detected photons N from the single molecule, (b) the background level b_k, and (c) the standard deviation of the readout noise σ_k. In (a)–(c), the numerical aperture is set to 1.4, the emission wavelength is set to 520 nm, the lateral magnification is set to 100, the pixel array (ROI) size is set to 25 × 25, and the pixel dimensions are set to 13 × 13 μm. In (a) and (c), b_k is 20 photons/pixel for all the pixels. In (a) and (b), the PLAM is calculated with $\sigma_k = 0\,e^-$/pixel for all the pixels. In (b) and (c), N is set to 500 photons.

Another approach to characterize the accuracy with which a single molecule can be localized has been proposed in [38] and [42]. Here, using a number of approximations, specific algorithms such as the least squares algorithm have been used, assuming a Gaussian image profile to obtain an expression for the standard deviation of the particular location estimator. However, great care needs to be taken in the use of these expressions as deviations from the actual performance of the algorithms have been observed when applied to images with Airy profiles [36].

The above analysis based on the Cramér–Rao lower bound has the advantage that it is independent of any particular estimation algorithm and gives bounds that any unbiased estimator needs to satisfy [31], [35]. From a practical point of view it is, however, important to know how well a particular algorithm performs in comparison to these bounds and whether there is an algorithm that attains the bounds. It is well known that, in general, assessing whether an algorithm attains the Cramér–Rao lower bound or to what extent it differs, is a theoretically difficult question and amenable to a theoretical analysis in only rare cases [35]. For the fundamental data model, in case the image is given by a Gaussian profile, it was shown in [12] that the maximum likelihood estimator reduces to the center of gravity estimator and attains the Cramér–Rao lower bound. For all other cases, no analytical analysis was possible but simulations have shown that the maximum likelihood estimator is consistently close to and, in some cases, attains the Cramér–Rao lower bound for a wide range of experimental conditions [12], [31], [36].

Classical resolution criteria for microscopy, such as Rayleigh's or Abbé's criterion, are heuristic criteria that were developed at a time when microscope samples were typically investigated by eye, rather than being recorded by a highly sensitive imaging detector [9], [10]. Therefore, the classical notions of resolution did not take into account the added benefits of a detailed analysis of the acquired data by sophisticated image and signal processing algorithms.

Resolution can be defined in a number of ways. One of the most fundamental ways relates to the question of the resolution of two point sources, which is the scenario Rayleigh's classical criterion addresses [9], [11], [43]. It states that two point sources can be resolved if they are separated by a distance of at least $0.61\lambda/n_a$ [9], [10]. Interestingly, this expression does not show any dependence on the amount of data that is acquired. In [11], this two-point resolution problem was cast in the aforementioned photon counting framework and the question was changed from "Can two points be resolved?" to the question of "How well can two points be resolved?" An expression for the limit on the standard deviation with which the distance d between two point sources can be estimated using the fundamental data model was then derived as [11]

$$\sigma_d = \frac{1}{\sqrt{4\pi N \Gamma_0(d)}} \frac{\lambda}{n_a}, \tag{4}$$

where N is the expected photon count on the infinite detector plane per point source and $\Gamma_0(d)$ is a nonlinear function of the distance between point sources (see [11]). Importantly this expression shows that arbitrarily small distances can be resolved, but the smaller the distance, the more photons need to be acquired to obtain the same accuracy. This approach can be generalized to multiple point sources in a relatively straightforward fashion (see, e.g., [11] and [44]).

The aforementioned information-theoretic resolution measure [i.e., (4)] is a powerful tool in determining how well two point sources can be resolved, and hence it is suited for applications where the structure of interest can be defined by a limited number of molecules. However, in other applications where continuous structures with a large number of potential labeling sites are imaged, the situation is more complex. One important aspect relates to the labeling density. As shown in [45], with decreasing density of the fluorescent labels, the structure of interest gradually becomes unresolvable in the acquired image, even when the previously discussed two-point resolution measure is appropriate. A resolution measure based on the Fourier ring correlation was recently published in [46]. This measure can be directly computed from the experimental data and takes into account the

localization accuracy, the density of fluorescent labels, and the spatial structure of the sample.

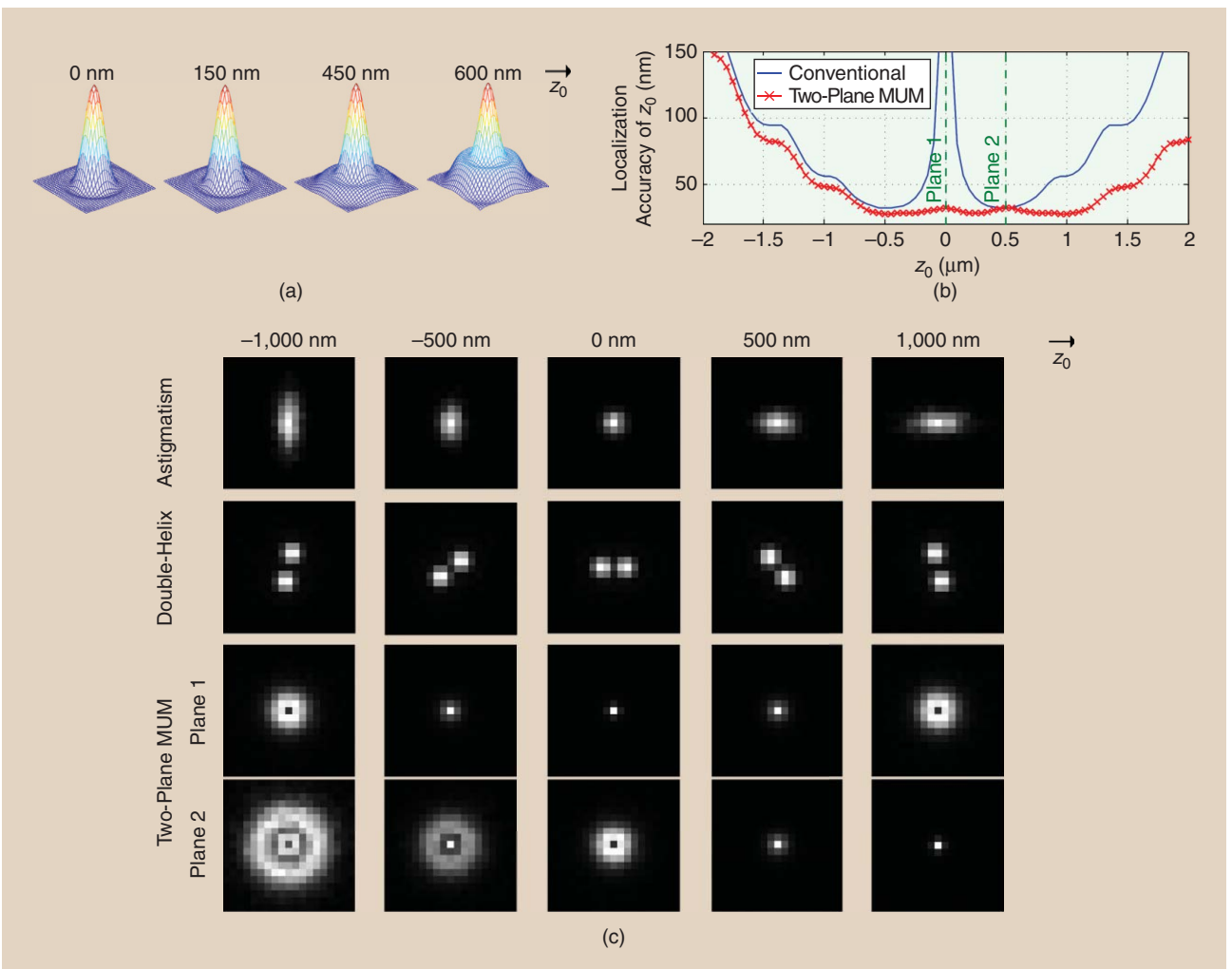
IMAGING IN THREE DIMENSIONS

Microscopy is, by its nature, a technique that is most suited to study phenomena that occur in one plane, i.e., the focal plane of the microscope [1], [12]. Cells, however, are three-dimensional (3-D) objects, and 3-D imaging of cellular processes poses several technical challenges, especially at the single-molecule level. In the previous section, we discussed results that showed that the x- and y-coordinates of an in-focus single molecule can be determined with very high accuracy. However, the situation changes dramatically when we are concerned with the estimation of the third spatial coordinate, i.e., the z-position of the single molecule. Considering the standard Born and Wolf 3-D PSF model [9], the image function, which now depends on the z_0-position, $z_0 \in \mathbb{R}$, of the single molecule, is given by [27]

$$q_{z_0}(\boldsymbol{r}) = A\left|\int_0^1 J_0(\alpha\rho\|\boldsymbol{r}\|)e^{j\frac{\pi n_a^2 z_0}{\lambda n_m}\rho^2}\rho d\rho\right|^2, \tag{5}$$

where $\boldsymbol{r} = (x, y) \in \mathbb{R}^2$, A is a normalization constant, $\alpha := 2\pi n_a/\lambda$, n_m denotes the refractive index of the immersion medium, and J_0 is the zeroth order Bessel function of the first kind [9]. As seen in Figure 5(a), if the single molecule is in focus, i.e., for $z_0 = 0$, the image of the single molecule is identical to the in-focus image we have seen in Figure 1(b). However, for out-of-focus positions, i.e., $z_0 \neq 0$, the image starts to depict out-of-focus rings with increasing z_0 and in general becomes flatter and more spread out.

Using the approaches based on the Cramér–Rao lower bound introduced in the section "Every Photon Counts: A Fisher Information Approach to Resolution and Localization Accuracy," we can also compute the accuracy with which the z-position of the single molecule can be determined, i.e., the



[FIG5] Single molecule imaging in 3-D. (a) An image profile of a point source at different z-positions acquired by a conventional (single plane) microscope. (b) A comparison of the localization accuracy, i.e., the PLAM, for the z coordinate of the single molecule along the z-axis for a conventional microscope and a two-plane multifocal place microscopy (MUM) setup. For a two-plane MUM setup, the PLAM predicts relatively constant z-localization accuracy for a range of z-positions including at the plane of focus (i.e., z_0= 0). (c) A comparison of 3-D single-molecule imaging approaches, which encode/deduce the z-position using different strategies.

PLAM, [see Figure 5(b)]. Inspecting this plot, we see that far away from the focal plane, i.e., above 1.5 μm, the localization accuracy of the z coordinate is very poor. Far from the focus, the spread out images are barely visible above the background [27]. It is therefore not surprising that little information can be obtained from them. This indicates that single molecules cannot be satisfactorily localized outside a certain distance from the focal plane.

What may, however, be surprising at first glance is that the accuracy of estimation for the z-position is also very poor when the single molecule is located close to the focal plane [see Figure 5(b)]. The reason for this phenomenon, which we refer to as the *depth discrimination problem*, is that, as can be seen in Figure 5(a), the images of a point source that is located close to the focal plane are barely distinguishable (compare profiles at z_0=0 and z_0=150 nm). Therefore, near the focus there is little information in the images of the single molecules about their precise z-positions. The images only start to show appreciable differences when the single molecule is farther from the focal plane [see Figure 5(a), where z_0= 450 and z_0= 600 nm]. Aberrations in the sample can reduce the depth discrimination problem [9], [24], [25], but the overall problem persists.

To address the depth discrimination problem, a number of approaches have been proposed. In [47], an astigmatic lens is used that introduces an elongation in the image of the single molecule when it is out of focus. As can be seen in Figure 5(c), this elongation occurs along different lateral axes depending upon whether the molecule is above or below the plane of focus. By determining the extent of elongation of the image profile, the z-location of the single molecule can be estimated. Approximate analytical expressions are proposed for the PSF of an astigmatic microscope, such as those based on 2-D elliptical Gaussian profiles [47].

In another approach [48], sophisticated optical designs have been employed to change the image of a single molecule. The result, shown in Figure 5(c), is a bimodal image profile that resembles a double helix and encodes the z-position as a rotation of the profile. The z-location of the single molecule is deduced by determining the change in the relative orientation of the bimodal peaks with respect to the in-focus image. A precise analytical expression is not available for the double helix PSF. Nevertheless, using approximate expressions, the double helix PSF has been shown to provide a relatively uniform z-localization accuracy along the z-axis [48].

Another approach, MUM, relies on the simultaneous imaging of several distinct focal planes within the sample [see Figure 5(c)] [27], [49]. This general approach, which is also known by slightly different terminology (e.g., [50] and [51]), produces multiple images of a single molecule that are acquired from different depths. The z-location of the single molecule is deduced by simultaneously fitting these images with appropriate 3-D PSF models [e.g., (5)]. Simultaneous imaging of different focal planes provides consistently more information about the z-position of the single molecule than a conventional microscopy image, even at the plane of focus [27]. This is possible since the Fisher information matrix for a MUM setup $\boldsymbol{I}_{\mathrm{MUM}}(\boldsymbol{\theta})$ is the sum of the Fisher information matrices of the individual focal planes $\boldsymbol{I}_k(\boldsymbol{\theta}), k = 1, \ldots, K_{\mathrm{pln}}$, due to the independence of data acquisition at each focal plane, i.e., we have

$$\boldsymbol{I}_{\mathrm{MUM}}(\boldsymbol{\theta}) = \boldsymbol{I}_1(\boldsymbol{\theta}) + \ldots + \boldsymbol{I}_{K_{\mathrm{pln}}}(\boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \Theta.$$

Therefore, the PLAM for MUM shows significant improvements in the z-localization accuracy when compared to a conventional microscope, as shown in Figure 5(b). Other approaches such as the iPALM are also proposed, which rely on interferometric optics [52]. All of the aforementioned approaches overcome the depth discrimination problem of conventional microscopy. A possibly competing criterion is related to the range of z-positions over which the single molecule can be localized to an acceptable accuracy. For a comparison of different 3-D imaging modalities, see, e.g., [53].

CURRENT CHALLENGES

Significant challenges remain in the analysis of single-molecule data. One of the assumptions that underlies the localization-based superresolution experiments is that during each acquisition only one single molecule is imaged in a ROI that allows for the localization of the single molecule [21], [22]. However, since the number of excited fluorophores in superresolution experiments is stochastic, it cannot be guaranteed that all imaged single molecules are isolated. Therefore, multiemitters might be present. Hence, there is a significant effort underway to find criteria to determine the number of single molecules in an ROI and to localize the individual single molecules that are present in the multiemitter region [44], [54]. It should be pointed out that these problems are highly nontrivial and are closely related to the resolution problem [11].

Additional problems arise from tracking experiments. Often it is assumed that the single molecules are stationary during each of the exposures that are taken to capture the single-molecule dynamics. While this can well be an appropriate assumption in many cases, in other experimental situations this is problematic [29]. To analyze this problem, the Fisher information matrix has been calculated in [55] for parameter estimation problems involving a deterministic trajectory during the exposure interval. Diffusion of single molecules on the plasma membrane is an important process that can reveal important biological information [27]. Clearly, diffusive behavior of a single molecule during the exposure of an image can have a significant impact on the resulting image. This process has been investigated in an approximate fashion in a series of papers [29], [56], and approaches have been proposed of how to infer the diffusion coefficient from the obtained images.

As discussed earlier, under specific imaging conditions, e.g., immobilized fluorophores, polarized excitation, and out-of-focus imaging, the dipole nature of a single molecule may become evident in the form of asymmetric image profiles [25], [39], [40]. This can be exploited to estimate the dipole orientation of the fluorophore. However, the analysis of such data is particularly challenging. For example, fitting an inappropriate image profile to the acquired data might lead to biased location estimates [26].

Conventional microscopy produces an image of a sample almost instantly so that the microscopist can immediately evaluate the outcome of the imaging experiment. Localization-based superresolution microscopy experiments, in contrast, require a large number of acquisitions and have a very significant computational overhead, as images of tens of thousands of single molecules often need to be analyzed and processed to produce the final reconstructed image [21], [22]. Single-molecule localization can be computationally complex and will by necessity require a nontrivial amount of computational time. To make the results of the analysis available to the microscopist as fast as possible, considerable efforts are made to speed up the calculations, e.g., by parallelizing the calculations on graphics processing units [30], [44].

THIS NEW MICROSCOPY MODALITY INHERENTLY RELIES ON IMAGE AND SIGNAL PROCESSING METHODOLOGIES SINCE THE CENTRAL COMPONENT OF THE APPROACH IS THE PRECISE DETERMINATION OF THE POSITIONS AND OTHER PARAMETERS OF THE IMAGED SINGLE MOLECULES.

CONCLUSIONS

We have reviewed a number of key quantitative aspects of single-molecule microscopy. Although this is a nascent field, it has created significant interest among biologists, biophysicists, and chemists who benefit tremendously from an imaging technique that allows molecular processes to be studied at the level of individual molecules. This new microscopy modality inherently relies on image and signal processing methodologies since the central component of the approach is the precise determination of the positions and other parameters of the imaged single molecules. This localization task is not trivial since the acquired image is characterized by a typically very low photon signal in the presence of significant noise sources. Estimation approaches and expressions for the Cramér–Rao lower bound were reviewed. While much progress has been achieved in a relatively short time, significant problems remain that can benefit from advanced signal processing algorithms.

ACKNOWLEDGMENT

This work was supported in part by the National Institutes of Health (R01 GM085575).

AUTHORS

Raimund J. Ober (ober@utdallas.edu) received the Ph.D. degree in electrical engineering from Cambridge University, United Kingdom, in 1987. In 1990, he joined the University of Texas at Dallas, Richardson, where he is currently a professor in the Department of Electrical Engineering. He is also an adjunct professor with the University of Texas Southwestern Medical Center at Dallas. His research interests include the development of signal and image processing algorithms for biomedical applications, particularly the estimation and tracking for single-molecule microscopy data. He is a Senior Member of the IEEE.

Amir Tahmasbi (a.tahmasbi@utdallas.edu) received the B.Sc. and M.Sc. degrees (with honors) in electrical engineering from Shiraz University of Technology, Iran, in 2008 and Iran University of Science and Technology, Tehran, in 2010, respectively. He is currently pursuing his Ph.D. degree in electrical engineering at the University of Texas at Dallas, Richardson. His research interests lie in the fields of statistical signal and image processing, and estimation theory with applications to single-molecule microscopy. He is a Student Member of the IEEE.

Sripad Ram (sripad.ram@utsouthwestern.edu) received the B.Sc. degree in applied sciences from PSG College of Technology, Coimbatore, India, in 1999; the M.Sc. degree in physics from the Indian Institute of Technology, Chennai, in 2001; and the Ph.D. degree in biomedical engineering from the University of Texas at Arlington in 2007. His research interests include statistical image processing of microscopy data and the development of imaging approaches for three-dimensional tracking of single molecules.

Zhiping Lin (ezplin@ntu.edu.sg) received the Ph.D. degree in information engineering from Cambridge University, United Kingdom, in 1987. Since 1999, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include multidimensional systems and signal processing, statistical and biomedical signal processing, and machine learning. He is currently the editor-in-chief of *Multidimensional Systems and Signal Processing*. He is a Senior Member of the IEEE.

Elizabeth Sally Ward (sally.ward@utsouthwestern.edu) received the Ph.D. degree from the Department of Biochemistry, Cambridge University, United Kingdom, in 1986. In 1990, she joined the University of Texas Southwestern Medical Center at Dallas, where she is a professor in the Department of Immunology and currently holds the Paul and Betty Meek-FINA Professorship in Molecular Immunology. Her research interests include antibody engineering, molecular mechanisms that lead to autoimmune disease, questions related to the in vivo dynamics of antibodies, and the development of advanced microscopy techniques.

REFERENCES

[1] W. E. Moerner, "New directions in single-molecule imaging and analysis," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 104, no. 31, pp. 12596–12602, 2007.

[2] J. W. Lichtman and J. Conchello, "Fluorescence microscopy," *Nat. Methods*, vol. 2, no. 12, pp. 910–919, 2005.

[3] E. Meijering, I. Smal, and G. Danuser, "Tracking in molecular bioimaging," *IEEE Signal Processing Mag.*, vol. 23, no. 3, pp. 46–53, 2006.

[4] N. Chenouard, I. Smal, F. de Chaumont, M. Maška, I. F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, A. R. Cohen, W. J. Godinez, K. Rohr, Y. Kalaidzidis, L. Liang, J. Duncan, H. Shen, Y. Xu, K. E. G. Magnusson, J. Jaldén, H. M. Blau, P. Paul-Gilloteaux, P. Roudot, C. Kervrann, F. Waharte, J.-Y. Tinevez, S. L. Shorte, J. Willemse, K. Celler, G. P. van Wezel, H. Dan, Y. Tsai, C. O. de Solórzano, J.-C. Olivo-Marin, and E. Meijering, "Objective comparison of particle tracking methods," *Nat. Methods*, vol. 11, no. 3, pp. 281–289, 2014.

[5] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, "Green fluorescent protein as a marker for gene expression," *Science*, vol. 263, no. 5148, pp. 802–805, 1994.

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

[6] N. C. Shaner, R. E. Campbell, P. A. Steinbech, B. N. Giepmans, A. E. Palmer, and R. Y. Tsien, "Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein," *Nat. Biotechnol.*, vol. 22, no. 12, pp. 1567–1572, 2004.

[7] X. Michalet, F. F. Pinaud, L. A. Bentolila, J. M. Tsay, S. Doose, J. J. Li, G. Sundaresan, A. M. Wu, S. S. Gambhir, and S. Weiss, "Quantum dots for live cells, in vivo imaging, and diagnostics," *Science*, vol. 307, no. 5709, pp. 538–544, 2005.

[8] D. S. Lidke and K. A. Lidke, "Advances in high-resolution imaging—Techniques for three-dimensional imaging of cellular structures," *J. Cell Sci.*, vol. 125, no. 11, pp. 2571–2580, 2013.

[9] M. Born and E. Wolf, *Principles of Optics*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[10] X. Qu, D. Wu, L. Mets, and N. F. Scherer, "Nanometer-localized multiple single-molecule fluorescence microscopy," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 101, no. 31, pp. 11298–11303, 2004.

[11] S. Ram, E. S. Ward, and R. J. Ober, "Beyond Rayleigh's criterion: A resolution measure with application to single-molecule microscopy," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 12, pp. 4457–4462, 2006.

[12] R. J. Ober, S. Ram, and E. S. Ward, "Localization accuracy in single-molecule microscopy," *Biophys. J.*, vol. 86, no. 2, pp. 1185–1200, 2004.

[13] W. E. Moerner and L. Kador, "Optical detection and spectroscopy of single molecules in a solid," *Phys. Rev. Lett.*, vol. 62, no. 21, pp. 2535–2538, 1989.

[14] M. Orrit and J. Bernard, "Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal," *Phys. Rev. Lett.*, vol. 65, no. 21, pp. 2716–2719, 1990.

[15] W. P. Ambrose and W. E. Moerner, "Fluorescence spectroscopy and spectral diffusion of single impurity molecules in a crystal," *Nature*, vol. 349, no. 6306, pp. 225–227, 1991.

[16] A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin, "Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization," *Science*, vol. 300, no. 5628, pp. 2061–2065, 2003.

[17] R. M. Dickson, A. B. Cubitt, R. Y. Tsien, and W. E. Moerner, "On/off blinking and switching behaviour of single molecules of green fluorescent protein," *Nature*, vol. 388, no. 6640, pp. 355–358, 1997.

[18] R. J. Ober, C. Martinez, X. Lai, J. Zhou, and E. S. Ward, "Exocytosis of IgG as mediated by the receptor, FcRn: An analysis at the single-molecule level," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 101, no. 30, pp. 11076–11081, 2004.

[19] K. A. Lidke, B. Rieger, T. M. Jovin, and R. Heintzmann, "Superresolution by localization of quantum dots using blinking statistics," *Opt. Express*, vol. 13, no. 18, pp. 7052–7062, 2005.

[20] A. M. van Oijen, J. Köhler, J. Schmidt, M. Müller, and G. J. Brakenhoff, "3-Dimensional super-resolution by spectrally selective imaging," *Chem. Phys. Lett.*, vol. 292, no. 1–2, pp. 183–187, 1998.

[21] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[22] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods*, vol. 3, no. 10, pp. 793–795, 2006.

[23] S. T. Hess, T. P. K. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophys. J.*, vol. 91, no. 11, pp. 4258–4272, 2006.

[24] S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *J. Opt. Soc. Am. A*, vol. 9, no. 1, pp. 154–166, 1992.

[25] F. Aguet, S. Geissbühler, I. Märki, T. Lasser, and M. Unser, "Super-resolution orientation estimation and localization of fluorescent dipoles using 3-D steerable filters," *Opt. Express*, vol. 17, no. 8, pp. 6829–6848, 2009.

[26] M. P. Backlund, M. D. Lew, A. S. Backer, S. J. Sahl, and W. E. Moerner, "The role of molecular dipole orientation in single-molecule fluorescence microscopy and implications for super-resolution imaging," *ChemPhysChem*, vol. 15, no. 4, pp. 587–599, 2014.

[27] S. Ram, P. Prabhat, J. Chao, E. S. Ward, and R. J. Ober, "High accuracy 3D quantum dot tracking with multifocal plane microscopy for the study of fast intracellular dynamics in live cells," *Biophys. J.*, vol. 95, no. 12, pp. 6025–6043, 2008.

[28] M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophys. J.*, vol. 81, no. 4, pp. 2378–2388, 2001.

[29] X. Michalet, "Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium," *Phys. Rev. E*, vol. 82, no. 4, pp. 041914, 2010.

[30] C. S. Smith, N. Joseph, B. Rieger, and K. A. Lidke, "Fast, single-molecule localization that achieves theoretically minimum uncertainty," *Nat. Methods*, vol. 7, no. 5, pp. 373–375, 2010.

[31] S. Ram, E. S. Ward, and R. J. Ober, "A stochastic analysis of performance limits for optical microscopes," *Multidim. Sys. Sig. Proc.*, vol. 17, no. 1, pp. 27–57, 2006.

[32] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White, "Compensation for readout noise in CCD images," *J. Opt. Soc. Amer. A*, vol. 12, no. 2, pp. 272–283, 1995.

[33] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Appl. Opt.*, vol. 46, no. 10, pp. 1819–1829, 2007.

[34] J. Chao, E. S. Ward, and R. J. Ober, "Fisher information matrix for branching processes with application to electron-multiplying charge-coupled devices," *Multidim. Syst. Signal Process.*, vol. 23, no. 3, pp. 349–379, 2012.

[35] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*, 2nd ed. New York: Springer-Verlag, 1991.

[36] A. V. Abraham, S. Ram, J. Chao, E. S. Ward, and R. J. Ober, "Quantitative study of single molecule location estimation techniques," *Opt. Express*, vol. 17, no. 26, pp. 23352–23373, 2009.

[37] J. Hynecek and T. Nishiwaki, "Excess noise and other important characteristics of low light level imaging using charge multiplying CCDs," *IEEE Trans. Electron Devices*, vol. 50, no. 1, pp. 239–245, 2003.

[38] R. E. Thompson, D. R. Larson, and W. W. Webb. "Precise nanometer localization analysis for individual fluorescent probes," *Biophys. J.*, vol. 82, no. 5, pp. 2775–2783, 2002.

[39] J. Enderlein, E. Toprak, and P. R. Selvin, "Polarization effect on position accuracy of fluorophore localization," *Opt. Express*, vol. 14, no. 18, pp. 8111–8120, 2006.

[40] S. Stallinga and B. Rieger, "Accuracy of the Gaussian point spread function model in 2D localization microscopy," *Opt. Express*, vol. 18, no. 24, pp. 24461–24476, 2010.

[41] S. B. Andersson, "Localization of a fluorescent source without numerical fitting," *Opt. Express*, vol. 16, no. 23, pp. 18714–18724, 2008.

[42] N. Bobroff, "Position measurement with a resolution and noise limited instrument," *Rev. Sci. Instrum.*, vol. 57, no. 6, pp. 1152–1157, 1986.

[43] M. Shahram and P. Milanfar, "Imaging below the diffraction limit: A statistical analysis," *IEEE Trans. Image Processing*, vol. 13, no. 5, pp. 677–689, 2004.

[44] F. Huang, S. L. Schwartz, J. M. Byars, and K. A. Lidke, "Simultaneous multiple-emitter fitting for single molecule super-resolution imaging," *Biomed. Opt. Express*, vol. 2, no. 5, pp. 1377–1393, 2011.

[45] H. Shroff, C. G. Galbraith, J. A. Galbraith, and E. Betzig, "Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics," *Nat. Methods*, vol. 5, no. 5, pp. 417–423, 2008.

[46] R. P. J. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grünwald, S. Stallinga, and B. Rieger, "Measuring image resolution in optical nanoscopy," *Nat. Methods*, vol. 10, no. 6, pp. 557–562, 2013.

[47] L. Holtzer, T. Meckel, and T. Schmidt, "Nanometric three-dimensional tracking of individual quantum dots in cells," *Appl. Phys. Lett.*, vol. 90, p. 053902, Feb. 2007.

[48] S. R. P. Pavani and R. Piestun, "Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system," *Opt. Express*, vol. 16, no. 26, pp. 22048–22057, 2008.

[49] P. Prabhat, S. Ram, E. S. Ward, and R. J. Ober, "Simultaneous imaging of different focal planes in fluorescence microscopy for the study of cellular dynamics in three dimensions," *IEEE Trans. Nanobiosci.*, vol. 3, no. 4, pp. 237–242, 2004.

[50] S. Abrahamsson, J. Chen, B. Hajj, S. Stallinga, A. Y. Katsov, J. Wisniewski, G. Mizuguchi, P. Soule, F. Mueller, C. D. Darzacq, X. Darzacq, C. Wu, C. I. Bargmann, D. A. Agard, M. Dahan, and M. G. L. Gustafsson, "Fast multicolor 3D imaging using aberration-corrected multifocus microscopy," *Nat. Methods*, vol. 10, no. 1, pp. 60–63, 2013.

[51] M. F. Juette, T. J. Gould, Mark. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Bewersdorf, "Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples," *Nat. Methods*, vol. 5, no. 6, pp. 527–529, 2008.

[52] G. Shtengel, J. A. Galbraith, C. G. Galbraith, J. Lippincott-Schwartz, J. M. Gillette, S. Manley, R. Sougrat, C. M. Waterman, P. Kanchanawong, M. W. Davidson, R. D. Fetter, and H. F. Hess, "Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 9, pp. 3125–3130, 2009.

[53] M. Badieirostami, M. D. Lew, M. A. Thompson, and W. E. Moerner, "Three-dimensional localization precision of the double-helix point spread function versus astigmatism and biplane," *Appl. Phys. Lett.*, vol. 97, no. 16, p. 161103, 2010.

[54] L. Zhu, W. Zhang, D. Elnatan, and B. Huang, "Faster STORM using compressed sensing," *Nat. Methods*, vol. 9, no. 7, pp. 721–723, 2012.

[55] Y. Wong, Z. Lin, and R. J. Ober, "Limits of the accuracy of parameter estimation for moving single molecules imaged by fluorescence microscopy," *IEEE Trans. Signal Processing*, vol. 59, no. 3, pp. 895–911, 2011.

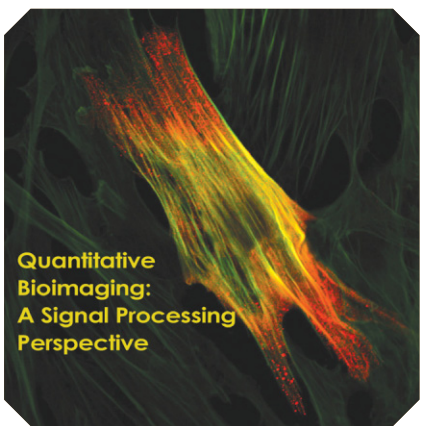
[56] X. Michalet and A. J. Berglund, "Optimal diffusion coefficient estimation in single-particle tracking," *Phys. Rev. E*, vol. 85, p. 061916, June 2012.

[SP]

Lei Qu, Fuhui Long, and Hanchuan Peng

3-D Registration of Biological Images and Models

Registration of microscopic images and its uses in segmentation and annotation



The registration, segmentation, and annotation of microscopy images and respective biological objects (e.g., cells) are distinct challenges often encountered in bioimage informatics. Here we present several studies in widely used model systems of the fruit fly, zebrafish, and *C. elegans* to demonstrate how registration methods have been employed to align three-dimensional (3-D) brain images at a very large scale and to solve challenging segmentation and annotation problems for 3-D cellular images. Specifically, we consider two types of registration between images and models: image-to-image registration and model-to-image registration, where a *model* consists of a description of the geometrical shape or the spatial layout of biological objects in the respective images.

INTRODUCTION

The registration of objects or patterns (e.g., cells with a globular shape, gene expression patterns, and highly irregular arborization patterns of neurons) is a commonly used technique in biological and medical data analysis. Generally speaking, registration is a process to map one image, object, or pattern to another (often obtained from different sensors, times, subjects, etc.) so that they can be compared, analyzed, or visualized directly within the same coordinate system. A spatial coordinate system is often considered. Along with the development of time-lapse light microscopy, the registration of a time series of images is also common and deemed important for many developmental biology studies. As an enabling technique in many applications such as building digital atlases, assessing the invariance (stereotypy) of patterns, profiling neuron connectivity, and studying the variation of cell populations, registration is essential in large-scale bioimage visualization, analysis, data mining, and informatics fields [1]–[3].

Segmentation and annotation of microscopy images and the respective biological objects are two challenging topics in bioimage analysis and informatics [1], [4], [5]. Segmentation refers to partitioning an image into multiple disjointed salient image regions, within each of which the image pixels share certain common characteristics. For 3-D cellular or brain images, the partitioned regions often represent interesting cells or compartments. In many cases, this partitioning process is realized by assigning a label to a group of pixels or by delineating the boundary of interesting objects and patterns. In contrast to segmentation, annotation is more closely related to the recognition of patterns or objects. Annotation often associates specific

Digital Object Identifier 10.1109/MSP.2014.2354060
Date of publication: 5 December 2014

1053-5888/15©2015IEEE

semantic properties such as the identities or categories to objects or patterns. Segmentation and annotation are critical to address important biological questions (e.g., quantification of gene expression patterns, generation of the ontology databases, and digital atlases of model animals).

FROM IMAGE-TO-IMAGE REGISTRATION TO MODEL-TO-IMAGE REGISTRATION

Registration is often needed to compare, fuse, or quantify objects or patterns in images. In many cases, registration is also required to map images to models and vice versa. In these latter situations, a *model* often consists of geometric shape description of the anatomy or spatial layout of biological objects in the respective images.

IMAGE-TO-IMAGE REGISTRATION

Many system biology studies rely on aligning images of gene expressions in different cell populations [6]–[8] or specimens that correspond to different developmental times [9]. In several recent brain mapping projects of the *Drosophila* (fruit fly), it became critical to align a number of 3-D confocal images of the insect's brains. Each fly had been genetically engineered to express fluorescent proteins in a specific population of neurons, which were aligned to a standard space so that they could be compared with each other [Figure 1(a)]. The FlyCircuit project in Taiwan [10] and the FlyLight project at the Janelia Research Campus of the Howard Hughes Medical Institute [11] each generated tens of thousands of 3-D fruit fly brain image stacks represented some of the biggest neuroscience efforts to date to understand the brain's structure. In each of these brains, some neuron populations are labeled using genetic methods. In both projects, registration of brain images is crucial. Registering images that correspond to the same population is useful to quantify the intrapopulation variability of neurons, which can further help define the meaningful neuron types. Registering images that correspond to different populations is useful to quantify the spatial proximity of neurons and thus helps estimate the putative connectivity of neurons. Similarly interesting results for the zebrafish (*Danio rerio*) were also reported recently [3], [4], [12].

Sophisticated volumetric image registration methods have been developed in the biomedical imaging field. Many methods, such as mutual information registration [13], spline-based elastic registration [14], invariant moment feature-based registration [15], and congealing registration [16], [17], have been widely used and extended to align molecular and cellular images. However, since many of them were originally designed for magnetic resonance imaging and computer tomography data, in many cases it remains challenging to use them easily and effectively in aligning the microscopy images that have larger-scale and fuzzier contents.

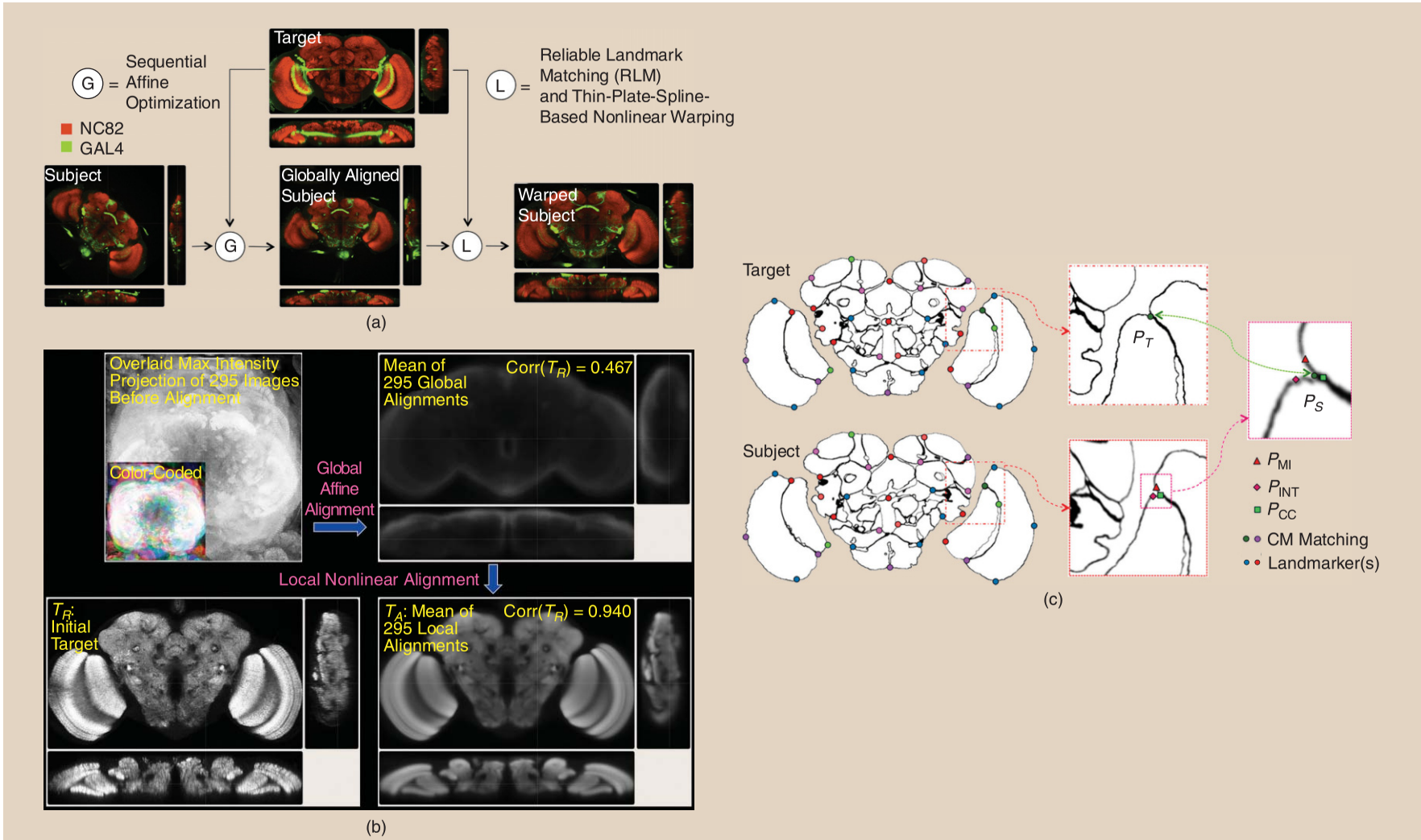
Two major challenges in biological image registration are the scale (in terms of the number and size of images) and variation of data (morphology or shape of patterns, image intensity, and noise level). For the first challenge, when the number of 3-D image stacks of brains increases to the order of tens of thousands and each image stack normally has the dimensions of 1,024 voxels (X) $\times$ 1,024 voxels (Y) $\times$ a few hundreds of voxels (Z), it will become exceedingly expensive to ask human annotators to supply even some simple prior knowledge of the data. The huge amount of image stacks requires that a successful registration scheme be highly automated, robust, and computationally efficient. These requirements limit the immediate applicability of many intensity-based registration methods in biomedical imaging field.

The second challenge is that the acquired microscopy image data not uncommonly display substantial variation of the appearance of the to-be-registered patterns. For instance, due to variable tissue labeling, light scattering, mismatching of reflective indexes of media along the light path, and many other issues in the automated image acquisition process, confocal microscopy data can exhibit a low signal-to-noise ratio. As in the fruit fly brain projects, an image normally comes with a neuropil staining that indicates the shape of the brain. Many times it is hard to threshold the neuropil image to segment the brain region from the image background. Therefore, it is often impractical to adopt boundary registration methods as used in the medical imaging field (see [15] for an example). In addition, complicated and varying shapes can arise from the flexible nature of specimens along with the sample preparation (e.g., tissue fixation). All these factors pose challenges to the image registration problem.

Many efforts were carried out to tackle these challenges. In an early effort of the FlyCircuit project, a simple affine transformation was used to align fruit fly brain images [10]. Unfortunately, the affine transformation is often not flexible enough to handle nonrigid deformations in images. In [18], 257 fruit fly brains are progressively registered using a method based on mutual information [19]. Such a method was also combined with multithreaded programming to accelerate the computation. However, nonsalient feature points used in registering different images can affect the accuracy of such a scheme.

BrainAligner [20] and ViBE-Z [3] are two programs developed recently to register sophisticated image patterns. ViBE-Z focuses on the registration of zebrafish brains. In such an application case, the image patterns consist of mainly line- and planelike structures [3]. ViBE-Z utilizes this feature by employing a trainable, rotation-invariant landmark detector. With 14 detected landmarks, a thin-plate spline transformation was used to perform a coarse but also elastic registration. Then, an intensity-based registration was used to realize a fine-scale elastic registration. In addition, a graph-based solver was used to determine the optimal deformation field in the fine elastic registration. This solver was shown to be efficient and less sensitive to local minima than commonly used gradient-descent methods.

We developed BrainAligner to detect the corresponding landmarks of any pair of images based on using a committee-machine algorithm [Figure 1(c)] to aggregate the feature matching results of a series of independent image feature analysis methods. In this way, the effect of pattern variation can be mitigated. The matched pairs of landmarks are further pruned using both the random sample consensus (RANSAC) algorithm [21] and tetrahedron pruning. RANSAC ensures all the corresponding landmark pairs form a globally consistent transform, which is the affine transform in our



[FIG1] The three-dimensional alignment of confocal image stacks of fruit fly brains. (a) The general scheme of image registration based on global alignment followed by nonlinear elastic local alignment. Red: reference channel (neuropil). Green: pattern channel (genetically labeled neurons or neuron-populations). (b) An example of aligning a number of fruit fly brain image stacks, each of which has a different orientation and size. Global alignment can only register these images to have approximately the similar size and orientation. Precise local alignment based on BrainAligner can produce sharply registered images. (c) An illustration of the first step in the reliable-landmark matching algorithm of the local alignment module in BrainAligner. Multiple independent matching criteria based on mutual information (MI), intensity (INT), and correlation (CC) are used to generate the initial candidates of matching locations (CM). Reliable-landmark matching will continue only when these candidate locations are close to each other. For more details regarding BrainAligner and reliable-landmark matching, see [20]. (Figure is adapted from [20] with permission.)

IEEE SignalProcessing
Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

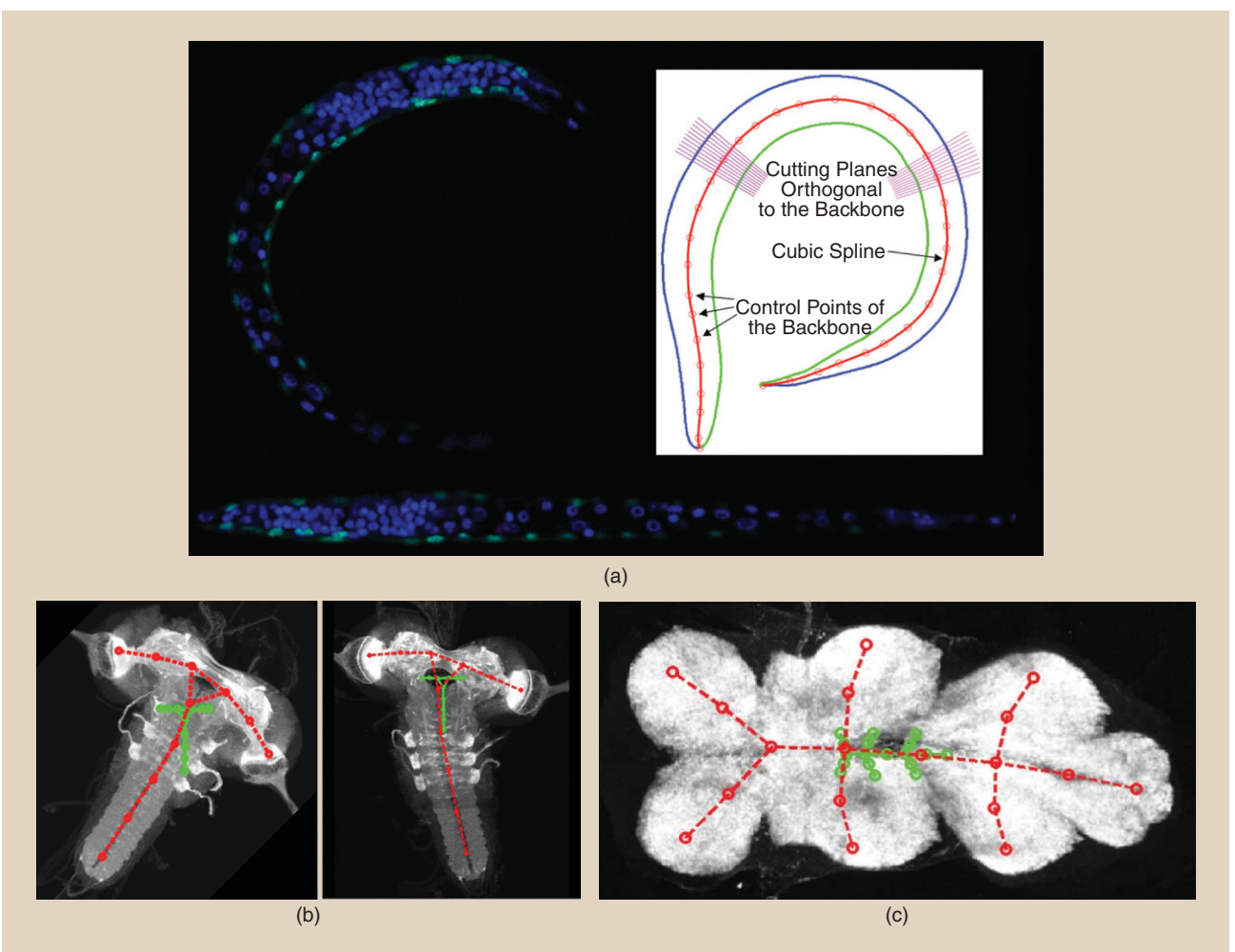
case. Tetrahedron pruning eliminates the cases of local self-intersection of corresponding landmark pairs and thus reduces the likelihood of occurrence of nonsmooth transform during registration. In addition, a hierarchical interpolation scheme for the 3-D thin-plate spline is employed in BrainAligner to quickly calculate the deformation field. Such an interpolation method considerably reduces both computation complexity and memory consumption of thin-plate spline warping. Together these components make BrainAligner robust to imperfect images (e.g., images of brains that have been partially damaged during sample preparation or images with fuzzy boundaries) and suitable for high-throughput processing. BrainAligner has aided a number of studies in fruit fly brain research by mapping neuron populations visualized using various genetic methods to a standard brain atlas model ([11], [20], [22]). This results in complete coverage of the fruit fly brain and a mesoscale connectome of the brain of the animal [23].

MODEL-TO-IMAGE REGISTRATION

PRINCIPAL SKELETON MODELS

Biological patterns often have highly curved, articulated, or branched structures. For instance, the bodies of *C. elegans* [Figure 2(a)] and zebrafish are usually curved. The fruit fly larval nervous system and ventral nerve cord of adult fruit fly have articulated shapes [Figure 2(b) and (c)]. The curved structure can be modeled as a lower-dimensional manifold pattern. A global affine transform is not suitable to globally register images of these patterns. Without being able to globally align these images, more detailed registration at local image regions will become impossible.

When the biological objects have an articulation or an embedded manifold, such patterns should be first globally standardized prior to the image-to-image registration (following the procedure discussed in the section "Image-to-Image Registration.")



[FIG2] Model-to-image registration and its use in standardization of articulated shapes that are often seen in microscopy images of model animals. This process is done via detecting the principal skeletons of these shapes followed by unbending the structures using a smooth warp. (a) Detecting the center "backbone" curve of a *C. elegans* image stack (top left) and straightening this image by restacking resampled image data (bottom) of all cross-sectional planes orthogonal to the backbone curve (top right). (b) Registering an initial model (green) of a fruit fly larval nervous system to two different confocal images of this animal. The red color indicates the final detected principal skeletons (the control nodes are marked as small circles). Note that the same model was used in both examples to generate the correct results. (c) Registering an initial model (green) of a fruit fly adult ventral nerve cord to a confocal image of this animal. The red color indicates the final deformed principal skeleton (the control nodes are marked as small circles).

The *standardization* refers to unfolding the embedding-manifold structures or globally aligning the articulated components of objects so that they possess similar scales, positions, and directions.

To standardize a shape in the image, we first explicitly model the curved or articulated shape. A principal skeleton model [24] is suitable for this goal. The principal skeleton is defined by a set of connected polylines with intrinsic shape constraints embedded (Figure 2). For different shapes, different principal skeleton models should be created. The principal skeleton model of a shape should correspond to the simplest skeleton that is complicated enough to capture the major structure and major deformation of this shape. In the simplest case, a principal skeleton model consists of only a polyline without any branch, which is sufficient to capture the smoothly curved shapes in *C. elegans* [Figure 2(a)] or zebrafish. In a more complicated case, a connected multipolyline model is used to define the principal skeleton. This fits well the cases of fruit fly larval nervous system and adult ventral nerve cord [Figure 2(b) and (c)] .

A principal skeleton model can be deformed to best register to the image content. This skeleton model, however, may not be easily produced using many approaches such as [25]–[30]. For instance, when the boundary of the animal's shape is not available [Figure 2(a)], a skeleton cannot be derived directly from the shape of the animal. Such cases are not uncommon in microscopy images. To solve this problem, we produced an optimized principal skeleton model for an image by iteratively mapping a predefined principal skeleton onto the image [24], [31]. Specifically, one can progressively update the control points in the principal skeleton while preserving the topology of the linkage between control points. To drive the deformation process, we defined a cost function to optimize two competing terms: one external force called *image force* and one internal force called *model force*. The image force is designed to push the principal skeleton to span as broadly as possible to cover the entire image pattern. This is realized by first generating the Voronoi partition using all control points and then minimizing the distance between each control point and the center of mass of its corresponding Voronoi region. The model force is designed based on the shape prior defined by the principal skeleton. Such a force is then minimized to attain the shortest overall length and the greatest smoothness of the principal skeleton. Figure 2 shows examples in which the initial model can deform to best register to images.

For multiple image patterns that have articulated structures, once their principal skeleton models have been generated, a thin-plate spline can be employed to warp these image patterns to a common coordinate system [24]. Such a method has been successfully applied to *C. elegans*, a fruit fly larval nervous system, and ventral nerve cord image data to perform more accurate global registration. Then local alignment methods such as BrainAligner can be used more effectively to generate high-resolution local registration.

SPATIAL LAYOUT ATLAS MODELS

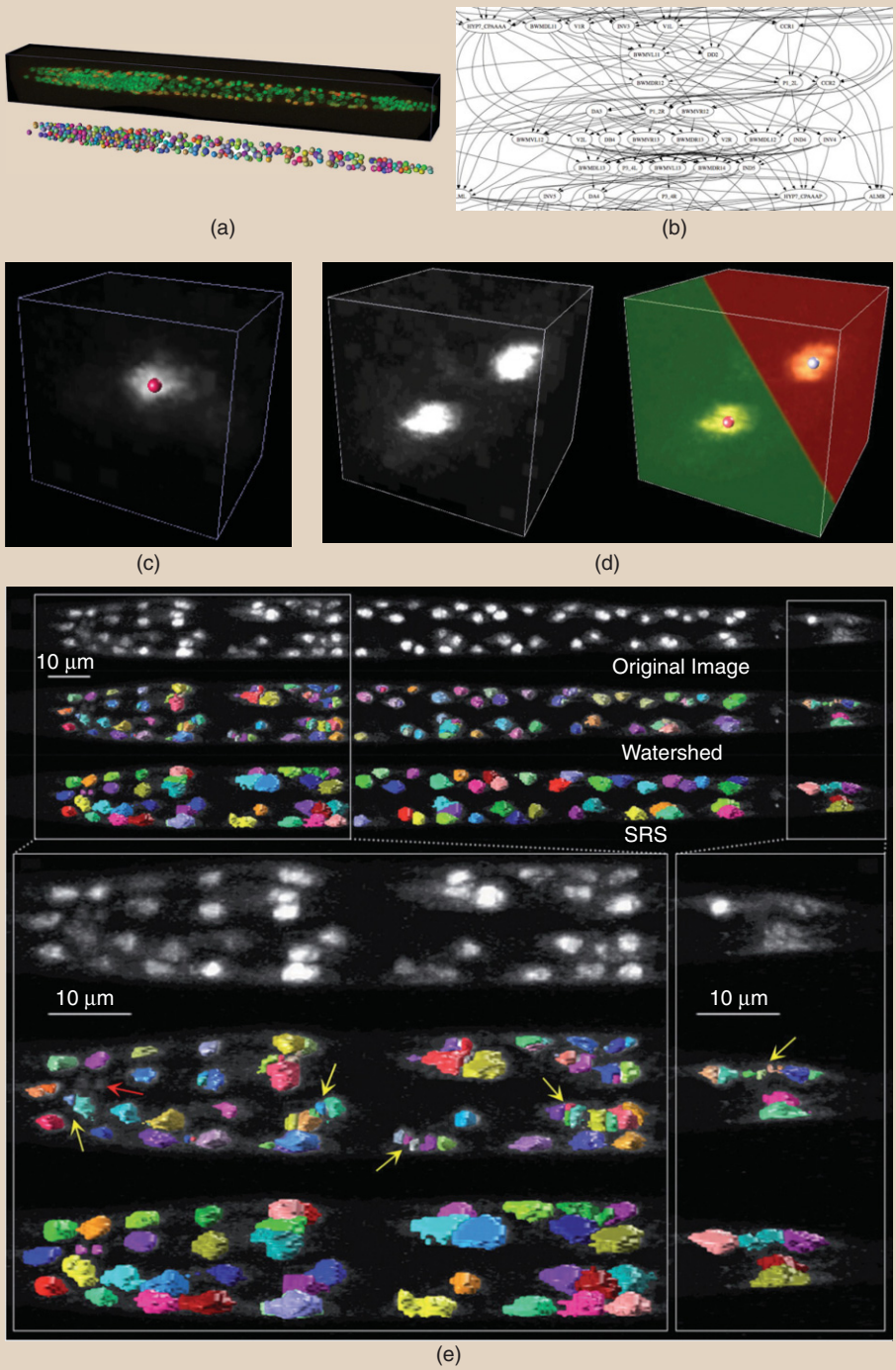
In some cases, the model may need to be much more complicated than the aforementioned principal skeleton. One piece of essential information is the complex 3-D spatial layout of objects. In addition, the model may also incorporate the objects' identities or some statistical information such as cell shape, size, and position variation, etc. [32]. With a complex version of the model, the model-to-image registration can be further extended to solve segmentation and annotation problems.

Here we restrict our discussion on *C. elegans* cell segmentation and annotation. For neuron- and whole-organism-level segmentation and annotation, we refer interested readers to [29], [33], and [34]. *C. elegans* is a model animal for a wide range of biological studies, from gene expression to brain function and even animal behavior [35]. This animal has an invariant number of cells, which also have invariant lineages during development. For the postembryonic *C. elegans*, a number of confocal images [Figure 3(a)] were segmented [32]. The results were further assembled as a 3-D digital atlas [Figure 3(b)] to describe the layout of cells at the single cell resolution [32]. This digital atlas can either be visualized in terms of a point-cloud [similar to Figure 3(a)] or a "wiring" graph of cells' relative locations [Figure 3(b)] in 3-D. The atlas was then used as a model to guide the recognition of cells in newly scanned 3-D image stacks of this animal.

Intuitively, recognition of these *C. elegans* cells could be achieved by first segmenting the cells in 3-D, followed by finding the correspondence between segmented cells in an image and the already standardized cells in the atlas model. Once cells have been segmented and recognized, useful information of cells, such as the expression level of specifically targeted cells, can be read out at these identified cellular locations. This routine was first developed in [36]. In the cell segmentation step, an optimized 3-D watershed algorithm was used. In the recognition step, since the relative locations of most cells are conserved from animal to animal, a graph-matching formulation of cell locations from the segmented cells to those recorded in the atlas was used. Both steps unavoidably had some errors. The biggest problem was that the information in the atlas (e.g., the number of cells, the variability of relative locations of cells) was not employed to help improve cell segmentation, which would also enhance the graph matching based recognition. In [37], the problem of over- and undersegmentation was alleviated by performing recognition on an abundant pool of segmentation hypotheses.

Instead of separating cell segmentation and recognition as two isolated processes, an alternative method is to perform segmentation and recognition in a simultaneous way with prior knowledge considered in both steps [38]. In short, this strategy was realized by "registering" the atlas to the image directly. The atlas itself in this case is a complex model that encodes both the identities and relative locations of all cells. The registration process is defined as deforming the 3-D locations of all cells in the model to best fit the cells in the image while keeping their relative locations. The cell segmentation in this case is implicitly realized via assigning a distinct group of image voxels to each cell.

To illustrate this idea, one may begin with a simplified case where there is only one cell in both the atlas and image. In this case, the best fit is apparently to move the cell's location to the center of mass of the image [Figure 3(c)]. In a slightly more complicated case where there are two cells (called u and v for



[FIG3] Three-dimensional segmentation and recognition of *C. elegans* cells. (a) Shown in the upper image is a 3-D confocal image stack of *C. elegans,* where different colors indicate different fluorescent labeling of cells; (a) (bottom) shows the point cloud representation of the 3-D segmentation result of this image stack, where different colors indicate different cells. The 3-D atlas is also often represented as a point cloud and visualized similar to the bottom of this picture. (b) A portion of the directed acyclic representation of the anterior-posterior location relationship in the 3-D atlas of *C. elegans*. The arrow from a cell *U* to a cell *V* means *U*'s location is always anterior of *V* in the atlas. Depicted in the middle of each circle (graph node) is the name of this cell. Similar left–right and dorsal–ventral graphs can be produced based on the atlas as well. (c) A schematic illustration of an image where there is only one cell and the optimal 3-D location of this cell should be the center of mass of image voxels. (d) A schematic illustration of an image where there are only two cells and the optimal 3-D locations of these two cells should be the centers of mass of the Vonoroi regions. (e) Results of simultaneous-segmentation and recognition of *C. elegans* via deforming an atlas model of all cells to best register to the 3-D image, and a comparison with the 3-D watershed segmentation, which has both under- and oversegmentation at different regions. (Image taken from [38] and used with permission.) For more details on the *C. elegans* atlas and the algorithm, see [35] and [38], respectively.

convenience) in the atlas and image [Figure 3(d)] and assuming the cell u is always in the left of the cell v. In this case, we would partition the image into two portions, each of which would be assigned to one cell, and move the cell's location to the center of mass of the respective partition. Finally, the constraint for cells' relative positions can be guaranteed by switching u and v if such a constraint is violated.

Biologically, this approach is suitable for the *C. elegans* cell recognition problem because the number of cells of the worm is a constant and the relative spatial locations of individual cells are highly constrained [35]. We formulated this approach into an algorithm called *simultaneous recognition and-segmentation of cells* [38]. Its optimization process consists of two iterative steps: 1) atlas-guided voxel classification and 2) voxel-classification-guided atlas deformation. A more detailed description is given in [38]. Interestingly, to make the algorithm more robust and efficient, several additional factors have also been considered [38]. First, because *C. elegans* is much more elongated along its anterior–posterior axis than the dorsal–ventral and left–right axes, the algorithm allows more flexible deformation of cells' locations along the anterior–posterior axis than the two other axes. Second, a "temperature"-modulated deterministic annealing optimization [39]–[41] was used to tackle the optimization problem by constraining the fuzziness of the classification probabilities. Thanks to this annealing method, simultaneous segmentation and recognition can even handle 180° flipped images [38]. Finally, to cope with the challenge of (usually) having an enormous amount of image voxels in a 3-D image, we downsampled the image before entering the iteration step. We also considered only sufficiently bright image pixels in the actual computation of likelihood and image partitioning (typically, only pixels with intensities greater than the average intensity of the image are included in the calculation). In the simultaneous segmentation and recognition result, the partition of the foreground image naturally translates to the segmented regions of cells.

Simultaneous segmentation and recognition has been applied to recognizing a number of cell types in *C. elegans*, including body wall muscle cells, intestinal cells, neurons, etc. It can recognize these cells reliably, even if the initial atlas of cells has a different orientation from the test image [38]. Simultaneous segmentation and recognition avoids many of the over- and undersegmentation problems [Figure 3(e)], compared to some widely used cell segmentation methods such as the watershed based [36], [42], graph-cut based [43], level-set based [44], and many other methods as mentioned in a recent review paper [5] and the many insight segmentation and registration toolkit methods wrapped up in the FARSIGHT project (see [45]). Such a feature indicates that this model-to-image registration-based approach can be used to solve challenging image segmentation in some situations.

DISCUSSION AND CONCLUSIONS

In this article, we introduced three cases of registration between 3-D images and models. We showed that registration-based approaches are useful for large-scale image alignment, as well as for the segmentation and annotation of 3-D cellular microscopy images. It is noteworthy that the generalization of registration-based approach can be further applied to other bioimage analysis problems. These analyzed results could be further visualized or annotated by widely used manual tools such as Vaa3D (http://vaa3d.org) [46] and CATMAID [47].

The model-to-image registration can be combined with image-to-image registration in a pipeline, thus the articulated objects in a bioimage can be meaningfully aligned. Model-to-image registration can also be combined with image tracking, a whole field of methods not discussed in this article, to analyze two-dimensional or 3-D video-based animal motion or development (e.g., *C. elegans* or zebrafish kinetic motion analysis). Another promising direction is to integrate all the steps of animal tracking, shape standardization, cell segmentation, and recognition with microscope hardware control to build an "intelligent" system that can simultaneously perturb cells and screen corresponding behaviors in vivo.

Despite the several examples we showed, we also found several challenges in registration methods and applications. There is also a lot of room to improve the related algorithms. For example, cell recognition and segmentation, despite the exploration of relative spatial location information and position variation statistics, still lacks an efficient method to use the cell shape and size priors embedded in the atlas. Sophisticated machine-learning techniques, such as supervised learning, can play interesting roles in its further development. Not limited to registration, another key factor of consideration in many bioimage analysis applications is whether or not the prior knowledge can be effectively modeled and utilized. We hope this article can inspire more research into signal processing, pattern recognition, and machine learning for robust bioimage analysis.

ACKNOWLEDGMENTS

We thank Xindi Ai, Katie Lin, and Rummi Ganguly for proofreading this manuscript during revision. Lei Qu was partially supported by Chinese Natural Science Foundation Project (61201396, 61301296, 61377006, U1201255); Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry; and the Technology Foundation for Selected Overseas Chinese Scholar, Ministry of Personnel of China.

AUTHORS

Lei Qu (qulei@ahu.edu.cn) received the Ph.D. degree in computer application techniques from Anhui University, China, in 2008. Between 2009 and 2011, he was a postdoctoral researcher at Howard Hughes Medical Institute–Janelia Farm Research Campus, Ashburn, Virginia, United States. He is currently an assistant professor at Anhui University, China. His research interests include computer vision, machine learning, and bioimage informatics. He is a member of three-dimensional imaging technology committee of the China Society of Image and Graphics.

Fuhui Long (fuhuil@alleninstitute.org) is currently with the Allen Institute for Brain Science in Seattle, Washingon, United States. She previously worked with Howard Hughes Medical Institute, Lawrence Berkeley National Lab, and Duke University. She is an associate editor of *BMC Bioinformatics*. Her research interests include big data and machine learning for neuroscience, brain research, health care, and education.

Hanchuan Peng (hanchuanp@alleninstitute.org) is the head of a research group at the Allen Institute for Brain Science in Seattle, Washingon, United States. His current research focuses on bioimage analysis and large-scale informatics as well as computational biology. His recent work includes developing novel and very efficient algorithms for three-dimensional (3-D) and higher-dimensional image analysis and data mining; building single-neuron, whole-brain level 3-D digital atlases for model animals; and Vaa3D. He was also the inventor of the mRMR feature selection algorithm. He received the 2012 Cozzarelli Prize. He is the founder of the annual Bioimage Informatics Conferences. He is currently a section editor of *BMC Bioinformatics*.

REFERENCES

[1] F. Long, J. Zhou and H. Peng, "Visualization and analysis of 3D microscopic images," *PLoS Comput. Biol.,* vol. 8, no. 6, p. e1002519, 2012.

[2] H. Peng, "Bioimage informatics: A new area of engineering biology," *Bioinformatics*, vol. 24, no. 17, pp. 1827–1836, 2008.

[3] O. Ronneberger, K. Liu, M. Rath, D. Rueβ, T. Mueller, H. Skibbe, B. Drayer, T. Schmidt, A. Filippi and R. Nitschke, "ViBE-Z: A framework for 3D virtual colocalization analysis in zebrafish larval brains," *Nat. Methods,* vol. 9, no. 7, pp. 735–742, 2012.

[4] N. Olivier, M. A. Luengo-Oroz, L. Duloquin, E. Faure, T. Savy, I. Veilleux, X. Solinas, D. Débarre, P. Bourgine and A. Santos, "Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy," *Science,* vol. 329, no. 5994, pp. 967–971, 2010.

[5] E. Meijering, "Cell segmentation: 50 years down the road" [Life Sciences], *IEEE Signal Processing Mag.,* vol. 29, no. 5, pp. 140–145, 2012.

[6] C. C. Fowlkes, C. L. L. Hendriks, S. V. Keränen, G. H. Weber, O. Rübel, M.-Y. Huang, S. Chatoor, A. H. DePace, L. Simirenko and C. Henriquez, "A quantitative spatiotemporal atlas of gene expression in the Drosophila blastoderm," *Cell,* vol. 133, no. 2, pp. 364–374, 2008.

[7] E. S. Lein, M. J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard, A. F. Boe, M. S. Boguski, K. S. Brockway and E. J. Byrnes, "Genome-wide atlas of gene expression in the adult mouse brain," *Nature,* vol. 445, no. 7124, pp. 168–176, 2006.

[8] S. W. Oh, J. A. Harris, L. Ng, B. Winslow, N. Cain, S. Mihalas, Q. Wang, C. Lau, L. Kuan and A. M. Henry, "A mesoscale connectome of the mouse brain," *Nature,* vol. 508, no. 7495, pp. 207–214, 2014.

[9] A. S. Forouhar, M. Liebling, A. Hickerson, A. Nasiraei-Moghaddam, H.-J. Tsai, J. R. Hove, S. E. Fraser, M. E. Dickinson and M. Gharib, "The embryonic vertebrate heart tube is a dynamic suction pump," *Science,* vol. 312, no. 5774, pp. 751–753, 2006.

[10] A.-S. Chiang, C.-Y. Lin, C.-C. Chuang, H.-M. Chang, C.-H. Hsieh, C.-W. Yeh, C.-T. Shih, J.-J. Wu, G.-T. Wang and Y.-C. Chen, "Three-dimensional reconstruction of brain-wide wiring networks in Drosophila at single-cell resolution," *Curr. Biol.,* vol. 21, no. 1, pp. 1–11, 2011.

[11] A. Jenett, G. M. Rubin, T.-T. Ngo, D. Shepherd, C. Murphy, H. Dionne, B. D. Pfeiffer, A. Cavallaro, D. Hall and J. Jeter, "A GAL4-driver line resource for Drosophila neurobiology," *Cell Rep.*, vol. 2, no. 4, pp. 991–1001, 2012.

[12] C. Castro-González, M. A. Luengo-Oroz, L. Duloquin, T. Savy, B. Rizzi, S. Desnoulez, R. Doursat, Y. L. Kergosien, M. J. Ledesma-Carbayo and P. Bourgine, "A digital framework to build, visualize and analyze a gene expression atlas with cellular resolution in zebrafish early embryogenesis," *PLoS Comput. Biol.,* vol. 10, no. 6, p. e1003670, 2014.

[13] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1997.

[14] K. Rohr, M. Fornefett and H. S. Stiehl, "Spline-based elastic image registration: Integration of landmark errors and orientation attributes," *Comput. Vis. Image Understand.,* vol. 90, no. 2, pp. 153–168, 2003.

[15] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Trans. Med. Imaging*, vol. 21, no. 11, pp. 1421–1439, 2002.

[16] E. G. Learned-Miller, "Data driven image models through continuous joint alignment," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 28, no. 2, pp. 236–250, 2006.

[17] L. Zöllei, E. Learned-Miller, E. Grimson and W. Wells, "Efficient population registration of 3D data," in *Computer Vision for Biomedical Image Applications*. New York: Springer, 2005, pp. 291–301.

[18] G. S. Jefferis, C. J. Potter, A. M. Chan, E. C. Marin, T. Rohlfing, C. R. Maurer, Jr., and L. Luo, "Comprehensive maps of Drosophila higher olfactory centers: Spatially segregated fruit and pheromone representation," *Cell,* vol. 128, no. 6, pp. 1187–1203, 2007.

[19] T. Rohlfing and C. R. Maurer, Jr., "Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees," *IEEE Trans. Inform. Technol. Biomed.*, vol. 7, no. 1, pp. 16–25, 2003.

[20] H. Peng, P. Chung, F. Long, L. Qu, A. Jenett, A. M. Seeds, E. W. Myers and J. H. Simpson, "BrainAligner: 3D registration atlases of Drosophila brains," *Nat. Methods,* vol. 8, no. 6, pp. 493–498, 2011.

[21] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM,* vol. 24, no. 6, pp. 381–395, 1981.

[22] H.-H. Yu, T. Awasaki, M. D. Schroeder, F. Long, J. S. Yang, Y. He, P. Ding, J.-C. Kao, G. Y.-Y. Wu, and H. Peng, "Clonal development and organization of the adult *Drosophila* central brain," *Curr. Biol.,* vol. 23, no. 8, pp. 633–643, 2013.

[23] H. Peng, J. Tang, H. Xiao, A. Bria, J. Zhou, V. Butler, Z. Zhou, P. T. Gonzalez-Bellido, S. W. Oh, and J. Chen, "Virtual finger boosts three-dimensional imaging and microsurgery as well as terabyte volume image visualization and analysis," *Nat. Commun.,* vol. 5, no. 4342, 2014. doi: 10.1038/ncomms5342.

[24] L. Qu and H. Peng, "A principal skeleton algorithm for standardizing confocal images of fruit fly nervous systems," *Bioinformatics,* vol. 26, no. 8, pp. 1091–1097, 2010.

[25] J. W. Brandt and V. R. Algazi, "Continuous skeleton computation by Voronoi diagram," *CVGIP: Image Understand.,* vol. 55, no. 3, pp. 329–338, 1992.

[26] J.-H. Chuang, N. Ahuja, C.-C. Lin, C.-H. Tsai, and C.-H. Chen, "A potential-based generalized cylinder representation," *Comput. Graph.,* vol. 28, no. 6, pp. 907–918, 2004.

[27] L. Lam, S.-W. Lee and C. Y. Suen, "Thinning methodologies-a comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 9, pp. 869–885, 1992.

[28] G. Malandain and S. Fernández-Vidal, "Euclidean skeletons," *Image Vis. Comput.,* vol. 16, no. 5, pp. 317–327, 1998.

[29] C. Wählby, L. Kamentsky, Z. H. Liu, T. Riklin-Raviv, A. L. Conery, E. J. O'Rourke, K. L. Sokolnicki, O. Visvikis, V. Ljosa and J. E. Irazoqui, "An image analysis toolbox for high-throughput C. elegans assays," *Nat. Methods,* vol. 9, no. 7, pp. 714–716, 2012.

[30] O. Ishaq, J. Negri, M.-A. Bray, A. Pacureanu, R. T. Peterson and C. Wahlby, "Automatic quantification of Zebrafish tail deformation for high-throughput drug screening," in *IEEE 10th Int.Symp. Biomedical Imaging (ISBI)*, 7–11 Apr. 2013, San Francisco, CA, pp. 902–905.

[31] H. Peng, F. Long, X. Liu, S. K. Kim and E. W. Myers, "Straightening *Caenorhabditis* elegans images," *Bioinformatics,* vol. 24, no. 2, pp. 234–242, 2008.

[32] F. Long, H. Peng, X. Liu, S. K. Kim and E. Myers, "A 3D digital atlas of C. elegans and its application to single-cell analyses," *Nat. Methods,* vol. 6, no. 9, pp. 667–672, 2009.

[33] A. White, B. Lees, H.-L. Kao, G. Cipriani, A. Paaby, E. Sontag, K. Erickson, D. Geiger, K. Gunsalus and F. Piano, "DevStaR: High-throughput quantification of C. *elegans* developmental stages," *IEEE Trans. Med. Imaging* vol. 32, no. 10, pp. 1791–1803, 2013.

[34] A. Cardona, S. Saalfeld, I. Arganda, W. Pereanu, J. Schindelin and V. Hartenstein, "Identifying neuronal lineages of *Drosophila* by sequence analysis of axon tracts," *J. Neurosci*., vol. 30, no. 22, pp. 7538–7553, 2010.

[35] L. R. Girard, T. J. Fiedler, T. W. Harris, F. Carvalho, I. Antoshechkin, M. Han, P. W. Sternberg, L. D. Stein and M. Chalfie, "WormBook: The online review of *Caenorhabditis elegans* biology," *Nucleic Acids Res.,* vol. 35, no. 1, pp. D472–D475, 2007.

[36] F. Long, H. Peng, X. Liu, S. Kim and G. Myers, "Automatic recognition of cells (ARC) for 3D images of *C. elegans*," in *Research in Computational Molecular Biology*, 2008, pp. 128–139.

[37] D. Kainmueller, F. Jug, C. Rother and G. Myers, "Active graph matching for automatic joint segmentation and annotation of *C. elegans*," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 81–88, 2014.

[38] L. Qu, F. Long, X. Liu, S. Kim, E. Myers and H. Peng, "Simultaneous recognition and segmentation of cells: Application in *C. elegans*," *Bioinformatics,* vol. 27, no. 20, pp. 2895–2902, 2011.

[39] A. Rangarajan, H. Chui and F. L. Bookstein, "The soft assign procrustes matching algorithm," in *Proc. Information Processing in Medical Imaging*, 1997, Poultney, VT, pp. 29–42.

[40] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA: SIAM, 1990.

[41] S. Gold, A. Rangarajan, C.-P. Lu, S. Pappu and E. Mjolsness, "New algorithms for 2D and 3D point matching: Pose estimation and correspondence," *Pattern Recognit.,* vol. 31, no. 8, pp. 1019–1031, 1998.

[42] Q. Wu, F. Merchant and K. Castleman, *Microscope Image Processing*. New York: Academic Press, 2010.

[43] Y. Al-Kofahi, W. Lassoued, W. Lee and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Trans. Biomed. Eng.* , vol. 57, no. 4, pp. 841–852, 2010.

[44] O. Dzyubachyk, W. A. van Cappellen, J. Essers, W. J. Niessen and E. Meijering, "Advanced level-set-based cell tracking in time-lapse fluorescence microscopy," *IEEE Trans. Med. Imaging,* vol. 29, no. 3, pp. 852–867, 2010.

[45] B. Roysam, W. Shain, E. Robey, Y. Chen, A. Narayanaswamy, C. Tsai, Y. Al-Kofahi, C. Bjornsson, E. Ladi and P. Herzmark, "The FARSIGHT project: Associative 4D/5D image analysis methods for quantifying complex and dynamic biological microenvironments," *Microsc. Microanal.*, vol. 14, no. S2, pp. 60–61, 2008.

[46] H. Peng, A. Bria, Z. Zhou, G. Iannello, and F. Long, "Extensible visualization and analysis for multidimensional images using Vaa3D," *Nat. Protocols*, vol. 9, no. 1, pp. 193–208, 2014.

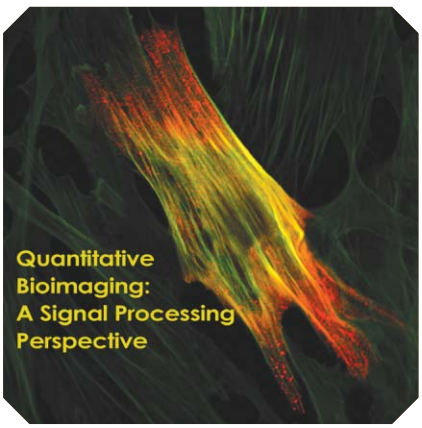
[47] S. Saalfeld, A. Cardona, V. Hartenstein and P. Tomancak, "CATMAID: Collaborative annotation toolkit for massive amounts of image data," *Bioinformatics*, vol. 25, no. 15, pp. 1984–1986, 2009.

[SP]

Michael T. McCann, John A. Ozolek, Carlos A. Castro, Bahram Parvin, and Jelena Kovačević

Automated Histology Analysis

Opportunities for signal processing



Histology is the microscopic inspection of plant or animal tissue. It is a critical component in diagnostic medicine and a tool for studying the pathogenesis and biology of processes such as cancer and embryogenesis. Tissue processing for histology has become increasingly automated, drastically increasing the speed at which histology labs can produce tissue slides for viewing. Another trend is the digitization of these slides, allowing them to be viewed on a computer rather than through a microscope. Despite these changes, much of the routine analysis of tissue sections remains a painstaking, manual task that can only be completed by highly trained pathologists at a high cost per hour. There is, therefore, a niche for image analysis methods that can automate some aspects of this analysis. These methods could also automate tasks that are prohibitively time-consuming for humans, e.g., discovering new disease markers from hundreds of whole-slide images (WSIs) or precisely quantifying tissues within a tumor.

In this article, we aim to acquaint the signal processing researcher with histology and review the current approaches to the fascinating and important signal processing problems associated with histology image analysis. Throughout, we focus on slides stained with the ubiquitous hematoxylin and eosin (H&E) stain and imaged with brightfield microscopy.

Digital Object Identifier 10.1109/MSP.2014.2346443
Date of publication: 5 December 2014

HISTOLOGY: THE PATHOLOGIST'S VIEW

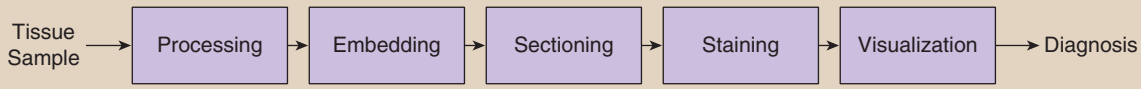
The main goal of the surgical pathologist in a diagnostic practice is to examine tissue and render a correct diagnosis that will ultimately translate to a therapeutic intervention for the patient. The therapeutic response may range from no action, in the case of a diagnosis of normal or unremarkable, to close follow-up, local excision, medical treatment only (benign diagnoses), or radical chemotherapy and/or surgery (malignant diagnoses).

Figure 1 presents the pipeline from tissue processing to diagnosis from the perspective of the diagnostic pathologist. Understanding this process will give the reader an appreciation of how images are derived from tissue and the associated sources of variability, noise, and artifacts; this information is critical to designing automated image analysis systems. This portion of the article is based on standard histology texts [1], [2], as well as years of experience in the field of pathology.

TISSUE COLLECTION

The clinical histology process begins when the treating physician, after assessing the patient by history, physical examination, and/or radiographic and laboratory studies, determines that treatment can proceed no further without histology confirmation. The treating physician then must obtain enough good-quality tissue to obtain a diagnosis. There are several possible approaches to tissue collection, including fine-needle aspiration, needle biopsy, excisional biopsy, or excision of the lesion in its entirety. The sensitivity (likelihood of

1053-5888/15©2015IEEE



[FIG1] A block diagram of the histology process.

getting the correct diagnosis) and specificity (likelihood of not getting the incorrect diagnosis) increase from fine-needle aspiration to excision of the entire lesion. This is because the larger biopsies preserve more cellular context and allow the pathologist to examine multiple slides from different areas of the sample.

After biopsy, the pathologist evaluates the tissue on the macroscopic scale, measuring it and recording a description of its color and characteristics. For larger tissues (e.g., tumor resections or colon resections), the tissue must be trimmed to fit into the tissue cassettes (approximately 10 × 10 × 3 mm) that will contain it for the subsequent processing steps (Figure 2).

PROCESSING

The next step in the diagnostic pipeline is tissue processing, which involves chemically and physically stabilizing the tissue. The tissue is first immersed into a fixative solution that is used to stop cells from breaking down and prevent microorganism growth. In general, tissue is fixed for a few hours (small biopsies) to about 24 hours (large biopsies). Fixation is critical because poorly fixed tissue leads to poor tissue sectioning and poor microscopic morphology.

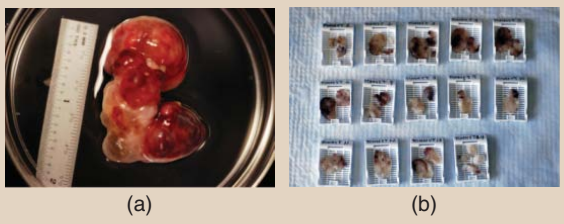
After fixation, the tissue is physically stabilized by one of several methods (freeze drying, microwave, chemical) with the end goal of preserving the cellular morphology. The most commonly employed method involves the use of alcohols and xylene and is automated in most laboratories: First, the tissue is dehydrated, which clears the water and aqueous fixative from it. Next, the tissue is cleared of the dehydrating agent, leaving the tissue ready for paraffin infiltration. Finally, the paraffin warms in the processor until it is liquefied, infiltrates the tissue under vacuum, and then cools so that the tissue becomes firm. This process takes approximately nine hours and, in many laboratories, is run overnight. One of the end results of processing tissue in this manner is that the tissue section is slightly smaller than the original fresh or fixed tissue prior to processing.

EMBEDDING

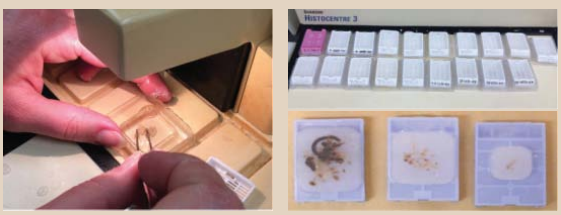
After processing, the tissue is embedded in a block of support material, as shown in Figure 3. To achieve this, the tissue is placed on the bottom of a mold and paraffin is poured over it. The original tissue cassette is placed over the mold and then onto a cooling plate to solidify the paraffin. The result is called a *tissue block*, tissue that is impregnated with and surrounded by hardened paraffin. Because the tissue will be sliced parallel to the cassette, orientation of the tissue during embedding is key; see Figure 4 for an example.

SECTIONING

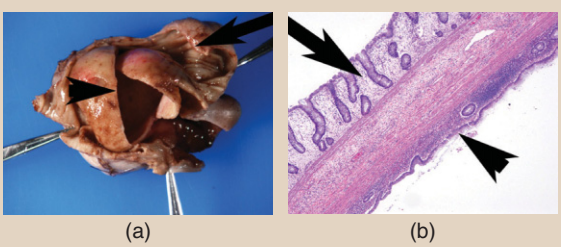
Sectioning is cutting thin slices of tissue that are mounted on microscope slides (Figure 5). This is achieved with a tool called a *microtome*, which operates like a deli slicer. This step can be manual, semiautomated, or automated. For example, the manual rotary microtome advances the block by a set amount with each turn of the wheel and thus produces a ribbon of tissue. Most tissue sections for diagnostic purposes are cut at 3–4-μm thickness. For certain applications (e.g., silver staining of kidney biopsies) thinner sections are necessary; these are more difficult to obtain without damaging the tissue. Thicker sections tend to make staining dark and obscure nuclear detail.



[FIG2] (a) A gross specimen is cut into smaller pieces and placed into (b) cassettes for further processing.



[FIG3] (a) Paraffinized tissue is oriented and embedded into a block of paraffin. (b) Tissue blocks after embedding.



[FIG4] The importance of proper orientation during embedding. (a) A gross specimen shows an opened enteric duplication cyst (arrowhead) intimately joined to the bowl (arrow). (b) A histology image of the cyst with normal bowel (arrow) abutting cyst wall (arrowhead) can only be observed when slicing occurs perpendicular to the wall.

(a) (b)

[FIG5] (a) During sectioning, the tissue block is sliced into 3–4-μm sections, which remain connected in a ribbon. (b) Collecting a ribbon from the water bath using a glass slide.

(a) (b)

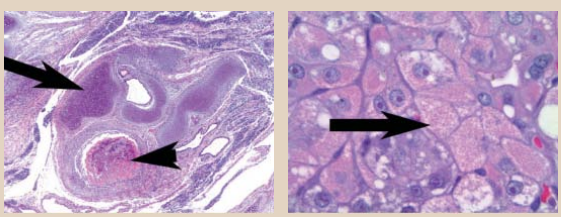
[FIG6] (a) Histology slides after staining. (b) The visualization setup including a microscope with attached camera and a desktop computer.

These thinly cut sections are floated out onto a water bath heated to about 10 °C below the melting point of the paraffin. This allows dispersion of any wrinkles generated by the microtome blade at sectioning. Once the sections are cut and floated, they are placed onto 25 × 75 × 1 mm glass microscope slides.

STAINING

At this point, the tissue slices are nearly invisible under a light microscope, so they must be stained to create contrast. Most staining procedures in the laboratory, aside from immunohistochemical (IHC) ones (antibody based), use chemicals or dyes that will bind or have affinity for certain components of the cells and extracellular components. The chemical properties of these dyes produce the visual appearance that is seen under the microscope.

The most widely used stains for both diagnostic and research histology are H&E. Hematoxylin stains nucleic acids and appears blue/purple, while eosin stains proteins and appears pink/red when visualized under a brightfield microscope. So for most tissues, cell nuclei are blue, while cytoplasm can vary from clear to red to purple depending on its constituents. The reason that H&E staining of tissue has persisted for decades as the primary tissue stain in diagnostic and research pathology is that these stains attach themselves to almost every cellular component, allowing for the visualization of whole cells and all tissue components. Another reason is that these stains provide excellent contrast between cellular constituents by having chemical properties that produce colors at opposite ends of the visual spectrum. These color perceptions are helpful in diagnosis, though not entirely necessary, since, even in grayscale, distinctions between and within tissues can be made and diagnoses rendered.



(a) (b)

[FIG7] Some examples of important visual cues in histology. (a) In this teratoma section, color makes distinguishing cartilage (blue/gray, arrow) from bone (pink/red, arrowhead) easy, even at low magnification. (b) In this liver section of a child with mitochondrial disorder, texture is important. Under high magnification, cell borders are accentuated and cytoplasm shows tiny red granules representing abnormal mitochondria (arrow).

VISUALIZATION

Once stained, the slides must be visualized (Figure 6). A growing trend in pathology is to digitize slides so that pathologists can make diagnoses based solely on the digital image [3], [4]. The advent and refinement of whole-slide scanners have made rapid scanning and high-resolution WSIs commonplace. These systems are now offered by many companies and offer spatial resolutions using the 40 × objective of approximately 0.23–0.25 μm/pixel. Storing digital images would be an attractive alternative to storing glass slides, since glass slides take considerably more space, can be damaged or lost, and fade over time, but currently institutions must keep their glass slides and tissue blocks for at least ten years.

Despite these advantages, in most medical centers, the pathologists still rely on visualization of the slide through a microscope. According to pathologists, a microscope offers faster panning, faster focusing, and an intangible sense of being closer to the tissue. This last advantage is more difficult to explain, but the sense is that the eye can capture greater detail through the microscope than from a digital image on a monitor. Even so, studies have shown no appreciable difference in diagnoses rendered by pathologists using digitized images compared to diagnoses rendered using a microscope [5]–[8].

ANALYSIS

The goals of the pathologist in the clinical domain versus the research domain can be very different; we discuss each separately.

CLINICAL PRACTICE

In the clinic, the goal of the pathologist is to render accurate and timely diagnoses. For a given slide, they analyze a wide variety of characteristics including tissue architecture, cellular color and texture (Figure 7), and cellular/nuclear morphometry (Figure 8) to produce diagnoses. Computer algorithms may automate some of these tasks. For example, in cytopathology, automated methods for screening Pap smear slides limit the number of cytotechnologists needed to complete these screenings. In the future, we could

hope to automate or semiautomate histology screenings, e.g., of the gastrointestinal tract. Such automation could greatly reduce health-care costs and potentially provide pathologists more time for challenging cases and important research.

RESEARCH

In the research domain, pathologists may aim to quantify differences between histology samples in terms of a variety of parameters including cellular/nuclear morphometry, amount of stroma (the connective tissue cells that support the function of cells around them) present, types of tissue present, etc. Although some basic analyses can be readily performed using available image analysis software (e.g., Photoshop, ImageJ, MetaMorph), most quantification in pathology remains semiquantitative: staining intensity may be rated as low, moderate, or strong; the amount of a certain cell type may be visually estimated as 0–25%, 25–50%, or >50% of the total population; and morphometric descriptions of cells are limited to semantic descriptions such as larger, thickened, pleomorphic, or cellular.

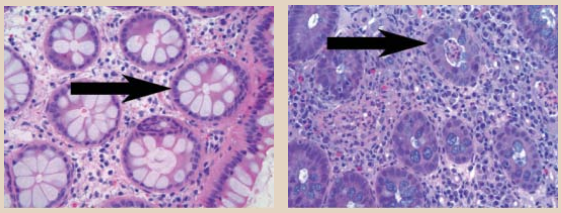
For some studies, this type of analysis is not accurate enough, because many biologically or clinically relevant features cannot be easily captured and processed by the human visual system. For example, given two tumors, how would a pathologist support the claim that the average nucleus size is different between them? Similarly, how can pathologists quantify, by eye, complex patterns such as chromatin distribution? Signal processing solutions to these quantification challenges would be extremely useful technology for researchers in almost any area of investigation that analyzes and quantifies observations from tissue specimens.

SOURCES OF VARIABILITY

There are three main sources of variability in a histology-based diagnosis: biological variability, interobserver variability, and technical variability. Biological variability encompasses the normal variability among people and the myriad of pathological processes that can affect any tissue group. Due to biological variability, slides generated from the same tissue in different patients can look different. Interobserver variability contends that two pathologists can look at the same tissue and render different interpretations. Finally, technical variability is the variability in a slide's appearance due to how it was prepared. We focus here on the details of technical variability.

In the best case, the slide that is generated from the above tissue processing pipeline shows tissue that is properly oriented, sectioned, stained, and coverslipped. Unfortunately, each of these steps can introduce variability into the final product, and differences in protocol between labs can greatly alter the appearance of even biologically similar tissue samples. Some variables such as fixation, specimen orientation in the block, and microtome sectioning are heavily dependent on human skill, and even though tissue processing, staining, and coverslipping are largely automated, they still depend on human monitoring, machine maintenance, and solution preparation.

For automated analysis systems, each source of variability presents a significant obstacle. Tissue that is poorly fixed will not cut well and has a blurry appearance, removing important cues such as edges. Tissue that has been dried out will be shrunken and have poor morphology and stain contrast, similar to poorly fixed tissue. Sectioning artifacts are some of the most commonly encountered artifacts and can produce folds in the tissue, chatter artifacts from a dull blade (seen as alternating light and dark regions), or missing pieces of the tissue.

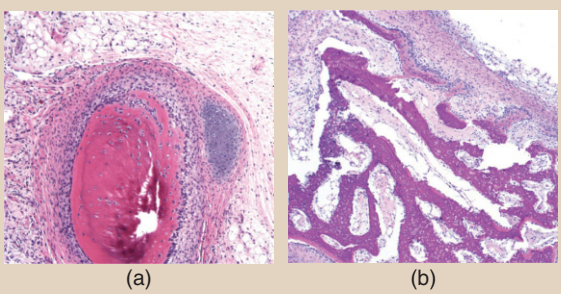


[FIG8] Some examples of the importance of nuclear shape and distribution in colon screening. (a) In normal colon tissue, nuclei are regular in size and distribution around the colon gland (arrow). (b) In colitis, nuclei become much less regular (arrow).

Staining is a critical source of variability because it produces the color and contrast on the slide (Figure 9). Many different formulations of H&E exist, each producing a slightly different appearance. The stain can also be applied in different manners. In progressive staining, the sample remains in the hematoxylin solution for a specified amount of time to render appropriate staining. In regressive staining, the sample remains in the hematoxylin long enough to overstain, and is then destained back to the desired contrast with an acid alcohol solution. Furthermore, both H&E solutions can have their staining capabilities altered by prolonged storage, contamination by other reagents or water, precipitation, and changes in pH. Staining artifacts can include light staining with either hematoxylin or eosin, precipitated hematoxylin (seen as blue chunks under the microscope), or lack of staining with either hematoxylin, eosin, or both. That these artifacts are well understood by pathologists is another reason why H&E remains so popular.

AUTOMATED HISTOLOGY IMAGE ANALYSIS

In this section, we survey the current methods and signal processing challenges in automated histology image analysis. We aim to



[FIG9] An example of staining variability. Both (a) and (b) show images made of bone, but the color of the bone varies from (a) pink to (b) purple due to staining variability. Such color variations present a challenge for automated analysis systems.

cover a broad range of signal processing topics, but we restrict the review to slides stained with the common and inexpensive H&E stain and imaged with brightfield microscopy. This is because H&E images are prevalent in clinical and research settings and because this focus allows us to survey a more coherent group of methods; we argue that the approaches to a single task, e.g., nucleus detection, vary considerably across staining and imaging modalities. This section is organized around a generalized block diagram of a hypothetical histology analysis workflow, shown in Figure 10. For each block in the diagram, we describe the signal processing challenge and review the current approaches in the literature.

We are aware of three previous reviews of automated histology analysis. The first, [9], deals with analysis of histology images including H&E images as well as fluorescence and multispectral images. It covers preprocessing; segmentation of glands, nuclei, and other subcellular components; feature extraction; dimensionality reduction; and classification. The second, [10], tackles the broader field that the authors term *computational pathology*, which includes histology as well as cytology analysis. It discusses issues of data and ground truth collection including variation among experts and publicly available data sets and describes automated analysis primarily from a statistical pattern recognition viewpoint. The third, [4], discusses histology WSI informatics, including quality control during image acquisition, feature extraction, region of interest (ROI) detection, and visualization.

While our survey does not include every paper on automated histology, we aim to give the reader a sense of what has been tried for the various histology analysis tasks we present. We have given priority to recent journal papers except where necessary. Also, note that we do not focus on results of or comparison between methods. This is because, at this stage, automated histology research is diffuse: most methods are tailored to private data sets and there is no consensus on what quantitative metrics should be reported. Moving toward shared data sets and metrics will be a critical step forward in the field.

PREPROCESSING

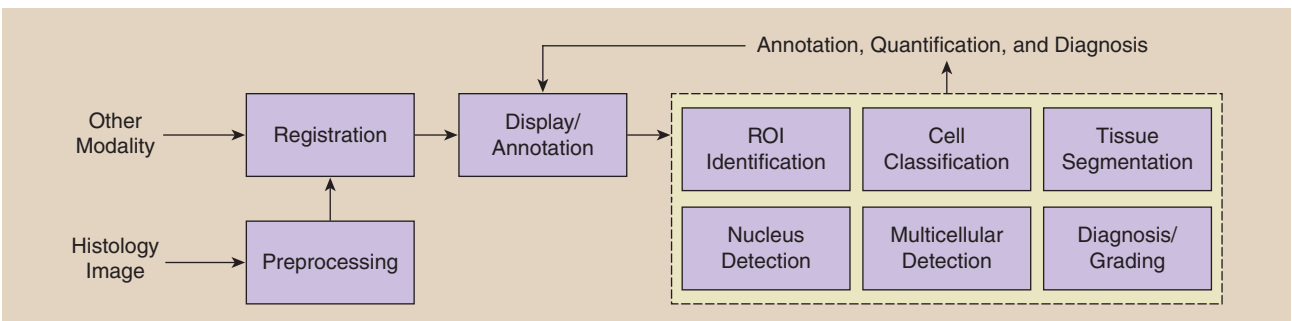
Histology images exhibit the same types of artifacts and noise as any digital microscopy image, as well as some novel ones introduced by tissue processing. These can affect analysis unless removed with appropriate preprocessing. A good overview of noise and artifacts in digital microscopy and methods to correct them are discussed in [11]; in this section we focus on the issues specific to H&E images of tissue processed as described above.

STAIN NORMALIZATION

As discussed in the section "Sources of Variability," staining variations affect the appearance of histology images; such variations are problematic for automated analysis because color is a critical feature in histology. Stain normalization is the process of taking two H&E images that have staining variation between them and removing this variation. Doing so has been shown to improve histology image segmentation [12], [13]. One approach to stain normalization is to use color normalization techniques from photography such as histogram equalization; e.g., in [12], the rank statistics of the input image are scaled to match those of the reference image separately in each color channel. The method has been successfully used in other automated histology work, including [14]. Such methods are especially suited for cases where the images to be normalized show approximately the same tissue, e.g., in the case of serial sections such as in [15].

An approach more specifically tailored for H&E image normalization is to first separate the image into H-only and E-only images (sometimes called *color deconvolution*), then normalize these images separately and recombine. This approach better handles cases where, e.g., the hematoxylin stain is too intense but the eosin stain is too weak. When the stain colors are known, the color deconvolution method in [16] solves the separation. When they are not known, the problem becomes more challenging. The approach taken in [13], [17], and [18] is to estimate the stain colors and deconvolve as before. In [17], the stain colors are estimated using the fact that all stained pixels will lie on a wedge in color space; this is a special case of the nonnegative matrix factorization problem studied elsewhere. Reference [18] uses expectation-maximization to find clusters in chromaticity space that correspond to stain colors, and [13] finds pixels stained with only one stain via supervised classification and uses their mean color to estimate the stain colors. A different approach is to estimate the single-stain images directly; e.g., [19] searches for an H-only image that removes most of the contrast from the red channel of the input, based on the assumption that most this contrast comes from the nuclei, which are stained only by hematoxylin.

These separation methods all produce qualitatively fair results, but it is unclear which comes closest to the correct separation.



[FIG10] A block diagram of a generic automated histology analysis workflow. Most current works address only one or a few of these blocks.

We have recently released a stain separation benchmark data set based on chemical destaining [20]; our comparison indicates that the method of [17] is superior to that of [19], but we did not evaluate any of the more recent approaches.

TISSUE DEFORMATION

Because the tissue slices are very thin, they can fold over on themselves during processing, creating a tissue area that is doubly thick. The authors of [21] proposed detecting these folds by their high color saturation, and the authors of [22] devised a method for selecting a good saturation threshold for this detection, but, to our knowledge, no work addresses correcting tissue folds digitally. In the same vein, differences in the water content of different tissues can cause them to pull apart when the tissue is dehydrated during processing; the result is white cracks that are not biologically meaningful. Again, we are not aware of any work that addresses correcting these artifacts.

STITCHING

High-resolution WSIs are often acquired by imaging several strips or tiles separately, which then must be stitched together to create the final image. Because the offsets between the strips or tiles are known, adequate stitching results are often provided by the microscope/scanner software or can easily be achieved with simple compositing techniques. A more challenging stitching problem occurs when a large gross specimen must be sectioned and imaged as several pieces. In [23], the authors describe a GUI that allows stitching of high-resolution images of such tissue fragments via hand-selected control points and linear transformation. We are not aware of a work that addresses histology image stitching with automatic control point selection or explores a richer set of transforms.

REGISTRATION

There are a variety of H&E histology image registration tasks, depending on the image modalities involved. In this section, we review three common registration tasks involving H&E images.

H&E TO H&E

Given serial sections of a tissue sample, each stained with H&E, one aim is to register them to create a tissue volume. This three-dimensional (3-D) reconstruction can more fully show the extent of a pathological process or show relationships of tissue types to one another. This is a difficult registration task: each slice undergoes nonrigid deformation during processing and may exhibit cracking or folding artifacts. The standard approach is to register in a coarse to fine manner as in [15], where the authors register images of serial tissue slices to explore the 3-D shape of cervical tumor fronts using a series of three registrations: 1) a rigid registration using a frequency domain method, 2) a polynomial registration using control points automatically selected with correlation matching on small patches, and 3) a registration consisting of unconstrained local displacements regularized by local curvature. An approach more tailored to histology is to use specific anatomical landmarks; e.g., [24] registers based on blood vessels.

H&E TO ANOTHER STAIN

One may also want to register images of two adjacent tissue slices that use different stains; e.g., H&E to IHC stains. This task is difficult because, by design, the stains will give contrast to different structures. One way to overcome this obstacle is with the selection of an appropriate pixel-wise similarity measure, e.g., the authors of [25] register H&E images to a variety of IHC images using mutual information and the elastix software package [26]. Another approach is given in [27], where distinctive landmarks such as blood vessels are segmented and used to compute the registration.

H&E TO MRI

Compared to histology, magnetic resonance imaging (MRI) has low contrast and resolution, but has the advantage of being noninvasive; registering histology images to magnetic resonance (MR) images could help train radiologists, provide better noninvasive diagnoses, and enable the development of MRI-based computer-aided design tools [28]. This task is difficult because MR images are 3-D while histology images are two dimensional, MR images are lower resolution than histology images, and because the contrast in MR and histology images is generated in different ways. When considering MR images collected in vivo, registration is even more difficult because surgical extraction and histology processing can greatly deform the tissue. In [29], histology images are registered to in vivo MR images via two intermediates, the block face photo and the ex vivo MR image, with the idea that the deformations between these intermediate stages are less drastic and therefore easier to estimate. Each step of the registration is completed using mutual information as the metric and thin plate splines to deform the image with control points initialized by hand. In [28], histology images are registered to MR slices to create a histology volume, then the MR volume is registered in 3-D to the histology volume and resliced. The process is iterated until convergence.

In each of these scenarios, comparison between methods is difficult because no ground truth exists. This problem is intensified in the first two scenarios because qualitatively good registrations may actually remove true differences between the adjacent tissue slices being registered.

DISPLAY AND ANNOTATION

A fundamental problem in the display of histology images is that they can be huge, easily several gigabytes for uncompressed WSIs. As a result, commercial slide scanners often save images in a proprietary format for which the manufacturers provides free viewing software, e.g., Aperio ImageScope. For the researcher interested in reading the images themselves, one solution is Openslide [30], a C library that aims to allow slides from any vendor to be opened, manipulated, or converted to other formats. One format that OpenSlide can convert to and is a natural fit for large histology images is the Deep Zoom (DZI) format, which creates from a large image a tiled image pyramid, allowing real-time viewing of images of arbitrary size, even streaming over the Internet. One example of this approach can be found in [31], which involves displaying WSIs from the Cancer Genome Atlas data set [32] on the Internet using the similar Zoomify format. The system in [33] is designed

specifically for viewing histology images and avoids the time-consuming calculation of image pyramids by creating the current view in real time from image tiles.

Beyond simply viewing large histology images, researchers may want to collect expert annotations for them for the purpose of training and testing their analysis algorithms. In the simplest case, the expert pathologist may label an entire image, e.g., as *normal* or *cancer*. For more fine-grained labeling, Aperio ImageScope, Sedeen Viewer, and Cytomine [34] provide freehand annotation tools for large images. The next step for histology image display and annotation is systems allowing the pathologist to interact with automated analyses, teaching and correcting the system in real time. Some efforts have been made in this area: after the automated segmentation in [35], experts can click on nuclei to indicate that they either need to be split or merged, and Cytomine is beginning to include collaborative proofreading of automated cancer segmentation.

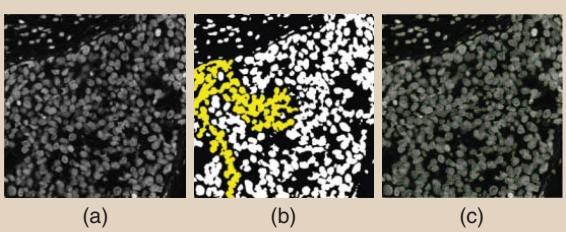
ROI IDENTIFICATION

A single biopsy can generate dozens of high-resolution WSIs, however, often, only a small region of this vast quantity of tissue is diagnostically useful. There is, therefore, a need for fast computational methods that can identify these ROIs in while-slide histology images. Once identified, these regions can be passed to a pathologist or to subsequent steps of a larger automated histology pipeline.

One approach is to simply downsample the input; e.g., [36] uses a low-resolution input to extract features based on color and sparse coding of subpatches. These features are classified via a support vector machine (SVM) to detect ROIs. Another approach is to process the image at multiple scales, such as in [37], where WSIs of breast cancer tissue are recursively partitioned via color clustering at increasingly fine resolutions to efficiently identify tissue versus nontissue and lesion versus normal regions. Yet another approach is to detect objects such as glands and identify ROIs based on these [38]. There are generally fewer objects on a slide than pixels, so processing objects can be more efficient than processing pixels.

NUCLEUS DETECTION

Nuclei are prevalent in histology images and their size, shape, distribution, and texture are relevant for many analysis tasks such as identifying inflammation, identifying and grading cancer, and determining tissue type. Therefore, locating nuclei is a critical step in many histology analysis systems. Nucleus detection in histology images is challenging because nuclei can be tightly clustered (Figure 11) and vary in size, shape, and color depending on their cell type. Even within a single cell type, the stain may not penetrate all nuclei equally, leaving some darkly stained and others lightly stained. Finally, human nuclei are around the same scale as the thickness of the tissue slice (5 μm), meaning that, for some nuclei, only a portion appears on the slide.



[FIG11] The clustering of nuclei makes nucleus segmentation challenging. (a) In this grayscale, H-only image, the closely packed and overlapping nuclei cannot be separated with threshold-based clustering, resulting in (b) large connected components (yellow cluster). The method discussed in [35] breaks apart these clusters via graph cuts, resulting in (c) a good separation of nuclei. (Figure adapted with permission from [35].)

Basic approaches to nucleus detection involve color clustering of the pixels; e.g., the work in [39] uses clustering in the lab color space to identify four subcellular components—nuclei, cytoplasm, neuropil, and background—and [40] uses expectation-maximization clustering to identify regions of lymphocyte nuclei, stroma, cancer nuclei, and background.

Using only local information omits the strong prior knowledge we have about nucleus size and shape. Approaches that move beyond local information include [36] and [40], which use active contours to refine their initial color segmentation, and [14], which uses a graph cut based on color and Laplacian of Gaussian features. Similarly, the authors in [35] use Laplacian of Gaussian filtering with clever scale selection to detect nucleus seed points followed by local maximum clustering to form a rough segmentation. Another approach is to frame nucleus detection as a classification problem; e.g., in [41], an SVM is trained to detect rectangular windows containing nuclei; the features used are pixel intensities and Laplacian of Gaussian edge intensities.

Finally, several methods use heuristics to separate clustered nuclei, including the curvature-based reasoning in [14] and concavity detection in [40]. The method in [35] separates nuclei via graph cuts (Figure 11). These approaches achieve impressive-looking results even when nuclei are clustered, but large, hand-annotated data sets will be necessarily to compare them and understand their strengths and weaknesses.

CELL CLASSIFICATION

Most histology images contain cells of several types; moving beyond nucleus detection to cell classification can provide valuable diagnostic information. For example, the method in [42] classifies cells as centroblasts or normal, which is useful for cancer grading. It uses color and Fourier-based texture features with quadratic discriminant analysis as the classifier. The nuclei found in [40] are classified as belonging to lymphocytes or other based on their color.

Other work focuses on detecting mitotic cells. Notably, [43] presents a contest data set for this task which comprises 50 images collected on each of two different slide scanners as well as a multispectral scanner. The highest-scoring approach at the time of the contest was based on a deep convolutional neural network [44].

MULTICELLULAR STRUCTURE DETECTION

Cells in histology images are not solitary, rather they are part of organized structures (e.g., glands, acini). Detecting and analyzing these structures is a unique challenge in histology image analysis.

Because there are many types of multicellular structures, these tasks tend to be more varied and application-driven than the ones we have discussed before, hence we give only a few examples here.

Glands are a multicellular structure common to many tissue types (e.g., salivary, breast, prostate, pancreatic, sinonasal, gastrointestinal tissues), and changes in their morphology can be an important indicator of disease. In histology images, most glands appear as clear areas surrounded by cells. The method in [38] leverages the clear areas to find seed pixels and then uses region growing to segment the glands. Going further, [38] classifies glands as malignant or normal based on their size with a Markov random field to impose spatial smoothness. In [45], a graph is built on top of a colon tissue image with nodes corresponding to either nucleus or nonnucleus objects. The colon glands can then be described by subgraphs around a user-selected point at their center. Matching these subgraphs to reference ones from healthy or diseased glands allows classification of the input image.

Some approaches to segmenting multicellular structures treat them like nuclei, e.g., [46] segments lymphoid follicles, which are organized groups of lymphoid cells, using active contours. Detected regions are split based on curvature and false detections are further trimmed based on color.

TISSUE SEGMENTATION

Tissues are organized groups of cells. Identifying them in an image is important for diagnosis or giving context to subsequent analyses. The wide variety of tissues and the complexity of their appearances makes this a challenging problem. The method in [47] uses local pixel intensities as features and is able to segment bone, cartilage, and fat tissue in teratoma tumor images. We presented a segmentation method [48] inspired by the lack of edges in histology images. It uses local color histograms rather than edge-based features and outperforms generic methods for tissue segmentation. The authors in [49] showed that segmentations seeking homogeneity of objects such as cells and crypts, rather than simply pixel homogeneity, perform well on colon tissue.

Some tissue-level analyses amount to diagnoses; e.g., in [50], small subregions of a prostate tissue image are classified as *normal, stroma*, or *prostatic adenocarcinoma*. Normal regions are those around glands, which are simple to segment because of their white centers. Stroma and cancer regions are distinguished using Haralick texture features. Another group [51] used a graph-based methodology to segment regions of cancer in colon images, where features were based on the frequency of co-occurrence of nodes.

DIAGNOSIS

Automated diagnosis is essentially image classification: given a histology image, what disease does it represent? As such, most approaches follow the paradigm of feature extraction followed by classification. The authors in [52] address the problem of classifying subtypes of renal tumor in expert-selected ROIs. They use Fourier shape descriptors extracted from binary masks of nuclei, cytoplasm, and unstained regions as features and a series of SVM classifiers arranged in a directed acyclic graph to distinguish between three types of renal cell carcinoma and one benign tumor. Based on which shape descriptors are most distinguishing during classification, the authors can identify which shapes are indicative of each tissue type. In [53], hand-selected ROIs of breast tissue are classified as *normal, in situ cancer,* or *invasive cancer*. They use generic features including local binary patterns, co-occurence matrix statistics, and curvelet coefficient statistics. Classification proceeds in two stages, first a random subspace ensemble of SVMs, then a random subspace ensemble of neural networks. Each stage may either classify an image or reject it as too difficult. Images rejected from the first stage move on to the second, while images rejected from the second stage are viewed by a human expert. The method achieves accuracy of over 99% with a 1.94% rejection rate. The recent conference paper [54] moves away from hand-designed features by learning features from the data using sparse representation.

Grading is diagnosing the severity of a disease. Established grading scales are useful guides for algorithm development, and automated grading promises increased repeatability over human grading. The method in [36] grades breast cancer by analyzing the size, shape, and texture of the nuclei inside an ROI using a Bayesian classifier. Another approach is to include some notion of cellular components without specifically segmenting nuclei. In, [55] grayscale thresholding and morphology are used to find blobs that are then classified based on intensity and size into three different nucleus types and stroma. The image is then segmented into areas of high and low nucleus density. Features including nucleus count, nucleus spacing, and tubule count are extracted for the high density areas only. The images are finally classified into grade one, two, or three with a quadratic classifier. In [39], Haralick features are extracted from cytoplasm and neuropil regions and used to determine the differentiation level of neuroblastoma tumors. The classification step is an ensemble of seven different classifiers and dimensionality reduction methods combined via weighted voting. The method handles WSIs by splitting them into small tiles that are processed in parallel as well as first classifying a downsampled version of the tile and using higher resolution data only if the classification certainty is low.

Finally, generic features can be used. In [56], prostate cancer is detected and graded based on color and color co-occurrence features and a random forest classifier. In [57], two different lung cancer subtypes are differentiated in images of tissue microarrays and hand-selected ROIs from full slides. The red and blue channels of the image are histogram-stretched to enhance the H&E contrast, and then Haralick and densitometric (e.g., mean pixel value, pixel center of mass, etc.) features are used in a boosting decision tree, achieving accuracy over 90%.

CONCLUSIONS

Histology is a critical tool in medicine and therefore automated histology analysis could have a profound effect on health-care quality, availability, and cost. At the same time, the field provides a host of fascinating signal processing challenges. How can tissue folds be detected and corrected? How can images of deformable tissue sections be registered to recover the 3-D shape of a tumor? What features are needed to explain the complex architecture of tissues such as skin? How can we detect rare cell types

with a scarcity of labeled data? And how can we run any of these methods efficiently on huge WSIs?

We hope this discussion has served both as a useful primer and as a call to action. Much has been accomplished in this field, but these systems are far from clinical acceptance. To get there, we need algorithms that demonstrate robustness and that solve relevant problems in medicine. Robustness will come from sharing data sets and algorithms so that they can be truly validated and compared. Relevance will depend on the continued efforts of pathologists and engineers to collaborate on defining and refining algorithms. Given the impact histology image analysis can make on the future of health care, it is well worth the effort.

AUTHORS

Michael T. McCann (mtmccann@cmu.edu) received the B.S. degree in biomedical engineering in 2010 from the University of Michigan. He is currently a Ph.D. student in the Center for Bioimage Informatics and the Department of Biomedical Engineering at Carnegie Mellon University, where he received a National Science Foundation Graduate Research Fellowship. His work focuses on developing signal processing tools for histology images, including methods for normalization, registration, and tissue segmentation/identification. He is a Student Member of the IEEE.

John A. Ozolek (ozolja@upmc.edu) graduated from Case Western Reserve University in 1985 with a B.S. degree in biology and obtained his M.D. degree from the University of Pittsburgh School of Medicine in 1989. He completed postgraduate residency training in pediatrics and anatomic pathology and fellowships in neonatal-perinatal medicine and pediatric pathology at the University of Pittsburgh. He is currently an assistant professor of pathology at the University of Pittsburgh and staff pathologist at the Children's Hospital of Pittsburgh. His research interests focus on the role of stem and neural crest cells, particularly in the pathogenesis of branchial cleft cysts, developmental abnormalities, and tumors of the head and neck in children. He has extensive collaborations with the Center for Bioimage Informatics as well as the Department of Biomedical Engineering and the Department of Mathematics at Carnegie Mellon University to develop algorithms for automating tissue recognition and correlation of histology and high-resolution magnetic resonance images.

Carlos A. Castro (ccastro@mwri.magee.edu) received the M.D. degree from Rosario University Medical School, Bogota, Colombia, in 1982 and the D.M.D. degree from Javeriana University Dental School in Bogota, Colombia, in 1988. He went on to do his maxillofacial surgery fellowship at Rosario University Medical School and subsequently did oral, orthognathic surgery, and gnathology specialties at the same universities as well as the Military University School of Medicine. He has worked in general pathology and subsequently surgical and perinatal pathology at the University of Pittsburgh and as a pathologist assistant at Magee-Womens Hospital in Pittsburgh, Pennsylvania. He has also worked as a senior research associate, visiting research instructor, senior faculty research instructor, and codirector of the Magee-Womens Research Institute and Foundation Histology Core. He is currently the director of the same histology core.

Bahram Parvin (b_parvin@lbl.gov) is a principal scientist at the Lawrence Berkeley National Laboratory and a professor of electrical and biomedical engineering at the University of Nevada, Reno. His laboratory focuses on technology development for realization of pathway pathology, elucidating molecular signatures of aberrant morphogenesis in normal and engineered matrices, and screening for probes for labeling and cargo delivery. He has published over 100 papers and was the general chair of the 2013 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. He is a member of the IEEE Bioimaging and Signal Processing Technical Committee and is an associate editor of *IEEE Transactions on Medical Imaging*.

Jelena Kovačević (j.kovacevic@ieee.org) received a Ph.D. degree from Columbia University. She then joined Bell Labs, followed by Carnegie Mellon University in 2003, where she is currently the Edward David Schramm Professor and head of the Department of Electrical and Computer Engineering, professor of biomedical engineering, and the director of the Center for Bioimage Informatics. She received the Belgrade October Prize and the E.I. Jury Award at Columbia University. She is a coauthor on an IEEE Signal Processing Society award-winning paper and is a coauthor of *Wavelets and Subband Coding* and *Foundations of Signal Processing*. She was editor-in-chief of *IEEE Transactions on Image Processing* (2002–2006). She was a keynote speaker at a number of meetings and has been involved in organizing numerous conferences. Her research interests include multiresolution techniques and biomedical applications. She is a Fellow of the IEEE.

REFERENCES

[1] K. S. Suvarna, C. Layton, and J. D. Bancroft, *Bancroft's Theory and Practice of Histological Techniques*, 7th ed. London, UK: Churchill Livingstone, Nov. 2012.

[2] D. J. Cook, *Cellular Pathology: An Introduction to Techniques and Applications*, 2nd ed. Oxfordshire, UK: Scion Publishing Ltd., June 2006.

[3] T. C. Cornish, R. E. Swapp, and K. J. Kaplan, "Whole-slide imaging: Routine pathologic diagnosis," *Adv. Anat. Pathol.*, vol. 19, no. 3, pp. 152–159, May 2012.

[4] S. Kothari, J. H. Phan, T. H. Stokes, and M. D. Wang, "Pathology imaging informatics for quantitative analysis of whole-slide images," *J. Am. Med. Inform. Assoc.*, vol. 20, no. 6, pp. 1099–1108, Aug. 2013.

[5] T. W. Bauer, L. Schoenfield, R. J. Slaw, L. Yerian, Z. Sun, and W. H. Henricks, "Validation of whole slide imaging for primary diagnosis in surgical pathology," *Arch. Pathol. Lab. Med.*, vol. 137, no. 4, pp. 518–524, Apr. 2013.

[6] S. Krishnamurthy, K. Mathews, S. McClure, M. Murray, M. Gilcrease, C. Albarracin, J. Spinosa, B. Chang, J. Ho, J. Holt, A Cohen, D. Giri, K. Garg, R. L. Bassett, and K. Liang, "Multi-institutional comparison of whole slide digital imaging and optical microscopy for interpretation of hematoxylin-eosin-stained breast tissue sections," *Arch. Pathol. Lab. Med.*, vol. 137, no. 12, pp. 1733–1739, Dec. 2013.

[7] L. Pantanowitz, J. H. Sinard, W. H. Henricks, L. A. Fatheree, A. B. Carter, L. Contis, B. A. Beckwith, A. J. Evans, A. Lal, and A. V. Parwani, "Validating whole slide imaging for diagnostic purposes in pathology: Guideline from the college of American pathologists pathology and laboratory quality center," *Arch. Pathol. Lab. Med.*, vol. 137, no. 12, pp. 1710–1722, Dec. 2013.

[8] S. Al-Janabi, A. Huisman, P. G. J. Nikkels, F. J. W. ten Kate, and P. J. van Diest, "Whole slide images for primary diagnostics of paediatric pathology specimens: A feasibility study," *J. Clin. Pathol.*, vol. 66, no. 3, pp. 218–223, Mar. 2013.

[9] M. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, Apr. 2009.

[10] T. J. Fuchs and J. M. Buhmann, "Computational pathology: Challenges and promises for tissue analysis," *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 515–530, Oct. 2011.

[11] M. W. Davidson. (2014, Feb.). Introduction to optical microscopy, digital imaging, and photomicrography. [Online]. Available: http://micro.magnet.fsu.edu/primer/index.html

[12] S. Kothari, J. H. Phan, R. A. Moffitt, T. H. Stokes, S. E. Hassberger, Q. Chaudry, A. N. Young, and M. D. Wang, "Automatic batch-invariant color segmentation of

histological cancer images," in *Proc. IEEE Int. Symp. Biomedical Imaging*, Chicago, IL, Mar. 2011, pp. 657–660.

[13] A. M. Khan, N. Rajpoot, D. Treanor, and D. Magee, "A non-linear mapping approach to stain normalisation in digital histopathology images using image-specific colour deconvolution," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1729–1738, June 2014.

[14] H. Chang, J. Han, A. Borowsky, L. Loss, J. W. Gray, P. T. Spellman, and B. Parvin, "Invariant delineation of nuclear architecture in glioblastoma multiforme for clinical and molecular association," *IEEE Trans. Med. Imag.*, vol. 32, no. 4, pp. 670–682, Apr. 2013.

[15] U. D. Braumann, J. P. Kuska, J. Einenkel, L. C. Horn, M. Loffler, and M. Hockel, "Three-dimensional reconstruction and quantification of cervical carcinoma invasion fronts from histological serial sections," *IEEE Trans. Med. Imag.*, vol. 24, no. 10, pp. 1286–1307, Oct. 2005.

[16] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," *Anal. Quant. Cytol. Histol.*, vol. 23, no. 4, pp. 291–299, Aug. 2001.

[17] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," in *Proc. IEEE Int. Symp. Biomedical Imaging*, Boston, MA, June 2009, pp. 1107–1110.

[18] M. Gavrilovic, J. C. Azar, J. Lindblad, C. Wahlby, E. Bengtsson, C. Busch, and I. B. Carlbom, "Blind color decomposition of histological images," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 983–994, June 2013.

[19] P. Tadrous, "Digital stain separation for histological images," *J. Microsc.*, vol. 240, no. 2, pp. 164–172, Nov. 2010.

[20] M. T. McCann, J. Majumdar, C. Peng, C. A. Castro, and J. Kovacevic, "Algorithm and benchmark dataset for stain separation in histology images," in *Proc. IEEE Int. Conf. Image Processing*, Paris, Oct. 2014.

[21] P. A. Bautista and Y. Yagi, "Improving the visualization and detection of tissue folds in whole slide images through color enhancement," *J. Pathol. Inform.*, vol. 1, p. 25, Nov. 2010.

[22] S. Kothari, J. H. Phan, and M. D. Wang, "Eliminating tissue-fold artifacts in histopathological whole-slide images for improved image-based prediction of cancer grade," *J. Pathol. Inform.*, vol. 4, p. 22, Aug. 2013.

[23] J. Chappelow, J. E. Tomaszewski, M. Feldman, N. Shih, and A. Madabhushi, "HistoStitcher([copyright]): An interactive program for accurate and rapid reconstruction of digitized whole histological sections from tissue fragments," *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 557–567, Mar. 2011.

[24] M. Schwier, T. Böhler, H. K. Hahn, U. Dahmen, and O. Dirsch, "Registration of histological whole slide images guided by vessel structures," *J. Pathol. Inform.*, vol. 4 (Suppl), no. 2, p. 10, 2013.

[25] D. Mueller, D. Vossen, and B. Hulsken, "Real-time deformable registration of multi-modal whole slides for digital pathology," *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 542–556, June 2011.

[26] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. Pluim, "elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, Jan. 2010.

[27] L. Cooper, O. Sertel, J. Kong, G. Lozanski, K. Huang, and M. Gurcan, "Feature-based registration of histopathology images with different stains: An application for computerized follicular lymphoma prognosis." *Comput. Methods Programs Biomed.*, vol. 96, no. 3, pp. 182–192, Dec. 2009.

[28] G. Xiao, B. N. Bloch, J. Chappelow, E. M. Genega, N. M. Rofsky, R. E. Lenkinski, J. Tomaszewski, M. D. Feldman, M. Rosen, and A. Madabhushi, "Determining histology-MRI slice correspondences for defining MRI-based disease signatures of prostate cancer," *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 568–578, Jan. 2011.

[29] H. Park, M. R. Piert, A. Khan, R. Shah, H. Hussain, J. Siddiqui, T. L. Chenevert, and C. R. Meyer, "Registration methodology for histological sections and in vivo imaging of human prostate," *Acad. Radiol.*, vol. 15, no. 8, pp. 1027–1039, Aug. 2008.

[30] A. Goode, B. Gilbert, J. Harkes, D. Jukic, and M. Satyanarayanan, "OpenSlide: A vendor-neutral software foundation for digital pathology," *J. Pathol. Inform.*, vol. 4, p. 27, Sept. 2013.

[31] H. Chang, G. Fontenay, J. Han, G. Cong, F. Baehner, J. Gray, P. Spellman, and B. Parvin, "Morphometic analysis of TCGA glioblastoma multiforme," *BMC Bioinform.*, vol. 12, no. 1, p. 484, Dec. 2011.

[32] TCGA Research Network. (2014, Aug.). The cancer genome atlas. [Online]. Available: http://cancergenome.nih.gov/

[33] W.-K. Jeong, J. Schneider, S. Turney, B. Faulkner-Jones, D. Meyer, R. Westermann, R. Reid, J. Lichtman, and H. Pfister, "Interactive histology of large-scale biomedical image stacks," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 6, pp. 1386–1395, Nov. 2010.

[34] R. Marée, B. Stévens, L. Rollus, N. Rocks, X. M. Lopez, I. Salmon, D. Cataldo, and L. Wehenkel. (2013, Sept.). A rich internet application for remote visualization and collaborative annotation of digital slides in histology and cytology. *Diagn. Pathol.*, 8(Suppl 1), p. S26. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3849538/

[35] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 841–852, Apr. 2010.

[36] C.-H. H. Huang, A. Veillard, L. Roux, N. Loménie, and D. Racoceanu, "Time-efficient sparse analysis of histopathological whole slide images," *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 579–591, Dec. 2011.

[37] V. Roullier, O. Lézoray, V.-T. Ta, and A. Elmoataz, "Multi-resolution graph-based analysis of histopathological whole slide images: Application to mitotic cell extraction and visualization." *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 603–615, May 2011.

[38] J. P. Monaco, J. E. Tomaszewski, M. D. Feldman, I. Hagemann, M. Moradi, P. Mousavi, A. Boag, C. Davidson, P. Abolmaesumi, and A. Madabhushi, "High-throughput detection of prostate cancer in histological sections using probabilistic pairwise Markov models," *Med. Image Anal.*, vol. 14, no. 4, pp. 617–629, Aug. 2010.

[39] J. Kong, O. Sertel, H. Shimada, K. L. Boyer, J. H. Saltz, and M. N. Gurcan, "Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation," *Pattern Recogn.*, vol. 42, no. 6, pp. 1080–1092, June 2009.

[40] H. Fatakdawala, J. Xu, A. Basavanhally, G. Bhanot, S. Ganesan, M. Feldman, J. E. Tomaszewski, and A. Madabhushi, "Expectation-maximization-driven geodesic active contour with overlap resolution (EMaGACOR): Application to lymphocyte segmentation on breast cancer histopathology," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1676–1689, July 2010.

[41] J. Han, T. Breckon, D. Randell, and G. Landini, "The application of support vector machine classification to detect cell nuclei for automated microscopy," *Mach. Vis. Appl.*, vol. 23, no. 1, pp. 15–24, June 2012.

[42] K. Belkacem-Boussaid, M. Pennell, G. Lozanski, A. Shana'ah, and M. Gurcan. (2010, Oct.). Computer-aided classification of centroblast cells in follicular lymphoma. *Anal. Quant. Cytol. Histol.*, 32(5), pp. 254–260. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3078581/

[43] L. Roux, D. Racoceanu, N. Loménie, M. Kulikova, H. Irshad, J. Klossa, F. Capron, C. Genestie, G. Le Naour, and M. N. Gurcan, "Mitosis detection in breast cancer histological images An ICPR 2012 contest," *J. Pathol. Inform.*, vol. 4, p. 8, May 2013.

[44] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *Proc. Med. Image Comput. Comput. Assist. Interv.*, Nagoya, Japan, 2013, pp. 411–418.

[45] E. Ozdemir and C. Gunduz-Demir, "A hybrid classification model for digital pathology using structural and statistical pattern recognition," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 474–483, Feb. 2013.

[46] K. Belkacem-Boussaid, S. Samsi, G. Lozanski, and M. N. Gurcan, "Automatic detection of follicular regions in H&E images using iterative shape index," *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 592–602, Oct. 2011.

[47] C. Chen, J. A. Ozolek, W. Wang, and G. K. Rohde, "A general system for automatic biomedical image segmentation using intensity neighborhoods," *Int. J. Biomed. Imag.*, vol. 2011, Mar. 2011.

[48] M. T. McCann, D. G. Mixon, M. C. Fickus, C. A. Castro, J. A. Ozolek, and J. Kovačević, "Images as occlusions of textures: A framework for segmentation," *IEEE Trans. Image Processing*, vol. 23, no. 5, pp. 2033–2046, May 2014.

[49] A. B. Tosun, M. Kandemir, C. Sokmensuer, and C. Gunduz-Demir, "Object-oriented texture analysis for the unsupervised segmentation of biopsy images for cancer detection," *Pattern Recogn.*, vol. 42, no. 6, pp. 1104–1112, June 2009.

[50] J. Diamond, N. H. Anderson, P. H. Bartels, R. Montironi, and P. W. Hamilton, "The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia," *Hum. Pathol.*, vol. 35, no. 9, pp. 1121–1131, Sept. 2004.

[51] A. Cagri, A. Burak, C. Aykanat, C. Sokmensuer, and C. Gunduz-Demir. (2012, June). Multilevel segmentation of histopathological images using cooccurrence of tissue objects. *IEEE Trans. Biomed. Eng.*, 59(6), pp. 1681–1690. [Online]. Available: http://dx.doi.org/10.1109/tbme.2012.2191784

[52] S. Kothari, J. Phan, A. Young, and M. Wang, "Histological image classification using biologically interpretable shape-based features," *BMC Med. Imag.*, vol. 13, no. 9, Mar. 2013.

[53] Y. Zhang, B. Zhang, F. Coenen, and W. Lu, "Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles," *Mach. Vis. Appl.*, vol. 24, no. 7, pp. 1405–1420, Oct. 2013.

[54] H. Chang, N. Nayak, P. Spellman, and B. Parvin, "Characterization of tissue histopathology via predictive sparse decomposition and spatial pyramid matching," in *Proc. Int. Conf. Medical Image Computing Computer-Assisted Intervention*, Nagoya, Japan, Sept. 2013, vol. 8150, pp. 91–98.

[55] S. Petushi, F. Garcia, M. Haber, C. Katsinis, and A. Tozeren. (2006, Oct.). Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer. *BMC Med. Imag.*, 6(1), p. 14. [Online]. Available: http://www.biomedcentral.com/1471-2342/6/14

[56] M. D. DiFranco, G. O'Hurley, E. W. Kay, R. W. Watson, and P. Cunningham, "Ensemble based system for whole-slide prostate cancer probability mapping using color texture features," *Comput. Med. Imaging Graph.*, vol. 35, nos. 7–8, pp. 629–645, Jan. 2011.

[57] C.-W. Wang and C.-P. Yu, "Automated morphological classification of lung cancer subtypes using H&E tissue images," *Mach. Vis. Appl.*, vol. 24, no. 7, pp. 1383–1391, Oct. 2013.

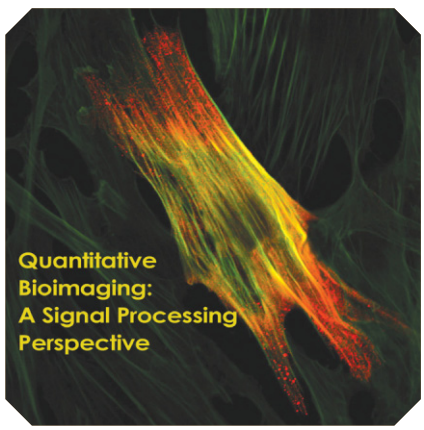
[SP]

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

Pouyan Mohajerani, Stratis Tzoumas, Amir Rosenthal, and Vasilis Ntziachristos

Optical and Optoacoustic Model-Based Tomography

Theory and current challenges for deep tissue imaging of optical contrast



© ISTOCK PHOTO.COM/BEANO5

Light offers a range of interactions with tissue that give rise to an extensive list of methods to sense physical, chemical, or biological processes. Combined with using safe and nonionizing radiation, optical imaging is considered as a fundamental tool in the biomedical sciences [1].

Enabling optical imaging beyond the reach of microscopy depths has been considered for bringing unique optical contrast in visualizing small animals or human tissues in vivo [1]–[5]. Optical methods can offer label-free interrogation of various physiological, metabolic, or disease states. From cellular and vascular imaging to sensing of hemoglobin oxygenation, melanin, and other intrinsic tissue chromophores and genetic reporters, bio-optics can reveal a wealth of functional, disease, and drug efficacy biomarkers [6]. In addition, optical agents such as targeted fluorochromes and photoabsorbing nanoparticles are investigated to impart contrast on specific cellular and subcellular tissue parameters [4], [7], [8]. The ability to resolve intrinsic chromophores, targeted agents and anatomical and functional characteristics in vivo is a particularly strong feature of the optical methods.

Visualizing versatile optical contrast in tissues is nevertheless confronted by photon diffusion, i.e., the random scattering of photons in tissue due to multiple elastic scattering events. Scattering rapidly reduces the imaging resolution that can be achieved with depth and complicates quantification. There are currently two main approaches in addressing photon diffusion beyond the microscopy penetration limit. One strategy is to use ultrasonic detection in the form of optoacoustic methods to avoid the effects of photon scattering. The second is to numerically account for the effects of photon propagation to decompose diffusion effects from the images generated.

A common methodological approach exists in both imaging technologies: increased imaging accuracy is granted when using mathematical models of the underlying physical phenomena and of the excitation and detection processes. These models are often employed in the framework of image reconstruction approaches and offer significant imaging improvements over analytical solutions, such as filtered back-projection algorithms, by accounting for system related parameters and reducing imaging artifacts.

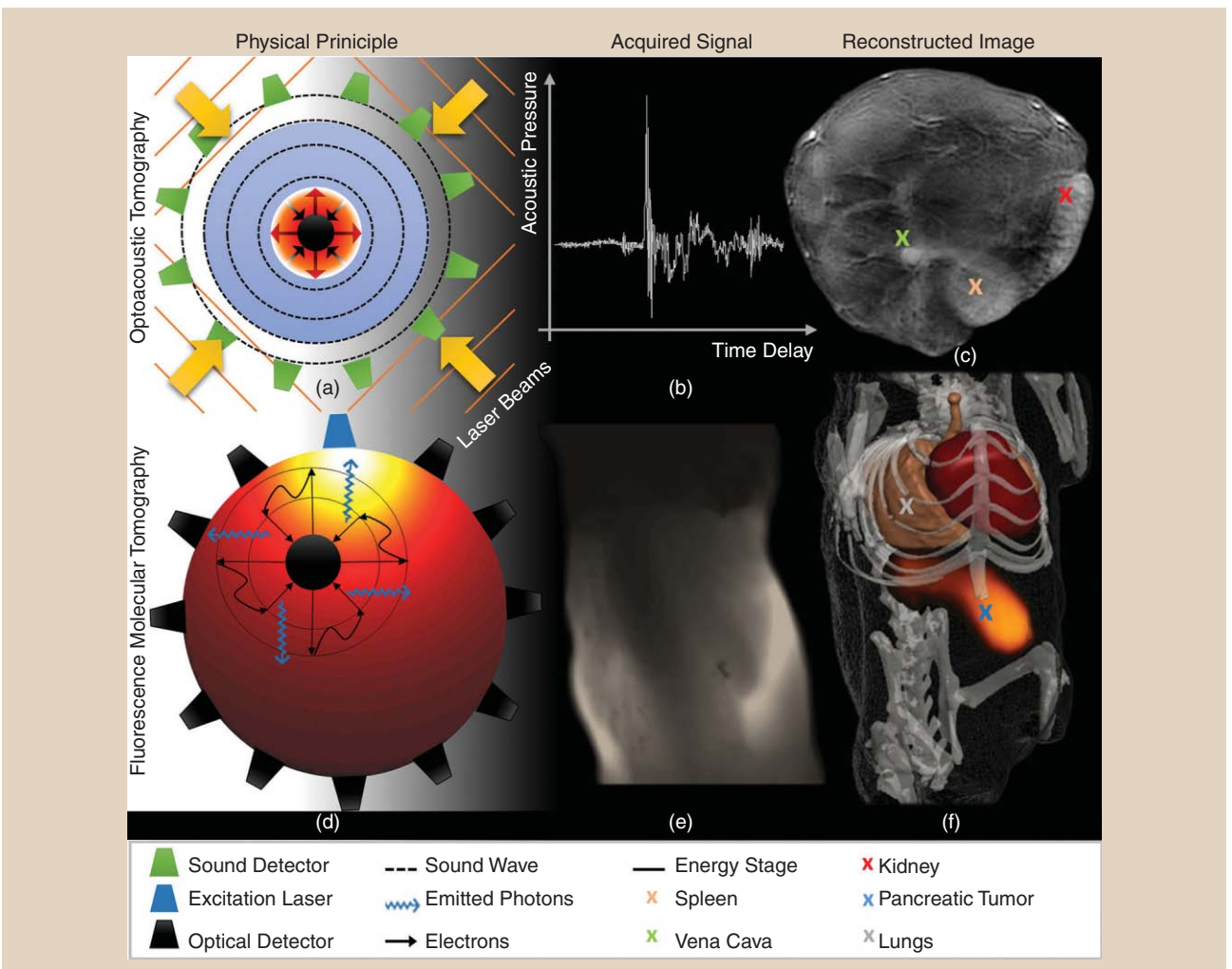
Optoacoustic tomography measures ultrasonic waves generated within tissues after absorption of light of transient intensity [9], [10]. The method forms optical absorption images with ultrasound resolution up to several millimeters to centimeters deep inside tissues. A number of optoacoustic imaging systems have been suggested ranging from microscopy [14] to tomographic systems for small animal imaging [2] or systems developed for clinical translations [3]. A powerful feature of optoacoustic techniques is the ability to obtain images at multiple wavelengths and spectrally unmix the absorption signatures of various tissue chromophores [4]. In particular, multispectral optoacoustic tomography (MSOT) offers remarkable anatomical, functional, and molecular imaging capabilities, enabling high-resolution visualization of photoabsorbing agents in tissues.

Digital Object Identifier 10.1109/MSP.2014.2352672
Date of publication: 5 December 2014

1053-5888/15©2015IEEE

Alternatively, images can be formed by utilizing optical detectors, for example, charge-coupled device (CCD) cameras, instead of acoustic detectors. Fluorescence molecular tomography (FMT) has been developed for molecular imaging of tissues [5], [13], [15], [16]. FMT aims at noninvasive and quantitative imaging of the biodistribution of fluorochromes in deep tissue. Hybrid modalities combining FMT with computed tomography (CT) and magnetic resonance imaging (MRI) as well as other anatomical or functional modalities have also gained significant momentum in the past five years [5], [13], [16]. FMT has been further deployed for preclinical and clinical applications using both specific and nonspecific probes [5], [16].

MSOT and FMT technologies aim at resolving tissue biomarkers in vivo for biological discovery and clinical applications (Figure 1). They share a common excitation problem, as they utilize light for the excitation of a secondary wave, which is eventually detected. Both technologies typically employ light in the near-infrared (NIR) range (650–900 nm) for excitation due to the relatively low tissue absorption in this region [17]. On the other hand, there are two major differences between the two modalities. First, MSOT typically applies extended illumination over a broad area allowing for real-time volumetric imaging, whereas FMT improves its accuracy by utilizing consecutive point illumination. Second, and most importantly, the nature of the secondary wave generated is different, i.e., ultrasound versus photon-density waves. This physical difference affects the numerical behavior of the inverse problem and largely defines the performance characteristics of the two modalities.



[FIG1] Principles of optical and optoacoustic tomography. (a) Themorelastic expansion of an optically absorbing object (black circle) within tissue (blue circle) upon illumination by pulsed laser beams. The object expands and contracts, due to temperature variation, and releases the absorbed energy as pressure waves (dotted circles). (b) Typical time-resolved optoacoustic signal detected using an ultrasound sensor. (c) A reconstructed transversal optoacoustic image of the abdominal region of a mouse, using a two-dimensional (2-D) circular measurement system geometry [11], [12]. (d) The principles of fluorescence, as electrons are excited to higher energy levels upon absorbing photons. Fluorescence photons are then emitted as the excited electrons vibrationally relax to their base states. (e) Fluorescence image acquired with a CCD camera from the dorsal side of a mouse. (f) A three-dimensional (3-D) image of a pancreatic tumor model reconstructed with concurrent X-ray CT and fluorescence molecular tomography (FMT-XCT), in 360° transillumination geometry [13].

This article offers a tutorial review of the major features of optical and optoacoustic tomography using model-based implementations. It further aims to convey the major challenges and performance characteristics of the technologies in a comparative manner and to survey pioneering works that address such challenges using signal processing techniques. Recent comprehensive surveys have addressed the hardware and system design aspects [2], [10], [16] as well as the biomedical applications [5], [8] of these technologies in detail. Therefore, these aspects are not covered in this article. Thorough reviews have also addressed the physical aspects of optoacoustic imaging [9], [10] and photoacoustic microscopy [14], the theory of light propagation in tissue [17], as well as the nonfluorescent diffuse optical tomography (DOT) [18]. These subjects are also accordingly avoided in our review. Finally, as time- and frequency-domain FMT systems share inversion challenges of continuous-wave (CW) systems and due to the prevalence of the latter, our discussion herein is focused on CW FMT systems.

FROM CELLULAR AND VASCULAR IMAGING TO SENSING OF HEMOGLOBIN OXYGENATION, MELANIN, AND OTHER INTRINSIC TISSUE CHROMOPHORES AND GENETIC REPORTERS, BIO-OPTICS CAN REVEAL A WEALTH OF FUNCTIONAL, DISEASE, AND DRUG EFFICACY BIOMARKERS.

MODEL-BASED TOMOGRAPHY

Tomographic imaging with optical and optoacoustic techniques involves excitation of the tissue with an optical source and the subsequent measurement of the converted energy that is emitted as a consequence. The propagation of light in tissue is generally governed by the radiative transfer equation (RTE), which can be approximated by the diffusion equation (DE) when the scattering events occur much more frequently than the absorption events [17]. The propagation of excitation photons as well as the converted energy (acoustic wave or fluorescence) is then governed in the frequency domain by the following partial differential equation (PDE):

$$\nabla^2 E + k^2 E = Q, \tag{1}$$

where k is designated as the wave number, E is the scalar field denoting the generated pressure wave or the emitted fluorescence intensity, and Q is the initial pressure or optical source term. Real values for k (valid for acoustic propagation in acoustically homogeneous, nonattenuating medium) denote the conventional Helmholtz equation. Complex values of k (valid for optical propagation under the diffusion approximation) lead to Helmholtz-like PDEs.

The general mathematical model relating the underlying image X to the measured emitted energy vector M can be described by a forward operator F as

$$M = F(\Delta, \{\Pi, X\}), \tag{2}$$

where Π denotes the set of all physical parameters of the tissue, such as energy conversion coefficients and optical or acoustic properties or constants. The set of spatial descriptions of the tissue, such as boundary conditions and measurement geometry, is further denoted by Δ.

The forward operator F can be highly nonlinear. Also, closed-form solutions of the underlying PDEs might not exist for arbitrary boundary conditions. Consequently, volumetric discretization methods, such as the finite element method (FEM), are typically considered to numerically solve the PDEs. Since solving nonlinear inverse problems is particularly challenging, simplified assumptions are often preferred toward the linearization of F. For instance, assuming an approximation of Π, which in itself is often part of the solution to be estimated, can be used to linearize F. The linearized problem is then expressed as

$$M = \mathbf{W}x, \tag{3}$$

where $\mathbf{W}$ is referred to as the *weight matrix* and the x is a discretized version of X over the spatial grid or mesh employed. The derivation of the weight matrix $\mathbf{W}$ is problem specific and is the subject of the sections "Optoacoustic Tomography" and "FMT."

The inversion of (3) involves estimation of the unknown image x, given the measurement vector M. When the model matrix $\mathbf{W}$ is well-conditioned, iterative techniques solving (3) in a least square sense, like conjugate-gradient methods, may be used for inversion. Alternatively, the pseudoinverse of $\mathbf{W}$ may be precalculated to define an inversion operator that is valid independent of the sample imaged. Conversely, for ill-conditioned $\mathbf{W}$, as is often the case in optical tomography [17], the inverse problem is frequently handled through approaches that limit the solution space, such as regularization methods. Tools from linear algebra are often utilized for characterizing the condition of the inverse problem and for imposing appropriate regularization.

The regularized inverse problem in the linearized model-based approach can be generally formulated as

$$X = \arg\min(\| \mathbf{G}(\mathbf{W}x - M)\|_2 + \lambda \| Lx \|_p), \tag{4}$$

where L is called the *regularization matrix*, λ is the regularization parameter, and $\mathbf{G}$ is determined using noise statistics. The parameter p defines the norm of the solution that is constrained through the regularization. The amount of regularization applied to the solution in the inversion is controlled by the parameter λ. Proper setting of λ, often performed based on L-curve analysis or cross-validation techniques, poses an important challenge and constitutes a tradeoff between model-fitting accuracy and a desired property of the solution (such as smoothness or sparsity).

OPTOACOUSTIC TOMOGRAPHY

The optoacoustic tomography problem aims at reconstructing the acoustic source distribution in a given volume from a set of acoustic measurements performed at multiple locations outside that

volume [19], [20]. Using the wave equation in the time domain, the pressure wave $p(\mathbf{r}, t)$ generated from an initial pressure distribution $p(\mathbf{r}, t = 0)$ can be written as [9], [21]

$$p(\mathbf{r}, t) = \frac{1}{4\pi v} \frac{\partial}{\partial t} \int_{|\mathbf{r}-\mathbf{r}'|=vt} \frac{p(\mathbf{r}', t = 0)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \tag{5}$$

where $\mathbf{r}$ and t are the spatial and temporal coordinates, respectively. The integration is performed over a sphere with its center at $\mathbf{r}$ and a radius of vt, with v denoting speed of sound. Using other terminology, for a given value of $\mathbf{r}$, $p(\mathbf{r}, t)$ is a projection of $p(\mathbf{r}, t = 0)$ over a sphere.

We restrict our discussion to the case in which the acoustic source distribution is a result of instantaneous energy deposition in the imaged object. We note that the acoustic inverse problem is agnostic to the nature of the energy deposited. Accordingly, the discussion herein is relevant to all fields of optoacoustics and thermoacoustics [19].

In the acoustic inverse problem, it is often assumed that $p(\mathbf{r}, t = 0)$ is nonzero only within a given volume V, where $p(\mathbf{r}, t)$ is known for all values of t over a detection surface S that encloses V. Under these conditions, the inverse problem has a unique solution; i.e., $p(\mathbf{r}, t = 0)$ may be uniquely determined from $p(\mathbf{r}, t)$ [22]. When $p(\mathbf{r}, t)$ is known over a sphere, a cylinder, or a plane, explicit solutions to the inverse problem exist in the form of frequency-domain or time-domain (so called back-projection) reconstruction formulas [23]. In other cases, although back-projection formulas may still be applied as an approximate solution, they may also lead to deformed images or accentuated image artifacts. However, practical considerations often lead to optoacoustic systems in which the detection surface S does not conform to the three detection geometries in which explicit solutions exist. In many cases S is open, and a unique solution is not guaranteed [24]. Additionally, the acoustic detectors employed in optoacoustic tomography are often larger than the typical acoustic wavelength of the signal, leading to projections that may not be well approximated by (5) [20], [25], [26]. Finally, in many cases, some a priori knowledge of the source is given, which may be valuable for improving the quality of the inversion. As an example, the quantity $p(\mathbf{r}, t = 0)$, as a measure of energy, is always nonnegative, a property that can be enforced in the inversion. However, such restrictions cannot be imposed on the source term when explicit solutions are used.

The mathematical operator that connects the source function and the projections is most often linear and therefore may be discretized to form a matrix $\mathbf{W}$. A matrix relation in (3) may then be applied for the optoacoustic system, for which M is a column vector representing the acoustic fields measured at a set of positions and time instants and x is a column vector representing the values of the initial pressure distribution on the grid. Each element of the weight matrix $\mathbf{W}$ relates to the initial pressure $p(\mathbf{r}, t = 0)$ at a position $\mathbf{r}$ in the volume to the specific pressure signals $p(\mathbf{r}_d, t)$ measured by the detector located at r_d at time point t.

To form the matrix $\mathbf{W}$, the initial pressure distribution is approximated as a sum of interpolation functions

$$p(\mathbf{r}, t = 0) \cong \sum_{j=1}^{n} f_j(\mathbf{r}) x_j, \tag{6}$$

where $x_j = p(\mathbf{r}_j, t = 0)$, $[r_1, \ldots, r_n]$ are the grid points on which the image is represented and $f_j(\mathbf{r})$ are the interpolation functions [19], [27]. By substituting (6) in (5), we obtain

$$m_i \cong \sum_{j=1}^{n} w_{ij} x_j, \tag{7}$$

where $m_i = p(\mathbf{r}_i, t_i)$, i.e., the pressure measured at a specific position $\mathbf{r}_i$ and time instant t_i, and

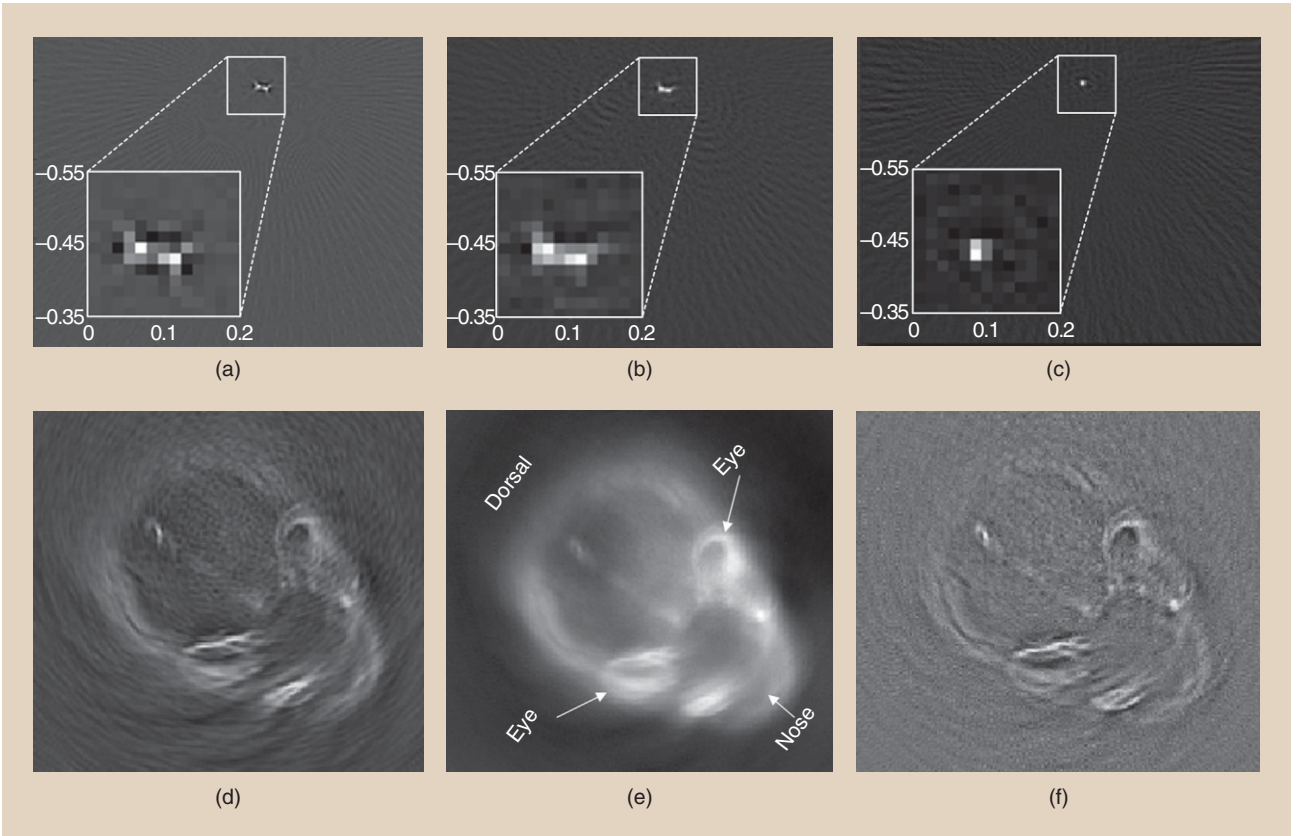
$$w_{ij} = \frac{1}{4\pi v} \frac{\partial}{\partial t} \int_{|\mathbf{r}_i-\mathbf{r}'|=vt_i} \frac{f_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \tag{8}$$

The matrix elements w_{ij} thus represent the response measured at position $\mathbf{r}_i$ and time instant t_i for the interpolation function $f_j(\mathbf{r})$. If the interpolation functions $f_j(\mathbf{r})$, $j = 1:n$ are chosen to have a finite support (i.e., to be spatially sparse), their corresponding pressure waves will be sparse in time and space. This property leads to sparsity in the matrix entries w_{ij}.

ONE OF THE MAIN ADVANTAGES OF THE MODEL-BASED APPROACH IS THE ABILITY TO INCLUDE IN THE INVERSION PROCESS A MODEL OF THE COMPLETE IMAGING SYSTEM.

One of the main advantages of the model-based approach is the ability to include a model of the complete imaging system in the inversion process. This approach has been shown to increase image quality over the one obtained with explicit formulas in cases of nonstandard detection surfaces or detectors larger than the typical acoustic wavelength generated by the imaged object [25] [Figure 2(a)–(c)]. Additionally, linear propagation effects such as acoustic attenuation and acoustic heterogeneity may be included to increase imaging fidelity [28]. However, such extensions to the model require either the use of a priori information or additional measurements. The quality of the reconstruction is not only determined by the performance of the imaging hardware but also by the accuracy of the discretization that leads to (3). In particular, when $p(\mathbf{r}, t = 0)$ is approximated with a finite sum of base functions that are not sufficiently smooth, the time derivative in (8) may lead to imaging artifacts [27], [29]. Alternatively, using a Fourier-domain formalism to model the optoacoustic operator leads to a nonsparse model matrix $\mathbf{W}$ and may lead to Gibbs-like artifacts owing to the finite number of harmonics used to approximate the Fourier transform [27]. Finite-support smooth interpolation functions have been therefore proposed to approximate $p(\mathbf{r}, t = 0)$ [27]. However, this approach comes at a price of longer calculation times compared to the use of discontinuous base functions.

Besides accurate modeling of system parameters, model-based tomography allows for imposing further restrictions in the inversion step by means of regularization. While regularization is not



[FIG2] Optoacoustic imaging under different reconstruction approaches (a)–(c) demonstrate the advantage of modeling system parameters (such as transducer shapes) using a pointlike source (human hair) with (a) filtered back-projection, (b) model-based inversion, and (c) improved model-based inversion incorporating the transducer shape. Tomographic reconstructions of a mouse head using (d) universal back-projection and (e) model-based inversion. (f) A high-pass filtered model-based image. While filtered back-projection only retains the high frequency anatomical features, model-based inversion shown in (e) offers a considerably enhanced imaging accuracy of absorbed energy density. [(a)–(c) reprinted from [26], and (d) and (e) reprinted from [30] with permission.]

always an essential step in optoacoustic image formation, it may significantly enhance the quality of the reconstructed image, especially in cases of limited-angle view geometries or under high noise levels [20]. Accordingly, many regularization techniques have been applied for optoacoustic tomography, including Tikhonov regularization (i.e., minimization of the L_2 norm of the solution), limited-iteration least squares (LSQR), truncated singular value decomposition, multiscale techniques, and total-variation regularization [20]. In general, the ability to seamlessly apply algebraic regularization techniques is a clear advantage of the model-based approach and often leads to images with less noise and artifacts than the ones obtained by analytical formulations or back-projection methods.

Despite the high versatility of model-based inversion algorithms, their use is far from ubiquitous. The reason is believed to be twofold. The first reason is the relative complexity involved in accurately modeling the optoacoustic setup. The use of a numerically inaccurate model may increase reconstruction error rather than decrease it. In contrast, many of the closed-form reconstruction formulas are very simple to implement. The second reason is the high computational complexity associated with inversion. In 2-D reconstruction, the memory required to store the model matrix is typically several gigabytes, whereas in 3-D imaging it is in the hundreds of gigabytes or more. For many of the algebraic inversion algorithms, such matrix sizes can be computationally prohibitive. This limitation can be mitigated via on-the-fly calculation of matrix elements [31] and utilization of model-separability in wavelet-packet domain [32].

While early applications of model-based algorithms focused mostly on numerical examples and phantom measurements, much of the current work is focused on application in high-resolution biological images [Figure 2(d)–(f)], which is an indication that the acceptance of these algorithms is on the rise. With the algorithms and computational resources available today, model-based inversion is already a viable option for real-time 2-D image formation and an acceptable option for 3-D reconstructions, when high throughput is not required. With expected improvements of these aspects, the role of model-based reconstruction in optoacoustic tomography is likely to increase in the future.

FMT

FMT illuminates tissue at multiple point-source locations and measures the light exiting tissue at multiple detector locations. Both excitation and emission occur in the NIR range. Structured

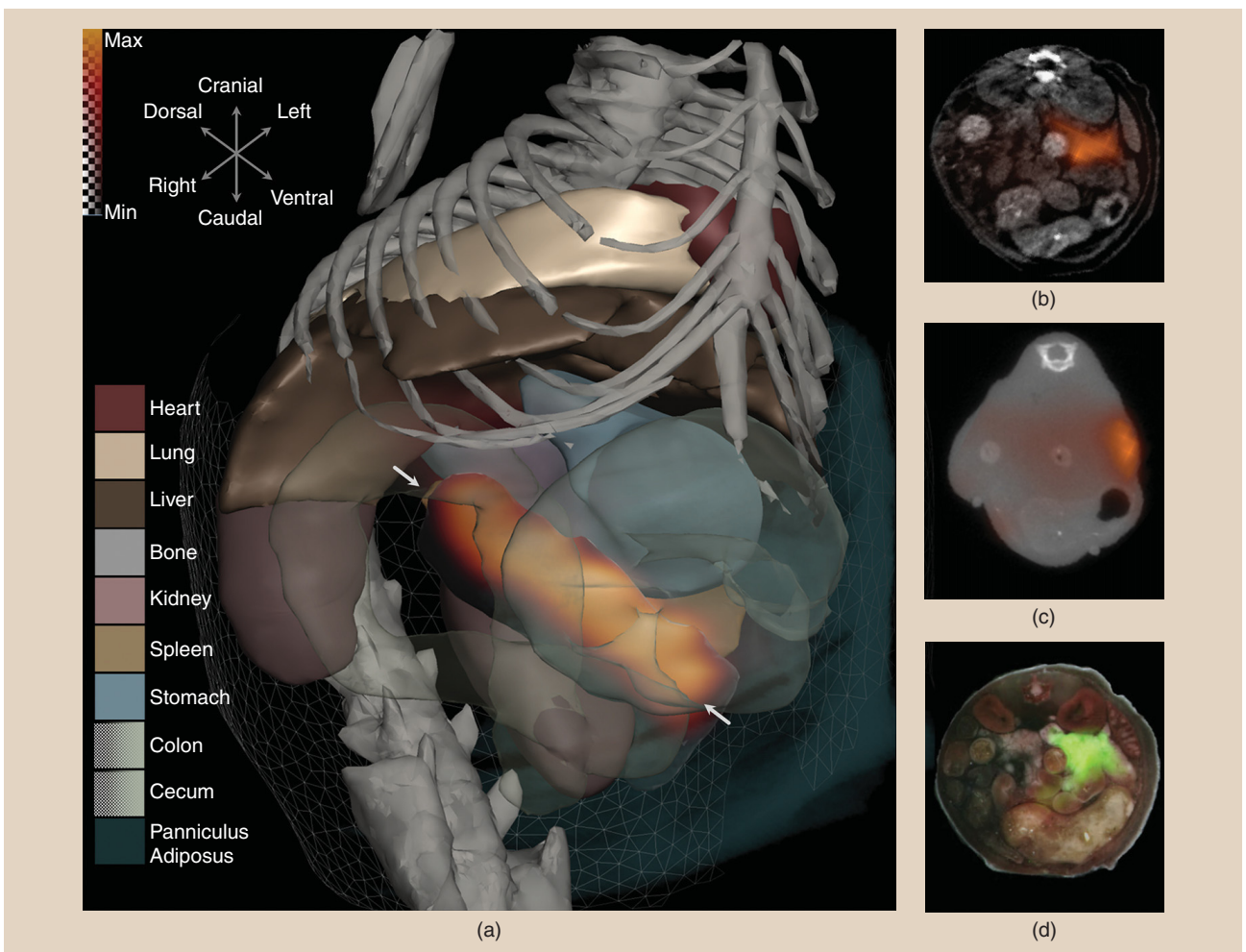
illumination using wide-field sources has been also proposed for FMT toward reducing imaging and inversion times [34]. The FMT forward model describes the relationship between the measurements collected from all the source-detector pairs implemented and the unknown fluorescence biodistribution. It is noted that both the unknown fluorophore distribution and tissue optical properties have an effect on the propagation of the excitation and the fluorescence photons. Therefore, the FMT inverse problem is essentially nonlinear. By assuming the effect of the unknown fluorophore distribution on the incident photon field to be negligible (as a Born-type approximation) and by assuming predetermined bulk values for endogenous tissue optical properties, the FMT problem can be linearized. In the linearized problem, the fluorescence fluence observed at the detector position $\mathbf{r}_d$ for illumination using the source located at $\mathbf{r}_s$ is given as

$$\phi_f(\mathbf{r}_d, \mathbf{r}_s) = a_f \int_V \frac{C(\mathbf{r}_d, \mathbf{r}_s)}{D} g_f(\mathbf{r}_d, \mathbf{r}) g_x(\mathbf{r}, \mathbf{r}_s)\, x(\mathbf{r})\, d^3\mathbf{r}, \quad (9)$$

where $x(\mathbf{r})$ is the unknown fluorophore concentration and D is the diffusion coefficient. The parameter a_f relates to physical constants, and $C(\mathbf{r}_d, \mathbf{r}_s)$ denotes the total coupling of light from the source to tissue and from tissue to detector. $g_x(\mathbf{r}, \mathbf{r}_s)$ ($g_f(\mathbf{r}_d, \mathbf{r})$) denotes the transported excitation (fluorescence) light from the position $\mathbf{r}$ to the source $\mathbf{r}_s$ (detector $\mathbf{r}_d$) [35]. The fluorescence signal is often normalized by the measurement at the excitation wavelength, to avoid the necessity of estimating the physical parameters a_f and $C(\mathbf{r}_d, \mathbf{r}_s)$ in (9) [35], as

$$\phi(\mathbf{r}_d, \mathbf{r}_s) := \frac{\phi_f(\mathbf{r}_d, \mathbf{r}_s)}{\phi_x(\mathbf{r}_d, \mathbf{r}_s)} = \beta \int_V \frac{g_f(\mathbf{r}_d, \mathbf{r}) g_x(\mathbf{r}, \mathbf{r}_s)}{g_x(\mathbf{r}_d, \mathbf{r}_s)} x(\mathbf{r})\, d^3\mathbf{r}, \quad (10)$$

where β can be found by experimental calibration. The inverse problem defined using this normalization is often referred to as the inversion of the Born ratio in the literature [15], [17], [35], [36]. Equation (10) can be discretized using a spatial mesh with N



[FIG3] A state of the art in hybrid optical molecular tomography: hybrid FMT–X-ray phase-contrast CT (PCCT) imaging of a pancreatic tumor using an αvβ3-targeting NIR probe in a pancreatic adenocarcinoma mouse model; (a) 3-D and (b) 2-D FMT-PCCT reconstructions in the form of fluorescence signal (in orange transparency) over the grayscale PCCT slice. Arrows in (a) show tumor extremities and organs were segmented from the PCCT scan. (c) FMT-XCT reconstruction showing low accuracy due to low soft-tissue contrast of CT in this case. (d) Ex vivo validation results, where green marks fluorescence overlaid on a cryosection photograph (reprinted from and used courtesy of [33]).

nodes. Assuming n illumination points (sources) and m detectors, the $(mn) \times N$ linearized weight matrix $\mathbf{W}$ for the Born ratio formulation is given as

$$\mathbf{W}((s,d),k) = \beta V_k g_f(\mathbf{r}_d, \mathbf{r}_k) g_x(\mathbf{r}_k, \mathbf{r}_s) / g_x(\mathbf{r}_d, \mathbf{r}_s), \qquad (11)$$

where V_k is the volume of the reconstruction voxel k.

The tissue optical properties are often assigned homogenous or organ-specific values using tabulated absorption and scattering values [37], which do not necessarily correspond to exact in vivo values. Hence, linearization of the FMT problem comes at the cost of a rather inaccurate model for light propagation. The effect of the forward problem on the imaging accuracy has been investigated [35], [37]. The Born ratio formulation has been shown to improve the imaging accuracy in thepresence of strong absorption deviations of the actual optical absorption of the tissue in respect to the assumed values [17], [35]. Strong scattering heterogeneities, however, cannot be compensated with Born ratio formulation [38]. For strong scattering heterogeneities, inversion solutions that use anatomical priors become accordingly important.

The FMT inverse problem consists of estimating the underlying fluorophore distribution x [(3)] from the surface optical measurements M [(3)], given the forward model $\mathbf{W}$ [(11)], and is often a highly ill-posed problem [17]. Accordingly, statistical or deterministic prior knowledge about the fluorophore distribution is often used in the inversion to mitigate the ill-posed nature of FMT. Such a priori information arises from specific statistical properties of the underlying distribution or is derived from data acquired simultaneously or sequentially using anatomical modalities. [An example of hybrid FMT imaging combined with phase-contrast CT, called *FMT-PCCT*, for imaging a mouse model of pancreatic ductal adenocarcinoma (PDAC) is presented in Figure 3.]

Various regularization methods are conventionally used to mitigate the ill-posed nature of the FMT inverse problem by constraining the solution space. While various regularization alternatives, or even regularization-free methods, have been proposed for FMT [16], the systems and methods developed and verified with in vivo data commonly use Tikhonov regularization or its variations [39], [40]. Moreover, FMT often uses target-specific probes, which generally localize in foci of disease and clear otherwise from healthy tissues. It is therefore possible to enforce the sparsity of the fluorophore distribution in the inversion process to increase imaging accuracy [41], [42]. Inversion based on minimization of the L_1 norm or the total variation of the solution, have been proposed as well. Nevertheless, the methods based on minimization of the L_p norm of the solution (or its gradient, in case of total variation methods) for $p < 2$, have been often limited to phantom data. Adapting these methods to in vivo applications presents an important opportunity for further development.

[TABLE 1] TWO-STEP DATA-DRIVEN INVERSION FOR HYBRID FMT; A FIRST STEP ESTIMATION USING A LOWER-DIMENSIONAL WEIGHT MATRIX $\mathbf{V}$ IS USED TO ESTIMATE A STATISTICAL MOMENT (HERE, MEAN) OF THE DISTRIBUTION IN EACH OF THE N ANATOMICAL SEGMENTS S_i. THESE ESTIMATES ARE THEN USED TO SHAPE THE REGULARIZATION MATRIX USING THE DECREASING FUNCTION ρ (DEFINED IN STEP 3) [13], [39].

1) FIRST STEP ESTIMATION: $x_1 = \text{argmin} \| (\mathbf{V}x - B) \|_2 + \lambda \| x \|_2$
2) CALCULATING AVERAGES: $\mu_1(i) = \text{mean}(x_1(v), v \in S_i)$
3) SHAPING THE REGULARIZER:
$\forall$ voxel $v \in S_i, w(v) = \rho(\mu_1(i)),\ \rho(\mu) = \dfrac{(\beta+1) \times \max(\mu_1)}{\mu + \beta \times \max(\mu_1)}$ (*)
4) SECOND STEP INVERSION:
$\text{Solution} = \text{argmin}(\| \mathbf{W}x - B \|_2 + \lambda \| \text{diag}(w)x \|_2)$.

(*) The parameter β controls the dynamic range of the regularization values ρ and is often set to a value close to 0.1.

Anatomical information, acquired simultaneously or sequentially using anatomical modalities (predominantly CT, MRI, and ultrasound), have been used as soft or hard priors to mitigate the ill-posed nature of the FMT inverse problem by constraining the solution space [5], [13], [40], [43]. Hard priors indicate enforcing a piece-wise constant solution, where the distribution has a constant value within each anatomical segment. Soft priors introduce various solution norms within each segment into the inversion optimization problem. Two methods are commonly used to enforce soft priors [39], [40]. The first approach, called the *method of weighted segments*, assigns different regularization values (segment weights) to the diagonal entries of the diagonal regularization matrix L, according to corresponding voxel-segment associations [13], [39]. The method of weighted segments relies on prior biological knowledge about preferential uptake of a given fluorescent probe in a specific volumetric segment (such as a lesion or organ delineable in the anatomical scans). Smaller regularization values are then assigned to voxels within that tissue segment, in comparison to other segments or the background tissue. The second approach makes use of the smoothness property within each segment, which can be enforced using the Laplacian operator [40]. Regularization based on the anisotropic diffusion function has been further proposed in conjunction with anatomic priors to preserve the anatomical boundaries in the reconstructed distribution [43].

Proper regularization based on anatomical priors can be a challenging task. The weighted-segments and Laplace-based methods are both prone to cause artificial accumulation or suppression of fluorescence in anatomical segments or to lead to oversmoothing, respectively. Instead, two-step data-driven inversion methods have been proposed to improve the approximations of the parameters of the structured regularization (Table 1). In these methods, the ill-posed nature of the FMT inversion is partially alleviated by performing a first inversion using a lower-dimensional 3-D inversion model (such as a low-resolution 3-D inversion mesh or a piecewise-constant model [39]). The results of the first inversion are then used to draw conclusions about the underlying fluorophore distribution. This knowledge is then used and enforced with the help of anatomical priors in a second step inversion performed on the full-resolution model [13], [39].

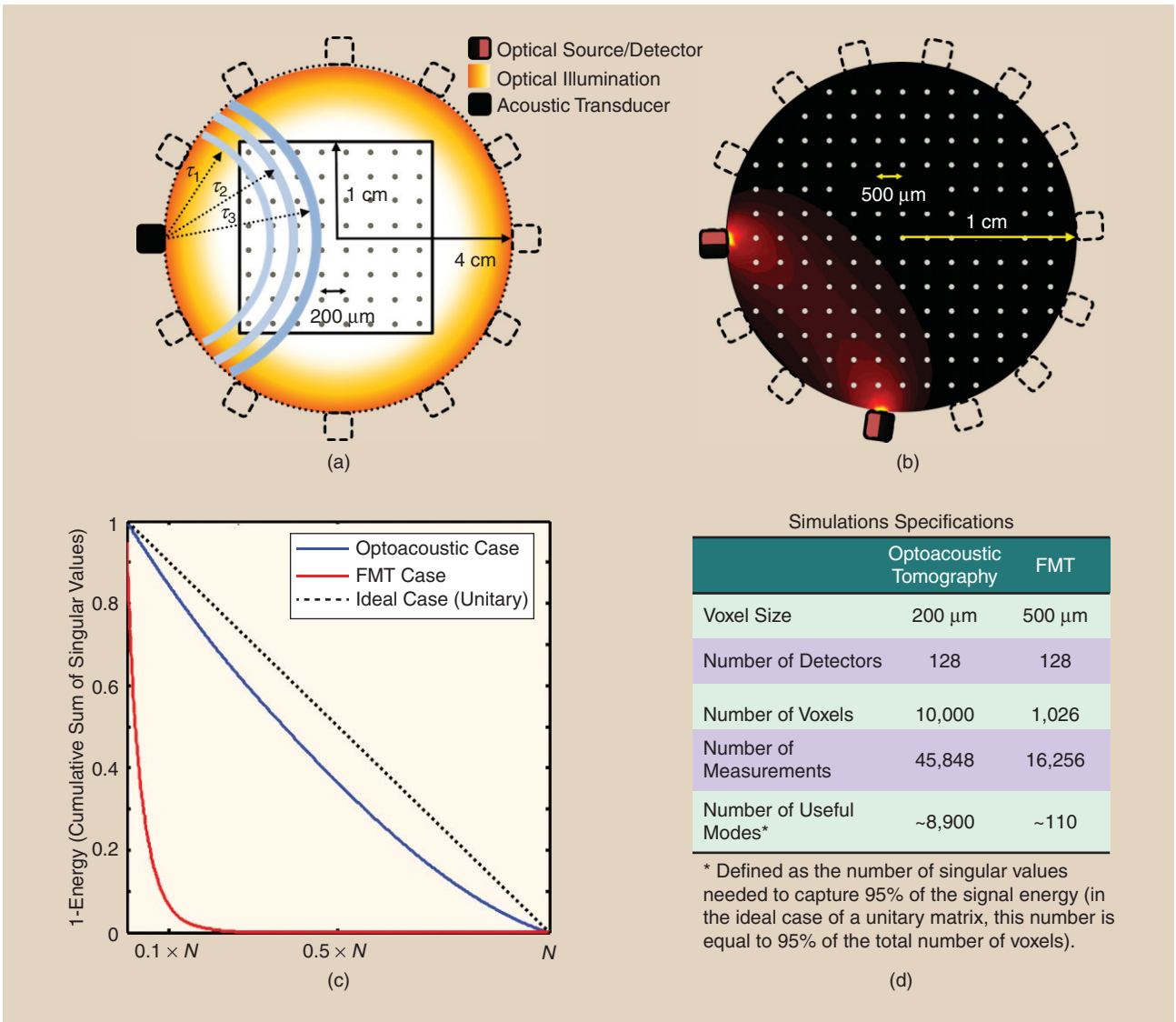
Statistical methods have been further applied toward FMT inversion. A nonlinear Bayesian optimization scheme was presented for extracting several biomarkers, including optical properties and fluorescence, in tissue [44]. Nevertheless, simultaneous inversion of the optical map and the fluorescence distribution is yet to be compared for in vivo data with application of predetermined

organ-specific optical properties—the common practice for in vivo imaging using FMT. Successful simultaneous inversion of such endogenous and exogenous biomarkers is challenging task, envisaged to improve the molecular imaging accuracy of FMT. Maximum-likelihood inversion of FMT using a stochastic model of the Born ratio has shown improvement over deterministic inversion for in vivo data, by taking advantage of a priori information derived from the statistical properties of the measurements [36]. Combinations of statistical inversion methods with anatomical priors offer as interesting possibility for further performance enhancement.

Due to the large data sizes in the FMT problem (tens of thousands of voxels and millions of data points), the FMT inversion is a computationally challenging problem. As a result, deterministic and statistical inversion techniques are often handled iteratively through algebraic reconstruction techniques or conjugate-gradient type methods (instead of through direction inversion). Adaptive mesh refinements have been further proposed for improving the reconstruction accuracy around the region of interest, while reducing the computational burden [45]. Data compression across sources as well as detectors has been further proposed and applied to in vivo data toward expediting the inversion [34], [46].

NUMERICAL BEHAVIOR OF OPTOACOUSTIC TOMOGRAPHY AND FMT

A comparative examination of the numerical behavior of the optoacoustic and the FMT inverse problems can offer a unique insight to the particular challenges faced in each imaging approach. The numerical behavior of linear inverse problems can be analyzed



Simulations Specifications

	Optoacoustic Tomography	FMT
Voxel Size	200 µm	500 µm
Number of Detectors	128	128
Number of Voxels	10,000	1,026
Number of Measurements	45,848	16,256
Number of Useful Modes*	~8,900	~110

* Defined as the number of singular values needed to capture 95% of the signal energy (in the ideal case of a unitary matrix, this number is equal to 95% of the total number of voxels).

(d)

[FIG4] A comparative numerical analysis of the inversion in optoacoustic tomography and FMT: (a) schematic illustration of the optoacoustic tomography simulation configuration, where blue arches correspond to different time delays τ and (b) the FMT simulation configuration. The sensitivity pattern of FMT between the shown source and detector positions is shown in (b). (c) One minus the cumulative sum of singular values of the respective weight matrices (N in abscissa is the number of nodes in the inversion grid). (d) Simulation settings and metrics are given. The results elucidate the ill-posed nature of the inverse problem in FMT compared to optoacoustic tomography.

using the singular value decomposition (SVD) of the respective weight matrices.

Two simplified 2-D imaging scenarios for optoacoustic tomography and FMT were accordingly analyzed (Figure 4). For the optoacoustic tomography case, a sparse weight matrix was calculated using the model-based approach presented in [27], while the FMT weight matrix was calculated using FEM-based discretization of the DE. A slower rate of decay in curves of Figure 4(c) corresponds to a larger number of useful (above the noise level) singular values and, hence, a better behaved inverse problem. For the FMT case, the first 10% of the singular values comprise around 95% of the weight matrix energy and are significant. For optoacoustic tomography, the system energy increases almost linearly with the number of singular values and is, hence, close to the ideal case of a unitary weight matrix (89% of singular values comprise 95% of the total energy). It should be noted that the linear systems in both cases were overdetermined, as indicated in Figure 4(d). Nevertheless, the FMT system was still highly ill-conditioned.

The numerical behavior of the associated inverse problems can be justified by examining the corresponding patterns of sensitivity of measurements to the solution. The acoustic signal measured on a given transducer location and a specific time point in optoacoustic tomography is sensitive only to a narrow (fewer than 100 μm) band of voxels [blue bands in Figure 4(a)]. This sparse sensitivity pattern leads to a well-behaved inverse problem. Conversely, a measurement point in FMT is sensitive to a much larger subset of voxels, due to strong tissue scattering. This fact leads to the substantially ill-posed nature of the FMT inverse problem, even given millions of data points acquired by CCD-based FMT systems.

QUANTITATIVE IMAGING AND MULTISPECTRAL UNMIXING

Molecular imaging aims at delivering an accurate mapping of the biodistribution of distinct molecules of biological interest. In this sense, the problem of differentiating between molecules for providing quantitative maps of molecular specificity is universal in the fields of optical and optoacoustic tomography. Spectroscopy techniques are commonly used in optical imaging to recover molecules according to their distinct spectral signatures. Multispectral imaging has been explored in FMT for separating the reporter molecules from the background fluorochromes to improve the molecular specificity [17]. Alternatively, the fluorescent lifetime contrast between different molecules can be used in time- or frequency-domain FMT to achieve molecular differentiation [47]. Conversely, multispectral tissue excitation plays a particularly important role for optoacoustic imaging since it is associated with both the functional (e.g., differentiation between oxy- and deoxy-hemoglobin) and molecular (i.e., extraction and quantification of extrinsic contrast agents) imaging character of the technology. Quantification and spectral unmixing in MSOT are particularly complex problems of nonlinear nature, as they are coupled to the description of light propagation in tissue.

QUANTIFICATION AND SPECTRAL UNMIXING IN MSOT ARE PARTICULARLY COMPLEX PROBLEMS OF NONLINEAR NATURE, AS THEY ARE COUPLED TO THE DESCRIPTION OF LIGHT PROPAGATION IN TISSUE.

OPTICAL FLUENCE AND SPECTRAL COLORING

Tissue absorption coefficient $\mu_a(\mathbf{r}, \lambda)$ at position $\mathbf{r}$ and excitation wavelength λ corresponds to a linear superposition of the concentrations of the local absorbers $c_i(\mathbf{r})$, weighted by their wavelength-dependent absorption coefficients $\varepsilon_i(\lambda)$ (absorption spectra):

$$\mu_a(\mathbf{r}, \lambda) = \sum_i c_i(\mathbf{r})\,\varepsilon_i(\lambda). \tag{12}$$

Given $\mu_a(\mathbf{r}, \lambda)$, the distinct tissue absorbers can be readily separated, and their concentrations $c_i(\mathbf{r})$ estimated using a linear regression method with the associated absorption spectra $\varepsilon_i(\lambda)$. In this way, valuable quantitative molecular information can be extracted. Physiological and metabolic information can be further inferred by computing the oxygen saturation of hemoglobin (sO_2) as the ratio between the oxygenated and total hemoglobin concentrations.

Multispectral optoacoustic images $P(\mathbf{r}, \lambda)$ are proportional to the spatial distribution of tissue optical absorption $\mu_a(\mathbf{r}, \lambda)$, and the unknown space and wavelength dependent optical fluence $\Phi(\mathbf{r}, \lambda)$ through the following relationship [48]:

$$P(\mathbf{r}, \lambda) = C(\mathbf{r})\Gamma(\mathbf{r})\Phi(\mathbf{r}, \lambda)\mu_a(\mathbf{r}, \lambda), \tag{13}$$

where $C(\mathbf{r})$ is a space-dependent scaling factor associated with effects not accounted for during the reconstruction (e.g., the system's spatial impulse response or ultrasound attenuation) and $\Gamma(\mathbf{r})$ refers to the Grüneisen coefficient of tissue, which may vary with tissue type [48]. Finally, the term $\mu_a(\mathbf{r}, \lambda)\Phi(\mathbf{r}, \lambda)$ is equivalent to the absorbed energy density (i.e., the energy per unit volume deposited in tissue), typically denoted as $H(\mathbf{r}, \lambda)$.

The spatial scaling factors $C(\mathbf{r})$, $\Gamma(\mathbf{r})$, and $\|\Phi(\mathbf{r})\|_2$ (the norm of optical fluence across λ) do not affect the underlying spectral responses, but rather the global intensity of the multispectral images. As such, they have no direct effect on the computation of sO_2, as a ratiometric value, or on the accurate separation of the distinct absorbers in tissue. The particular complications of spectral unmixing are typically introduced due to the wavelength dependence of the optical fluence. Different excitation wavelengths are attenuated differently as light propagates in tissue. Hence, the spectral pattern of optical fluence at different tissue positions can become highly nonuniform. The wavelength dependence of optical fluence alters the perceived spectral features with depth in a nonstraightforward manner. This change of spectral features, commonly referred in literature as *spectral coloring* [11], [48], depends on a multitude of parameters such as optical properties, imaging depth, and tissue physiology.

QUANTIFICATION OF ABSORBING COMPONENTS

The separation of the unknown optical fluence $\Phi(\mathbf{r}, \lambda)$ from $\mu_\alpha(\mathbf{r}, \lambda)$ comprises the main challenge for accurately estimating tissue optical absorption and realization of accurate quantitative physiological and molecular imaging with MSOT. Quantification approaches seek to model the optical fluence $\Phi(\mathbf{r}, \lambda)$ using light propagation models or heuristic fluence characterization approaches. Such algorithmic approaches may operate in the single or the multiwavelength domain. In both cases, multispectral excitation is favorable for obtaining bio-medically relevant information.

The problem of quantitative optoacoustic imaging is commonly formulated as a nonlinear optical property estimation problem [49], as the optical fluence $\Phi(\mathbf{r}, \lambda; \mu_a(\mathbf{r}, \lambda), \mu_s(\mathbf{r}, \lambda))$ is related to the spatial distribution of the optical absorption $\mu_a(\mathbf{r}, \lambda)$ and the reduced scattering coefficients of tissue $\mu_s(\mathbf{r}, \lambda)$. This relation is established with the help of a light propagation model such as the DE [49]. As such, estimations of $\mu_a(\mathbf{r}, \lambda)$ and $\mu_s(\mathbf{r}, \lambda)$ can be acquired by solving the following optimization problem using an iterative nonlinear optimization approach:

$$\arg \min_{\mu_a, \mu_s} \| H(\mathbf{r}, \lambda) - \mu_\alpha(\mathbf{r}, \lambda)\Phi(\mathbf{r}, \lambda; \mu_\alpha(\mathbf{r}, \lambda), \mu_s(\mathbf{r}, \lambda)) \|_2. \quad (14)$$

Model-based optical property estimation approaches may operate in the single wavelength domain by assuming known or uniform (reduced) scattering distribution μ_s. Under this assumption, a fixed-point iteration approach [49] and a regularized Newton method [50] were proposed for acquiring $\mu_a(\mathbf{r})$ values given $H(\mathbf{r})$. The assumption of known scattering may restrict the application of such methods in experimental images, since the optimization has been shown to diverge under an inaccurate selection of μ_s [51]. Furthermore, it has been shown that different combinations of $\mu_a(\mathbf{r})$ and $\mu_s(\mathbf{r})$ may result in the same $H(\mathbf{r})$, implying nonuniqueness in the solution [52]. Additional information acquired from multispectral excitation, in combination with prior knowledge on tissue absorption spectra and the wavelength dependence of scattering has been shown to alleviate this effect, as shown in [52], where both $\mu_a(\mathbf{r}, \lambda)$ [or, alternatively, the concentrations of tissue chromophores $c_i(\mathbf{r})$] and $\mu_s(\mathbf{r}, \lambda)$ can be simultaneously estimated under (14) using a Gauss–Newton optimization approach. Combining acoustic and optical inversion has also been considered for addressing further challenges related to the inaccurate estimation of the absorbed energy density $H(\mathbf{r}, \lambda)$ [53].

Model-based inversion approaches offer the advantage of relying on light propagation theory for solving the quantification problem. Proof-of-principle demonstration of their application has been showcased in phantom studies under a number of simplified assumptions [53]. However, the large number of unknown parameters as well as limitations in accurately estimating the absorbed energy density $\mathrm{H}(\mathbf{r}, \lambda)$, may restrict their applicability in the general case [48]. Hence, the adaptation of such methods to allow for robust application in structurally complex tissue images presents an important goal for further research.

Alternatively, heuristic approaches may seek to separate $\Phi(\mathbf{r})$ from $\mu_a(\mathbf{r})$ by relying on structural characteristic of the optical fluence rather than light propagation theory. By representing the optical fluence using a sparse Fourier basis and the absorption with a Haar wavelet basis, Rosenthal et al. achieved their separation using an orthogonal matching pursuit approach on the logarithmic transformation of the image [51]. The application of the method was demonstrated on experimental phantom images. However, its general applicability may also be limited as the performance of the method highly relies on the quality of the reconstructed image.

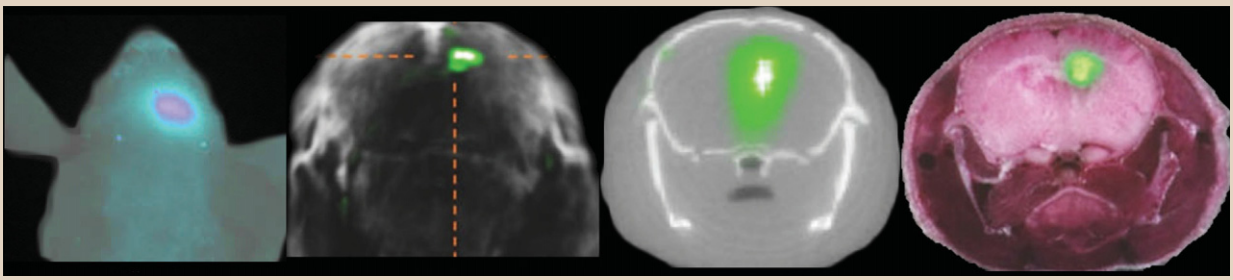
DETECTION OF CHROMOPHORES WITH MSOT

Accurate quantification of the concentrations of all chromophores within tissue presents a longstanding challenge in MSOT. However, multispectral excitation can be effectively used to extract molecular agents from the absorbing tissue background. When the spectral unmixing problem of MSOT is formulated as a molecular target detection problem, the goal is to extract the biodistribution of an extrinsic chromophore in tissue with high sensitivity and specificity [11]. In this context, sensitivity is typically defined as the minimum agent amount necessary for detection, while specificity relates to the minimization of false positives in detection. Molecular target detection approaches aim at modeling the spectral variability of the target molecule as well as the spectral variability of tissue background.

A straightforward approach to molecular contrast enhancement is the subtraction of two optoacoustic images acquired at two different excitation wavelengths, where the target molecule demonstrates significant variation in absorbance. This method assumes minor spectral variation of background tissue absorption for these particular wavelengths. Alternatively, a number of studies use a linear approximation of (13) by assuming constant $\Phi(\mathbf{r}, \lambda)$ throughout tissue and invert (12) for unmixing several chromophore using least squares methods [54]. Spectra of oxy- and deoxygenated hemoglobin are typically used for modeling the tissue absorption. Such simplifying assumptions may lead to inaccuracies thus, limiting the performance of such approximations in molecular target detection.

Blind source separation via independent component analysis (ICA) has also been considered in the context of multispectral optoacoustic imaging [55] and was found to outperform linear approximations due to the uncertainties associated with the optoacoustic spectral responses. In this respect, ICA and its variations [12] were used in a number of studies to extract molecular agents that were not detectable using conventional dual wavelength methods or linear approximations [one example is presented in Figure 5(b)].

More recently, statistical subpixel detection techniques were introduced in the context of MSOT and were shown to improve upon previous approaches in cases where the molecular target is sparsely present within the data (such as in tumor targeting or expression of fluorescent proteins) [11]. Statistical subpixel detection methods, like the adaptive matched filter (AMF), model the combined effect of tissue absorption and optical fluence, as well as



[FIG5] Demonstrating the potential of optical and optoacoustic model-based tomography via consecutive in vivo imaging of an NIR fluorescent protein (iRFP)-expressing brain tumor in a glioblastoma mouse model. (a) Epi-illumination fluorescence imaging. (b) MSOT transversal image with the unmixed iRFP signal (green transparency) overlaid to the anatomical optoacoustic image (gray). (c) FMT-XCT transversal image with the fluorescence signal (green transparency) overlaid on the CT image. (d) Ex vivo cryosectioning image, with the iRFP fluorescence signal (green transparency) overlaid on the color photograph of the cryosection. (Figure reprinted from and courtesy of [12].)

any other background spectral perturbations induced by noise and reconstruction artifacts, using a multivariate Gaussian model. The parameters of this model (i.e., the mean and covariance matrix of the Gaussian distribution) are estimated from the available data with maximum likelihood. Furthermore, the AMF features a constant false alarm rate (CFAR) property that allows for the definition of an application-independent global detection threshold, enabling molecular imaging of high specificity [11]. The combination of high sensitivity with high specificity makes the statistical subpixel detection method a very promising approach for accurate molecular imaging, while the adaptation of such algorithms to the particular challenges of MSOT presents an interesting area for further research.

TOWARD ROBUST MOLECULAR TOMOGRAPHIC IMAGING

MSOT and FMT are the two most extensively studied techniques currently employed for achieving in vivo optical molecular imaging beyond the microscopic regime. As such, although characteristically different in the detection process, the cross-examination of the two approaches holds promise for improving the respective imaging performances, for optimally selecting the most appropriate modality based on the application, and for combining the two approaches for robust molecular imaging solutions.

From a methodology standpoint, model-based tomographic reconstruction and tools stemming from linear and nonlinear inversion theories, regularization, and compressed sensing are frequently utilized for improving the performance of both modalities. As quantitatively shown in Figure 4, the tomographic reconstruction problem of FMT is substantially more complex due to the different nature of light versus ultrasound propagation in tissue. In MSOT, on the other hand, the effects and challenges of light propagation are identified in the spectral unmixing and quantification steps.

Light propagation in tissue defines complex nonlinear inversion problems in both of MSOT and FMT modalities. Nevertheless, nonlinear optimization methods for the simultaneous inversion of optical properties and target molecules have often been limited to simulation and phantom studies and should be yet investigated for in vivo data. Due to the complexity of the nonlinear approaches, the inversion problems are often handled via linearized or statistical models that frequently rely on heuristic assumptions on the distribution of the molecular agents (e.g., the one of sparsity).

Optical properties of tissue significantly affect the performance characteristics of these modalities, in distinct manners (Table 2). FMT resolution and quantification are especially affected by light scattering while tissue absorption plays a lesser role and mainly affects sensitivity. Optical scattering and absorption play a minor role in accuracy of optoacoustic tomography as long as quantification of the absorbers' concentrations is not pursued. However, they largely affect the quantification and spectral unmixing step through the space- and wavelength-dependent optical fluence. Robust quantification in multispectral optoacoustic imaging is considered an open problem and an active field of research. Moreover, detection of extrinsic molecules with high sensitivity is directly affected by high tissue absorption, which limits the signal-to-background interference ratio. Finally, molecular imaging sensitivity is also indirectly influenced by the wavelength-dependent optical fluence that increases the spectral variability of background and target molecules.

The sensitivity of MSOT in detecting molecular agents depends on a number of parameters such as the depth and volume of the lesion [4], the absorption coefficient of the molecular agent, or the spectral unmixing approach utilized [11]. While a thorough

[TABLE 2] ADVERSE IMPACT OF OPTICAL SCATTERING μ_s AND ABSORPTION μ_a ON PERFORMANCE METRICS OF FMT AND MSOT ("—" STRONGLY NEGATIVE IMPACT, "–" NEGATIVE IMPACT, AND "O" NO EFFECT).

	RESOLUTION		SENSITIVITY		QUANTIFICATION	
	μ_a	μ_s	μ_a	μ_s	μ_a	μ_s
FMT	–	—	—	–	–	—
MSOT	O	O	—	–	—	—

analysis of the sensitivity capabilities of the technology is yet to be performed, many studies report sensitivities in the lower micromolar scale for NIR fluorescent dyes such as ICG and IRDye [62]. Alternative contrast mechanisms like melanin or photoabsorbing nanoparticles may offer two to three orders of magnitude better sensitivity due to their high absorption coefficients [4]. FMT has demonstrated sensitivity in the nanomolar scale for in vivo studies [15]. This high sensitivity owes to the availability of very sensitive optical detectors and fluorescence agents with large quantum efficiency and binding affinity, as well as the low autofluorescence of tissue in NIR. FMT sensitivity can vary with depth, due to excitation light attenuation, and can be adversely affected by background fluorescence.

MSOT HAS FOR THE FIRST TIME ENABLED IN VIVO OPTICAL MOLECULAR IMAGING IN DEEP TISSUE AT HIGH SPATIAL RESOLUTION AND LOCALIZATION ACCURACY.

In summary, selecting or designing the right tool for the application at hand necessitates not only experimental evaluation of the different options but also a thorough theoretical and numerical cross-examination of the available technologies. Due to its high sensitivity and penetration depth, the role of FMT in the preclinical development and clinic is likely to grow in the future, especially as targeted probes for clinical use are under development. However, reliable deployment by the biologists depends upon robust inversion techniques and efficient incorporation of the anatomical priors to increase localization accuracy without introducing artifacts. Development and thorough evaluation of methods that take advantage of the specific deterministic or statistical properties of the underlying target signal, in conjunction with anatomical priors, are expected to further enhance performance and robustness of FMT. MSOT has for the first time enabled in vivo optical molecular imaging in deep tissue at high spatial resolution and localization accuracy. Such unique features are expected to offer new insights in preclinical molecular imaging studies. Of high significance for the deployment of the technology are the molecular imaging sensitivity and specificity that can be achieved, which both largely rely on performance of the used algorithms. Quantification is another long-standing open problem that is yet to be treated with robust solutions for structurally complex tissues.

ACKNOWLEDGMENTS

We thank Neal C. Burton and Christian Lutzweiler from our institute for their valuable comments.

AUTHORS

Pouyan Mohajerani (pouyan.mohajerani@gmail.com) received the B.Sc. degree in 2001 from Sharif University of Technology, Iran, the M.Sc. degree in 2004 from the Georgia Institute of Technology, United States, and the Ph.D. degree in 2014 from Technische Universität München, Germany, all in electrical engineering. He is currently a research fellow at the Institute for Biological and Medical Imaging at Helmholtz Zentrum München, Munich, Germany. His research interests are in signal and imaging processing, inverse problems, fluorescence molecular tomography, and frequency-domain optoacoustic tomography.

Stratis Tzoumas (strtzoumas@gmail.com) is a Ph.D. student at the Institute for Biological and Medical Imaging at Technische Universität München and Helmholtz Zentrum München, Munich, Germany. He studied electrical and computer engineering at the National Technical University of Athens, Greece. His current research interests are in molecular multispectral optoacoustic imaging, spectral unmixing, and inverse problems.

Amir Rosenthal (eeamir@gmail.com) received his B.Sc. and Ph.D. degrees, both from the Department of Electrical Engineering, the Technion–Israel Institute of Technology, Haifa, in 2002 and 2006, respectively. From 2009 to 2010, he was a research fellow at the Cardiovascular Research Center at Massachusetts General Hospital and Harvard Medical School, Boston. He is currently a faculty member and group leader at the Institute for Biological and Medical Imaging, Technische Universität München and Helmholtz Zentrum München, Germany. His research interests include optoacoustic imaging, interferometric sensing, intravascular imaging, inverse problems, and optical and acoustical modeling.

Vasilis Ntziachristos (v.ntziachristos@tum.de) studied electrical engineering at Aristotle University of Thessaloniki, Greece, and received his master's and doctorate degrees from the University of Pennsylvania, United States. He served as an assistant professor at Harvard University and Massachusetts General Hospital. He is currently a professor of medicine and electrical engineering and chair for biological imaging at Technische Universität München, Germany, and the director of the Institute for Biological and Medical Imaging at the Helmholtz Zentrum München. His research concentrates on basic research and translation of novel optical and optoacoustic in vivo imaging for addressing unmet biological and clinical needs. He has received several awards, including the 2013 Leibniz Prize and the 2011 Schrödinger Award. He is an author of over 230 peer-reviewed publications.

REFERENCES

[1] V. Ntziachristos, "Going deeper than microscopy: The optical imaging frontier in biology," *Nat. Methods*, vol. 7, no. 8, pp. 603–614, July 2010.

[2] J. Xia and L. Wang, "Small-animal whole-body photoacoustic tomography: A review," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1380–1389, May 2014.

[3] S. Zackrisson, S. van de Ven, and S. Gambhir, "Light in and sound out: Emerging translational strategies for photoacoustic imaging," *Cancer Res.*, vol. 74, no. 4, pp. 979–1004, 2014.

[4] V. Ntziachristos and D. Razansky, "Molecular imaging by means of multispectral optoacoustic tomography (MSOT)," *Chem. Rev.*, vol. 110, no. 5, pp. 2783–2794, Apr. 2010.

[5] F. Stuker, J. Ripoll, and M. Rudin, "Fluorescence molecular tomography: Principles and potential for pharmaceutical research," *Pharmaceutics*, vol. 3, no. 2, pp. 229–274, 2011.

[6] R. Weissleder and M. J. Pittet, "Imaging in the era of molecular oncology," *Nature*, vol. 452, no. 7187, pp. 580–589, 2008.

[7] S. Luo, E. Zhang, Y. Su, T. Cheng, and C. Shi, "A review of NIR dyes in cancer targeting and imaging," *Biomaterials*, vol. 32, no. 29, pp. 7127–7138, 2011.

[8] G. P. Luke, D. Yeager, and S. Y. Emelianov, "Biomedical applications of photoacoustic imaging with exogenous contrast agents," *Ann. Biomed. Eng.*, vol. 40, no. 2, pp. 422–437, 2012.

[9] A. Oraevsky and A. Karabutov, "Optoacoustic tomography," in *Biomedical Photonics Handbook*, T. Vo-Dinh, Ed. Boca Raton, FL: CRC Press, 2003, pp. 1–34.

[10] C. Li and L. V. Wang, "Photoacoustic tomography and sensing in biomedicine," *Phys. Med. Biol.*, vol. 54, no. 19, p. R59, 2009.

[11] S. Tzoumas, N. Deliolanis, S. Morscher, and V. Ntziachristos, "Unmixing molecular agents from absorbing tissue in multispectral optoacoustic tomography," *IEEE Trans. Med. Imaging*, vol. 33, no. 1, pp. 48–60, 2014.

[12] N. C. Deliolanis, A. Ale, S. Morscher, N. C. Burton, K. Schaefer, K. Radrich, D. Razansky, and V. Ntziachristo, "Deep-tissue reporter-gene imaging with fluorescence and optoacoustic tomography: A performance overview," *Mol. Imag. Bio.*, vol. 16, no. 5, pp. 652–660, Oct. 2014.

[13] R. B. Schulz, A. Ale, A. Sarantopoulos, M. Freyer, E. Soehngen, M. Zientkowska, and V. Ntziachristos, "Hybrid system for simultaneous fluorescence and x-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 29, pp. 465–473, Feb. 2010.

[14] J. Yao and L. V. Wang, "Sensitivity of photoacoustic microscopy," *Photoacoustics*, vol. 2, no. 2, pp. 87–101, pp. 87-101, June 2014.

[15] V. Ntziachristos, C. H. Tung, C. Bremer, and R. Weissleder, "Fluorescence molecular tomography resolves protease activity in vivo," *Nat. Med.*, vol. 8, no. 7, pp. 757–761, July 2002.

[16] C. Darne, Y. Lu, and E. M. Sevick-Muraca, "Small animal fluorescence and bioluminescence tomography: A review of approaches, algorithms and technology update," *Phys. Med. Biol.*, vol. 59, no. 1, p. R1, Jan. 2014.

[17] J. Ripoll, *Principles of Diffuse Light Propagation: Light Propagation in Tissues with Applications in Biology and Medicine*. Singapore: World Scientific, 2012.

[18] D. A. Boas, D. H. Brooks, E. L. Miller, C. A. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang, "Imaging the body with diffuse optical tomography," *IEEE Signal Processing Mag.*, vol. 18, no. 6, pp. 57–75, Nov. 2001.

[19] K. Wang and M. A. Anastasio, "Photoacoustic and thermoacoustic tomography: Image formation principles," in *Handbook of Mathematical Methods in Imaging*. New York: Springer, 2011, pp. 781–815.

[20] A. Rosenthal, V. Ntziachristos, and D. Razansky, "Acoustic inversion in optoacoustic tomography: A review," *Curr. Med. Imag. Rev.*, vol. 9, no. 4, pp. 318–336, Nov. 2013.

[21] L. V. Wang, *Photoacoustic Imaging and Spectroscopy*: Boca Raton, FL: CRC Press, 2009.

[22] P. Burgholzer, G. J. Matt, M. Haltmeier, and G. Paltauf, "Exact and approximative imaging methods for photoacoustic tomography using an arbitrary detection surface," *Phys. Rev. E*, vol. 75, no. 4, p. 046706, Apr. 2007.

[23] M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Phys. Rev. E*, vol. 71, no. 1, p. 016706, Jan. 2005.

[24] Y. Xu, L. V. Wang, G. Ambartsoumian, and P. Kuchment, "Reconstructions in limited-view thermoacoustic tomography," *Med. Phys.*, vol. 31, no. 4, pp. 724–733, Apr. 2004.

[25] K. Wang, S. A. Ermilov, R. Su, H.-P. Brecht, A. A. Oraevsky, and M. A. Anastasio, "An imaging model incorporating ultrasonic transducer properties for three-dimensional optoacoustic tomography," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 203–214, Feb. 2011.

[26] A. Rosenthal, V. Ntziachristos, and D. Razansky, "Model-based optoacoustic inversion with arbitrary-shape detectors," *Med. Phys.*, vol. 38, no. 7, pp. 4285–4295, July 2011.

[27] A. Rosenthal, D. Razansky, and V. Ntziachristos, "Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1275–1285, June 2010.

[28] J. Jose, R. G. H. Willemink, W. Steenbergen, C. H. Slump, T. G. van Leeuwen, and S. Manohar, "Speed-of-sound compensated photoacoustic tomography for accurate imaging," *Med. Phys.*, vol. 39, no. 12, pp. 7262–7271, Dec. 2012.

[29] K. Wang, R. Schoonover, R. Su, A. Oraevsky, and M. Anastasio, "Discrete imaging models for three-dimensional optoacoustic tomography using radially symmetric expansion functions," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1180–1193, May 2014.

[30] C. Lutzweiler, X. L. Deán-Ben, and D. Razansky, "Expediting model-based optoacoustic reconstructions with tomographic symmetries," *Med. Phys.*, vol. 41, no. 1, pp. 013302-1–013302-10, Jan. 2014.

[31] K. Wang, C. Huang, Y.-J. Kao, C.-Y. Chou, A. A. Oraevsky, and M. A. Anastasio, "Accelerating image reconstruction in three-dimensional optoacoustic tomography on graphics processing units," *Med. Phys.*, vol. 40, no. 2, p. 023301, Feb. 2013.

[32] A. Rosenthal, T. Jetzfellner, D. Razansky, and V. Ntziachristos, "Efficient framework for model-based tomographic image reconstruction using wavelet packets," *IEEE Trans. Med. Imag.*, vol. 31, no. 7, pp. 1346–1357, July 2012.

[33] P. Mohajerani, A. Hipp, M. Willner, M. Marschner, M. Trajkovic-Arsic, X. Ma, N. C. Burton, U. Klemm, K. Radrich, V. Ermolayev, S. Tzoumas, J. T. Siveke, M. Bech, F. Pfeiffer, and V. Ntziachristos, "FMT-PCCT: Hybrid fluorescence molecular tomography—X-ray phase-contrast CT imaging of mouse models," *IEEE Trans. Med. Imag.*, vol. 33, no. 7, pp. 1434–1446, July 2005.

[34] N. Ducros, A. Bassi, G. Valentini, G. Canti, S. Arridge, and C. D'Andrea, "Fluorescence molecular tomography of an animal model using structured light rotating view acquisition," *J. Biomed. Opt.*, vol. 18, no. 2, p. 20503, Feb. 2013.

[35] A. Soubret, J. Ripoll, and V. Ntziachristos, "Accuracy of fluorescent tomography in the presence of heterogeneities: Study of the normalized Born ratio," *IEEE Trans. Med. Imag.*, vol. 24, no. 10, pp. 1377–1386, Oct. 2005.

[36] D. Hyde, E. Miller, D. H. Brooks, and V. Ntziachristos, "A statistical approach to inverting the Born ratio," *IEEE Trans. Med. Imag.*, vol. 26, no. 7, pp. 893–905, July 2007.

[37] D. Hyde, R. Schulz, D. Brooks, E. Miller, and V. Ntziachristos, "Performance dependence of hybrid x-ray computed tomography/fluorescence molecular tomography on the optical forward problem," *JOSA A*, vol. 26, no. 4, pp. 919–923, Apr. 2009.

[38] T. Pyka, R. Schulz, A. Ale, and V. Ntziachristos, "Revisiting the normalized Born approximation: Effects of scattering," *Opt. Lett.*, vol. 36, pp. 4329–4331, Nov. 2011.

[39] D. Hyde, E. L. Miller, D. H. Brooks, and V. Ntziachristos, "Data specific spatially varying regularization for multimodal fluorescence molecular tomography," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 365–374, Feb. 2010.

[40] S. C. Davis, H. Dehghani, J. Wang, S. Jiang, B. W. Pogue, and K. D. Paulsen, "Image-guided diffuse optical fluorescence tomography implemented with Laplacian-type regularization," *Opt. Express*, vol. 15, no. 7, pp. 4066–4082, Apr. 2007.

[41] J. C. Baritaux, K. Hassler, and M. Unser, "An efficient numerical method for general L(p) regularization in fluorescence molecular tomography," *IEEE Trans. Med. Imag.*, vol. 29, no. 4, pp. 1075–1087, Apr. 2010.

[42] V. Pera, E. Zettergren, D. H. Brooks, and M. Niedre, "Maximum likelihood tomographic reconstruction of extremely sparse solutions in diffuse fluorescence flow cytometry," *Opt. Lett.*, vol. 38, no. 13, pp. 2357–2359, July 2013.

[43] T. Correia, J. Aguirre, A. Sisniega, J. Chamorro-Servent, J. Abascal, J. J. Vaquero, M. Desco, V. Kolehmainen, and S. Arridge, "Split operator method for fluorescence diffuse optical tomography using anisotropic diffusion regularisation with prior anatomical information," *Biomed. Opt. Exp.*, vol. 2, no. 9, pp. 2632–2648, Sept. 2011.

[44] A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane, "Fluorescence optical diffusion tomography," *Appl. Opt.*, vol. 42, no. 16, pp. 3081–3094, June 2003.

[45] J. H. Lee, A. Joshi, and E. M. Sevick-Muraca, "Fully adaptive finite element based tomography using tetrahedral dual-meshing for fluorescence enhanced optical imaging in tissue," *Opt. Express*, vol. 15, no. 6, pp. 6955–6975, June 2007.

[46] P. Mohajerani and V. Ntziachristos, "Compression of Born ratio for fluorescence molecular tomography/x-ray computed tomography hybrid imaging: Methodology and in vivo validation," *Opt. Lett.*, vol. 38, no. 13, pp. 2324–2326, July 2013.

[47] J. Chen, V. Venugopal, and X. Intes, "Monte Carlo based method for fluorescence tomographic imaging with lifetime multiplexing using time gates," *Biomed. Opt. Exp.*, vol. 2, no. 4, pp. 871–886, Apr. 2011.

[48] B. Cox, J. G. Laufer, S. R. Arridge, and P. C. Beard, "Quantitative spectroscopic photoacoustic imaging: A review," *J. Biomed. Opt.*, vol. 17, no. 6, pp. 0612021–06120222, June 2012.

[49] B. Cox, S. Arridge, K. Kostli, and P. Beard, "Quantitative photoacoustic imaging: fitting a model of light transport to the initial pressure distribution," in *Proc. Biomedical Optics 2005*, pp. 49–55.

[50] Z. Yuan, Q. Wang, and H. Jiang, "Reconstruction of optical absorption coefficient maps of heterogeneous media by photoacoustic tomography coupled with diffusion equation based regularized Newton method," *Opt. Express*, vol. 15, no. 26, pp. 18076–18081, Dec. 2007.

[51] A. Rosenthal, D. Razansky, and V. Ntziachristos, "Quantitative optoacoustic signal extraction using sparse signal representation," *IEEE Trans. Med. Imag.*, vol. 28, no. 12, pp. 1997–2006, Dec. 2009.

[52] B. Cox, S. Arridge, and P. Beard, "Estimating chromophore distributions from multiwavelength photoacoustic images," *JOSA A*, vol. 26, no. 2, pp. 443–455, Feb. 2009.

[53] J. Laufer, B. Cox, E. Zhang, and P. Beard, "Quantitative determination of chromophore concentrations from 2D photoacoustic images using a nonlinear model-based inversion scheme," *Appl. Opt.*, vol. 49, no. 8, pp. 1219–1233, Mar. 2010.

[54] M. Li, J. Oh, X. Xie, G. Ku, W. Wang, C. Li, G. Lungu, G. Stoica, and L.-V. Wang, "Simultaneous molecular and hypoxia imaging of brain tumors in vivo using spectroscopic photoacoustic tomography," *IEEE Proc.*, vol. 96, no. 3, p. 481, Mar. 2008.

[55] J. Glatz, N. C. Deliolanis, A. Buehler, D. Razansky, and V. Ntziachristos, "Blind source unmixing in multi-spectral optoacoustic tomography," *Opt. Express*, vol. 19, no. 4, pp. 3175–3184, Feb. 2011.

[SP]

Margaret H. Pinson, Lucjan Janowski, and Zdzisław Papir

VIDEO QUALITY ASSESSMENT

Subjective testing of entertainment scenes

TECHNOLOGY PHOTO—© ISTOCK.COM/VIOLETKAIPA
FILM PHOTO—© ISTOCK.COM/F9PHOTOS

This article describes how to perform a video quality subjective test. For companies, these tests can greatly facilitate video product development; for universities, removing perceived barriers to conducting such tests allows expanded research opportunities. This tutorial assumes no prior knowledge and focuses on proven techniques. (Certain commercial equipment, materials, and/or programs are identified in this article to adequately specify the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Telecommunications and Information Administration, nor does it imply that the program or equipment identified is necessarily the best available for this application.)

Video is a booming industry: content is embedded on many Web sites, delivered over the Internet, and streamed to mobile devices. Cisco statistics indicate that video exceeded 50% of total mobile traffic globally or the first time in 2012 and predict that over two-thirds of the world's mobile data traffic will be video by 2018 [1]. Each company must make a strategic decision on the correct balance between delivery cost and user experience. This decision can be made by the engineers designing the service or, for increased accuracy, by consulting users [2].

Digital Object Identifier 10.1109/MSP.2013.2292535
Date of publication: 5 December 2014

Video quality assessment requires a combined approach that includes objective metrics, subjective testing, and live video monitoring. Carefully conducted video quality subjective tests are extremely reliable and repeatable, as is shown in [3, Sec. 8]. This article provides an approachable tutorial on how to conduct a subjective video quality experiment. Our goal is to encourage more companies and universities to perform subjective tests.

A subjective video quality test uses a small set of short video sequences (e.g., 8–20 s) to measure people's opinions of the quality of different video processing options. These tests focus on people's current opinion, as opposed, e.g., to opinions of an entire movie. The goal is to make an impartial judgment about opinion trends. Example applications include choosing between different coding algorithms, comparing one coder at different bit rates, comparing two implementations of the same algorithm, optimizing coder parameters, improving an error concealment algorithm, or selecting a maximum packet loss rate for a service. Video quality subjective tests isolate one factor: video quality. Issues that might confound the experiment data should be excluded [e.g., audio, scene composition, aesthetics, display, environment, device interface, two-way communication, and quality of experience (QoE)].

INTERNATIONAL TELECOMMUNICATIONS UNION RECOMMENDATIONS

The International Telecommunications Union (ITU) recommendations most directly applicable to this tutorial are ITU-R Rec. BT.500 (2012), *Methodology for the Subjective Assessment of the Quality of Television Pictures*; ITU-T Rec. P.910 (2008), *Subjective Video Quality Assessment Methods for Multimedia Applications*; and ITU-R Rec. BT.1788 (2007), *Subjective Assessment of Multiple Video Quality (SAMVIQ)*. ITU-R Rec. BT.500 focuses on video quality and image quality in a home television environment; ITU-T Rec. P.910 focuses on video quality, videotelephony, videoconferencing, and storage/retrieval applications; and ITU-R Rec. BT.1788 identifies one particular rating method. The current version of each recommendation is distributed freely on the ITU Web site (http://www.itu.int/). These procedures remove all distractions from the environment to eliminate variables that might bias the test. The environment is basically an idealized living room: quiet and devoted to this one task. These ITU recommendations assume the reader has some prior knowledge.

The scope of ITU-R Rec. BT.500 is broadcast television and, therefore, entertainment video in either standard-definition or high-definition format. BT.500 specifies highly controlled monitor calibration and lighting conditions (e.g., the ratio of luminance of inactive screen to peak luminance should be $\leq$ 0.02). The monitor calibration techniques focus on the needs of broadcasters, so an amateur may have difficulty calibrating consumer-grade equipment.

ITU-T Rec. P.910 was designed for video systems at lower bit rates and quality than broadcast television. The wording of P.910's focus may look odd today because terminology has changed. P.910 is appropriate for high-definition television (HDTV) through quarter common intermediate format (QCIF) resolution (176×144). P.910 specifies exact lighting conditions, but they are easier to recreate than BT.500's lighting conditions. The monitor is not calibrated, which is more appropriate for computers, mobile devices, and consumer-grade televisions. ITU-R Rec. BS.1788 is commonly referred to by the acronym of its title: SAMVIQ. This recommendation defines a particular rating scale and method.

Work is underway in the ITU to develop recommendations better suited to new technologies. One example is the newly approved ITU-T Rec. 913 (2014) Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment. This recommendation describes techniques for situations not covered by BT.500 and P.910, including the use of natural lighting or a distracting environment (e.g., a cafeteria or bus). Published at the end of 2014, [51] contains a summary of the differences between the traditional techniques found in this article and P.913.

EXPERIMENT DESIGN

TERMS AND DEFINITIONS

- *Source sequence* (SRC) is the unimpaired video sequence (i.e., the content).
- *Original* refers to the original version of each SRC (e.g., broadcast quality).
- *Processed video sequence* (*PVS*) is the impaired version of a video sequence.
- *Clip* refers to any video sequence, SRC or PVS.
- *Hypothetical reference circuit* (HRC) is a fixed combination of a video encoder operating at a given bit rate, network condition, and video decoder. The abbreviation HRC is preferred when vendor names should not be identified.
- *Full matrix design* consists of n SRCs and m HRCs. All combinations of SRCs and HRCs are included in the experiment for a total of $(n \times m)$ PVSs.
- *Partial matrix design* splits the experiment into two or more smaller matrixes. For example, a two-matrix experiment would have two scene pools (pool A and pool B, with n_A and n_B SRCs, respectively) and two HRC pools (pool A and pool B, with m_A and m_B HRCs, respectively). All combinations of pool A SRC and HRCs are included, plus all combinations of pool B SRC and HRCs, for a total of $(n_A \times m_A + n_B \times m_B)$ PVSs.

GOAL OF EXPERIMENT AND DESIGN CONSEQUENCES

The first goal of a video quality subjective test is to answer a specific question about video encoding, transmission, or decoding. These questions are typically posed as comparisons between one or more variables. The analysis will directly compare pairs of HRCs using identical SRCs, typically via a full matrix design. For example, Younkin and Corriveau [4] use a full matrix design to analyze the impact of playback error severity on quality perception.

The full matrix design allows all HRCs to be directly compared and produces improved accuracy for some analysis techniques (see the section "Choosing a Subjective Scale"). The disadvantage is that less information is obtained about the impact of different source material on the HRCs. This is undesirable because codecs yield very different quality depending upon the scene content, as

can be seen in Figure 1, taken from [5]. The x-axis displays the bit rate, and the y-axis displays the mean opinion score (MOS). The boxes and whiskers in Figure 1 show the distribution of PVSs within an HRC. Some of the plotted HRCs span more than half of the absolute category rating (ACR) scale.

Twice as many SRCs can be included in a partial matrix design of two matrices, compared to a full matrix design. This alleviates the subjects' boredom. The partial matrix design allows direct comparisons only of HRCs within an HRC pool. For example, HRCs from pool *A* cannot be directly compared with HRCs from pool *B*. Pinson et al. [5] use a partial matrix design to compare the quality of the H.264 and MPEG-2 coders, both with and without packet loss.

Both full matrix and partial matrix designs depend upon two variables: SRC and HRC. A third variable—the environment—is needed to answer questions about interactions between the video signal and the viewing environment. For example, Brunnström et al. [6] explore the relationship between video quality and the viewing angle of the subject to the screen. Thus, Brunnström's HRC definitions specified the viewing angle.

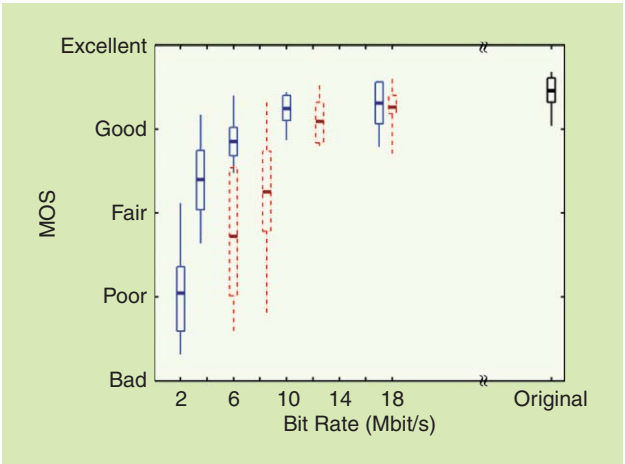
The second video quality subjective test goal is to train a metric or algorithm. For example, an objective video quality metric estimates the quality ratings that would result from a subjective experiment. The accuracy of the resulting metric depends upon the quantity and variety of training data. Thus, the optimal experiment design maximizes the number of SRCs and HRCs for the available number of PVSs. A random pairing of each SRC and a different HRC will accomplish this goal, though most engineers are troubled by the asymmetry. Voran and Wolf [7] provide an example of a subjective experiment designed specifically to train a metric. A full matrix or the partial matrix design is usually less effective, because fewer SRC and HRC can be analyzed. Some experimenters choose a full or partial matrix design anyway, because they want to use the same subjective test for two purposes: to answer a question and to train a metric. Huynh-Thu and Ghanbari [8] provide an example.

The third video quality subjective test goal is to analyze the performance of an existing objective video quality metric or algorithm. The constraint here is not the design of the test—both the full matrix and the partial matrix designs are suitable—but rather the fact that training data cannot be used to test the model's performance. Ideally, this prohibition includes scene content, coder implementations, and coder/network settings (e.g., packet loss rate, and bit rate). Voran and Catellier [9] provide an example of how to design a subjective test to both train and test a metric. This article describes a speech quality experiment; however, the experimental design issues are the same.

Video quality subjective tests can be used in combination with other subjective tests to understand larger quality implications. ITU-T Rec. P.1301 [10] demonstrates this idea for telemeeting systems, and an applied example can be found in [11].

SRC, HRC, AND PVS SELECTION

Video quality subjective tests typically use short sequences (e.g., 8–10 s duration). Pinson et al. [12] provide guidance on choosing a balanced and well-designed set of scenes for a subjective test.



[FIG1] For coding-only impairments, a quality comparison of H.264 (solid blue) and MPEG-2 (dotted red): the box-plot identifies minimum, 25%, mean, 75%, and maximum MOS.

This guidance includes avoiding offensive content, choosing scenes that evenly span a wide range of coding difficulty, deciding whether or not scene cuts should be allowed, and selecting scenes with unusual properties. It is important to use high-quality footage because otherwise the quality impairments in the SRC can obscure any effects of the HRC in the test results. Niu and Liu [13] explain the differences between professional and amateur videography and provide objective criteria for identifying professional video sequences. Amateur footage typically contains aesthetic problems that trigger low video quality ratings (e.g., focus control, color palette, camera motion, shot length, and visual continuity). The Consumer Digital Video Library (www.cdvl.org) provides free downloads of broadcast-quality footage for research and development purposes. Another Web site that offers free footage is http://www.irccyn.ec-nantes.fr/spip.php?article541.

The range of PVS quality should span the scale used to conduct the test (see the section "Choosing a Subjective Scale"). Experiments that contain a narrow range of quality will be frustrating for the subjects and researcher alike since the data are unlikely to show any significant results. It is better to design experiments that span a wide range of quality—or at least a wide enough range that meaningful results can be found.

The goal of many subjective experiments is to compare and contrast the quality of various video impairments. This analysis is only possible when HRC creation is limited by two constraints: 1) the definition of each HRC is constant throughout the experiment and 2) if two HRCs are to be compared, then those HRCs must be paired with the same set of SRCs.

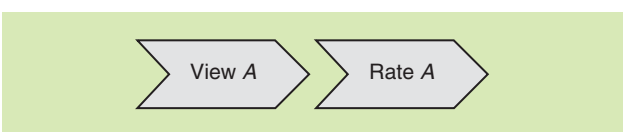
The term HRC implies that all PVSs associated with that HRC were created using a constant set of control parameters. A particularly popular HRC definition specifies codec *A*, profile *B*, constant bit rate *C*, and packet loss rate *D*. As a side effect of this design, difficult-to-encode SRCs will yield a very wide range of quality (excellent to bad), while easy-to-encode SRCs will yield a narrow range of quality (excellent to fair). An alternative approach tries to produce equivalent quality for all PVSs

associated with a single HRC. This can be done with variable bit rate encoding or a constant quantization profile value. The problem is that it becomes difficult to reach conclusions about the coder's behavior at different bit rates.

TEST ENVIRONMENT

Traditionally, subjective video quality tests are performed in a controlled laboratory environment. This reduces the effect of extraneous variables on the experiment without requiring a specialized space or great expense. While the potential impact of some elements is debatable, the traditional controlled environment demonstrates your expertise to the research community.

- *Walls*: The walls of the test chamber should be plain white and not show potentially distracting objects (e.g., pictures, clock, and wires). Windows must be covered with light-blocking curtains. Temporary room dividers encourage the illusion of a nondistracting chamber.
- *Floor*: The floor should be a neutral, nondistracting color. Solid gray is traditional.
- *Furniture*: Only necessary furniture should be in the test chamber. The chair provided to subjects should not have wheels. This will encourage the subject to keep a constant viewing distance throughout the test. An upright chair helps to encourage attention on the task.
- *Lighting*: See ITU-R Rec. BT.500 clause 2.1 or ITU-T Rec. P.910 clause 7.1 for lighting conditions. The listed specifications can be met inexpensively using a light meter, full spectrum bulbs, and variable intensity lamp controls.
- *Viewing distance*: See ITU-R Rec. BT.500 clause 2.1 or ITU-T Rec. P.910 clause 7.1 for details. For most experiments, the monitor and chair should be positioned at a defined viewing distance. The viewing distance is traditionally measured in picture heights: four to six times picture heights (H) for standard definition television (i.e., 4H to 6H), 2H to 3H for HDTV, and 8H for smartphones and other very small monitors [14].
- *Monitor*: BT.500 encourages the use of a professional quality monitor to eliminate a potentially confounding variable. For P.910, choose a monitor that matches the application.
- *Background noise*: The test chamber must be quiet, with minimum background noise. If a computer is used to play the videos, the computer should be outside the test chamber.
- *Bystanders*: While a subject is running through the test, the chamber should be used for no other purpose. In some cases, the test chamber will have two or more subjects and the experimenter. People who are interested in seeing the test results come out a certain way should not interact with the subjects, perform the data analysis, or design the test (e.g., product managers).



[FIG2] **The ACR rating cycle: the subject watches video clip *A* and then rates *A*.**

NUMBER OF SUBJECTS, STIMULI, AND TEST SESSIONS

The reliability of ratings depends upon averaging the data across multiple subjects. While BT.500 recommends a minimum of 15 subjects, a recent study by Pinson et al. [15] endorses a minimum of 24 subjects. Fewer subjects may be used to indicate trending.

Subjects have a limited attention span, and so the typical challenge is fitting all impairments of interest into a set of test sessions that one person can reasonably watch. Preferably, each session should last no more than 20 min, and each subject should spend no more than 1 h rating video. Longer experiments require additional motivation or variety to keep the subjects alert. Payment is the traditional motivator. The best way to add variety into an experiment is to increase the number of SRCs.

Subjective experiments should use at least eight different SRCs. Differences between SRCs are a major variable for every subjective experiment. A large and varied pool of scenes minimizes the risk that the subjective experiment will reach an erroneous conclusion. This effect is demonstrated in Pinson et al. [12]. Whenever possible, we advocate the use of the partial matrix test design over a full matrix design. Each full matrix within the partial matrix test is associated with a different set of SRCs and HRCs. This adds much needed visual interest for the subject. HRCs that need to be directly compared should be put within the same matrix.

CHOOSING A SUBJECTIVE SCALE

The choice of subjective scale is surprisingly contentious. Each scale has strengths and weaknesses. Choose the scale that best matches your goal. There different scales for single stimulus (SS) and double stimulus (DS) experiments. In an SS test, the subject watches and rates each video sequence separately. In a DS test, the subject watches two or more versions of the same source video sequence during the rating process.

LISTING OF SUBJECTIVE SCALES

For the ACR from ITU-T Rec. P.910, the subject watches a video sequence and then is asked to rate it on a discrete, five-level scale (see Figure 2). Each level is associated with a word and a number: excellent = 5, good = 4, fair = 3, poor = 2, and bad = 1. Variations of the scale include nine levels, 11 levels, and a continuous scale with labels only at the end points. A continuous scale is a continuous line when presented to the subject and converted to a 100-level scale for the purposes of data analysis. Alternative labels may be needed for some experiments. ITU-T Rec. P.800, a speech quality subjective testing standard, provides alternate ACR wording examples for a listening-effort scale, and a loudness-preference scale. ACR is an SS method.

ITU-T Rec. P.910 identifies a variant, ACR with hidden reference (ACR-HR). In ACR-HR, each original is included in the experiment but not identified as such. The ratings for the originals are removed from the scores of the associated PVSs during data processing. High-quality originals are critical when using ACR-HR. If the quality of the original SRC drops from excellent (score 5) to fair (score 3), the available ACR-HR scale decreases from four to two units. This causes an inherent bias in the data, when comparing PVSs associated with two different SRCs.

The degradation category rating (DCR) method from P.910 also appeared in an old version of BT.500 under the name DS impairment scale (DSIS). DCR presents stimuli to subjects in pairs (see Figure 3). The original is presented first, and the subject is told that this is the original. The stimulus to be rated is presented second. The subject rates the difference in quality on a discrete, five-level, impairment scale: imperceptible = 5, perceptible but not annoying = 4, slightly annoying = 3, annoying = 2, and very annoying = 1. DCR is a DS method.

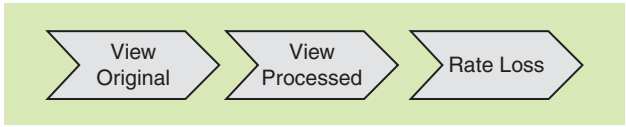
The pair comparison method from BT.500 also appears in P.800 under the name *comparison category rating* and in an old version of BT.500 under the name *double stimulus comparison scale* (*DSCS*). A pair of stimuli is presented to the subject; however, the order of stimuli is random (see Figure 4). If pair comparison is used to compare original and processed sequences (like DCR), then the original would be played first for approximately half of the trials, and the PVS would be played first for the rest of the trials. Pair comparison is the only method that can directly compare two different impaired versions of the same video sequence.

The subjects rate the quality of the second stimulus compared to the quality of the first on a discrete, seven-level scale: much better = 3, better = 2, slightly better = 1, about the same = 0, slightly worse = –1, worse = –2, and much worse = –3. Variations include a continuous scale (100 levels) and a discrete, two-level preference (better or worse).

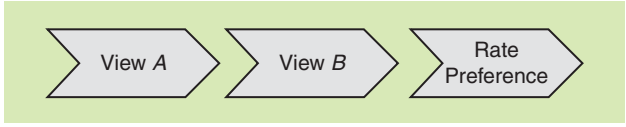
The double stimulus continuous quality scale (DSCQS) method from BT.500 involves four presentations of two stimuli, A and B (see Figure 5). One of these is the original, assigned randomly to position A or B. The subject is presented with stimulus A, then B, then A again, and then B again. Afterward, the subject rates A and B separately, each on a continuous scale showing the ACR labels (excellent, good, fair, poor, or bad).

The SS continuous quality evaluation (SSCQE) method from BT.500 presents the subject with a stimulus of long duration (e.g., 5–30 min). The subject has a slider that is constantly moved to reflect the subject's current opinion of the video quality (see Figure 6). Ratings are sampled every half second. SSCQE was intended for the analysis of monitoring applications and uses a continuous scale.

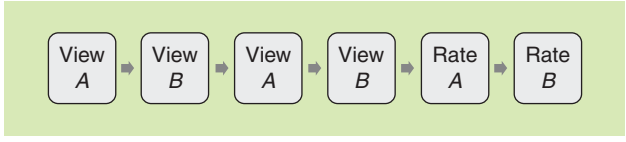
The SAMVIQ method from ITU-R Rec. BT.1788 uses a continuous scale, marked with the ACR labels. The test uses a computer interface, which presents the subject with multiple versions of the same SRC (see Figure 7). The subject may play each stimulus multiple times and may choose the order in which stimuli are rated. One of the stimuli is the original and explicitly labeled as such. Another stimulus is a hidden reference—identical to the original, but not labeled. The subject rates each version of one SRC and adjusts the ratings relative to each other.



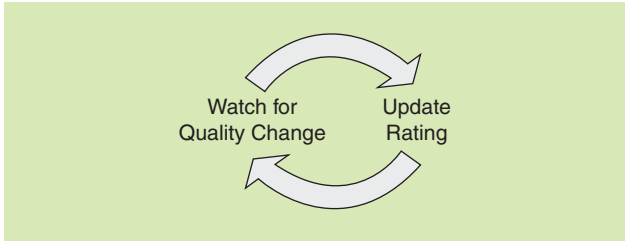
[FIG3] **The DCR rating cycle: the subject watches the original video, then watches a processed version of that video, and finally rates the level of impairment.**



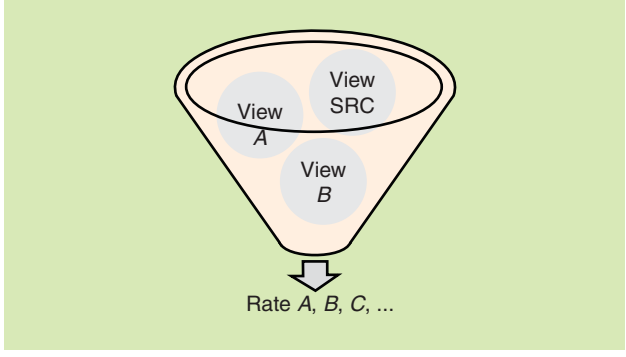
[FIG4] **Pair comparison rating cycle: the subject watches video sequence *A*, then watches video sequence *B*, and rates *B* relative to *A*.**



[FIG5] **The DSCQS rating cycle: the subject watches video clip *A*, then *B*, *A* again, *B* again, and then rates *A* and *B*.**



[FIG6] **The SSCQE rating cycle: watch a long video sequence and continuously update a slider to reflect current opinion of the video quality.**



[FIG7] **The SAMVIQ rating cycle: the subject watches several versions of one SRC in any order. The ratings are adjusted until the subject is satisfied.**

ANALYSIS OF SINGLE STIMULUS RATING METHODS

ACR with a five-level scale maximizes cognitive ease and the number of video sequences rated each minute [16]. ACR produces very repeatable subjective results, even across different groups of subjects, provided that the test design and instructions are carefully prepared [17]. Studies [16]–[18] compared ACR ratings with ratings gathered from DSCQS, DCR, and SAMVIQ. The rating scale choice had a minor impact on data accuracy.

The SS methods (ACR and SSCQE) have two weaknesses. The first is that some types of impairments are difficult to detect

without an explicit comparison between two video sequences. For example, small color shifts are difficult to detect using ACR tests.

The second weakness is that SS ratings do not differentiate between impairments in the SRC and impairments intentionally added to the processed sequence. SS scores are often biased by subjects' opinions of the scene's aesthetics, composition, production quality, and subject matter, despite instructions to the contrary. ACR-HR offsets this flaw by removing the original video's rating during data analysis. Thus, ACR-HR ratings look more like DS ratings (i.e., a score of "5" on the ACR-HR scale means that the original video and PVS have identical quality). This technique only works better with high-quality SRCs.

SSCQE has strengths and weaknesses similar to ACR. SSCQE has the potential for allowing the most evaluations from a single subject in a short time, because there are no pauses between stimuli for ratings. Pinson and Wolf [19] demonstrated that SSCQE can be as accurate as DSCQS and pair comparison for rating short video sequences by using multiple randomizations, hidden reference removal, and the SSCQE score at the end of each sequence. A variety of devices have been used to implement an SSCQE slider, including a game station steering wheel [20] and a sensory glove [21]. True SSCQE data analysis is complex, as it requires time series analysis.

ANALYSIS OF DOUBLE STIMULUS RATING METHODS

DS methods address both problems of the SS method, at least to some extent. DS methods ask one of two basic questions.

The first DS question is "Which of these two sequences do you like better?" Pair comparison is the only DS method that directly answers this question. Inversion errors can occur but seem to be rare (e.g., the subject marks "the second sequence is much better" when they intended to mark "the second sequence is much worse"). Pair comparison takes approximately twice as long as ACR for the same number of stimuli.

Pair comparison is the obvious choice for detecting very small differences between two different impaired versions of the same video sequence. Doherty et al. [22] demonstrate the use of pair comparison to detect differences between frame rate conversion algorithms.

The second DS question is "How well does the impaired sequence reproduce the reference sequence?" DCR answers this question explicitly. DSCQS, SAMVIQ, and ACR-HR answer this question implicitly.

DCR takes approximately twice as long as ACR for the same number of stimuli. This is as fast as any DS method can claim, and inversion errors do not occur. Tominaga [16] concludes that DCR is more desirable than DSCQS or SAMVIQ, because of improved speed and ease of use, without loss of accuracy. DCR cannot be used to measure quality improvements—the rating scale does not allow a subject to say that the PVS is of higher quality than the source. DCR is the only method where subjects are unambiguously instructed to rate the perceptual difference between an original sequence and an impaired version of that sequence.

A quirk of DCR is that the original video will not be scored as perfect. That is, if the original video is played identically as both the "original" and "processed" video in Figure 3, the rating will be slightly lower than five. This imposes a systematic downward shift on all scores that bothers some researchers. (Pair comparison is the only subjective method that is likely to yield perfect scores for original video sequences.) This bias causes no problems for the data analysis.

DSCQS takes approximately four times as long as ACR for the same number of stimuli. The repeated viewings of stimuli A and B are intended to yield improved accuracy per subject for small quality differences, but this has not been proven.

Inversion errors are a problem for DSCQS. Inversion errors are impossible to detect and remove from the data, and this is perhaps the reason why DSCQS failed to show improved accuracy in [16] and [18]. It should be possible to avoid DSCQS inversion errors using an automated subjective testing system that swaps the order of the fourth and fifth steps in Figure 5.

SAMVIQ takes approximately twice as long as ACR for the same number of stimuli. SAMVIQ is slowed down by the ability of subjects to repeatedly play and compare sequences, yet sped up by presenting all versions of each SRC to the subject simultaneously. SAMVIQ is the only method that allows subjects to directly compare multiple versions of a single SRC.

The advantage of SAMVIQ is improved accuracy. SAMVIQ with 15 subjects is as precise as ACR with more than 22 subjects [17]. An open question is whether or not SAMVIQ's improved accuracy per subject yields an additional advantage for an experiment that focuses on a narrow range of quality.

DISCRETE VERSUS CONTINUOUS

Discrete levels are used for ACR, DCR, and pair comparison. Continuous scales are used for SSCQE and DSCQS. Researchers have explored continuous scales and different numbers of discrete levels for ACR and pair comparison. Tominaga et al. [16] showed that a five-level discrete scale provides a much easier cognitive task for the subject than an 11-level discrete scale or a continuous scale. Studies [16]–[18] demonstrate that the continuous scales do not improve measurement accuracy. This makes sense; research on human thought indicates that people can only hold about seven items in immediate memory [23]. (The Harvard Mind Brain Behavior Event Video Archive provides a nice summary of [23] at minute 8:30 of the video "The Cognitive Revolution at Fifty Plus or Minus One: A Conversation with Jerome Bruner, Susan Carey, Noam Chomsky, and George Miller—Part 1.")

When using a discrete scale, researchers disagree on whether or not the level numbers should be displayed to the subject. No consensus exists on this subject.

IMPLEMENTATION

PLAYING BROADCAST-QUALITY VIDEO ON A TELEVISION

The ideal video playback system plays uncompressed video flawlessly. Why use uncompressed video? With uncompressed video and perfect rendering come the guarantee that the playback system does not add new impairments to the video.

This ideal television playback/capture system costs about US$10,000 excluding the monitor, a cost that has been fairly consistent for the past decade. The components consist of a multiple core computer, a redundant array of independent disks (RAID), and a specialized board to play video from the RAID to the serial digital interface (SDI) high-definition SDI (HD-SDI) or high-definition multimedia interface (HDMI). The following companies currently produce professional grade video capture and playback cards that are compatible with professional editing suites: AJA, Bluefish444, Blackmagic, and Matrox. Bit-perfect playback and capture must be proven, which requires two systems (i.e., system 1 plays the video, system 2 captures it, and then a pixel-by-pixel comparison is performed). Common problems include insufficient RAID speed, operating system interruptions, antivirus software interruptions, and driver incompatibilities.

The alternative is to use dedicated hardware. There are too many professional-grade devices available to list in this article. Most of these devices compress the video slightly (e.g., four to ten times). Professional video devices ensure reliable video play and record capability, usually with no perceptual impairment.

PLAYING VIDEO ON A COMPUTER MONITOR

The ideal video playback system plays uncompressed, progressive video flawlessly from a computer hard drive to its monitor. If compressed playback is acceptable, the computer setup is simplified and the price drops. For some devices, only compressed video playback is easily available (e.g., smartphones), and reliable playback requires substantial compression (e.g., 30–250 times). Any added impairment from the coder, decoder, or display will confound the research data and may cause the data analysis to be misleading. To avoid this problem, identically compress all video sequences for the purposes of playback only. That is, after the videos have been impaired as specified in the experiment design, re-encode all videos at the same (higher) bit rate for playback purposes. The goal is that any added playback impairments will be imposed identically on all videos.

When using a computer video playback system, you must calculate the appropriate level of compression yourself. The highest bit rate that guarantees flawless playback will minimize the perceptual impact of the recompression. To find this bit rate, encode a large variety of high-quality SRC, and play them to the target display repeatedly, while looking for playback problems (e.g., intermittent pauses and reduced frame rate).

THE TEST SESSIONS: AUTOMATED, EDITED, OR MANUAL

There are three options for playing video and recording scores during the actual subjective test: automated playback and scoring, edited sessions, and manual sessions.

When playing uncompressed or lightly compressed progressive video to a computer monitor, automated software provides an elegant solution for subjective testing. The software should identify subjects by ID number (see the section "Conducting the Experiment"), generate a unique randomized order of sequence presentation for each subject, implement the chosen method's rating cycle, ensure flawless playback for all subjects, present the rating scale after video playback finishes, record scores to a file, run each session separately, prompt for breaks, remove visual clutter from the screen, and either allow or disallow video replay. Three freely available software packages are AcrVQWin [24], Tally [25], and SubjectivePlayer [26].

If the subjects are allowed to replay the video, some subjects will, and this impacts the ratings. No consensus exists on the advisability of the replay option. Allowing replay provides the subject with an option other than guessing when their attention wandered—but conflicts arise with the usage paradigm for some video systems where rewind is not available. Thus, any article that describes a subjective test must specify whether or not videos could be replayed.

When subjective video quality testing began, the only option available was editing test tapes. Edited sessions work as well today as they did then, though the playback system is likely to be DVD, Blu-ray, three-dimensional (3-D) Blu-ray, or simply a long video file. No specialized software is required, and equipment costs are minimal. DVD or Blu-ray ensures consistent playback quality.

The concept is to edit together a long video sequence for each session. For example, when conducting an ACR test with 10-s sequences, the sequence would alternate between playing a 10-s sequence and playing 8 s of midlevel gray while the subject scores. This editing is simple yet prone to errors. A minimum of two different sequence orderings must be created to minimize the impact of ordering effects (i.e., the quality of clip N influences the perceived quality of clip $N+1$). Order effects can be reduced by randomizing the sessions (e.g., one subject sees session A, B, and C; another subject sees session C, B, and A). Ratings are entered either on a paper score sheet or on a small mobile device. Unlike an automated test, the ratings are not synchronized with the video playback. The audio track and text overlays keep the subject synchronized (e.g., please score clip 1). Subjects will sometimes make a mistake, and get off by one in their scores (e.g., record the quality of clip 9 where they were supposed to score clip 8). Data entry errors can occur when copying paper rating sheets into a spreadsheet.

The last option is a fully manual experiment. The experimenter can manually play each sequence in the desired order, ask the subject to choose a rating aloud, and record that rating themselves. This approach seems inelegant and the experimenter's behavior could influence ratings. It is, however, quite inexpensive and very practical.

We recommend playing midlevel gray between video sequences (i.e., Y = 128, Cb = 0, Cr = 0). The following MATLAB code will create this JPEG image, for video graphics array (480×640):

```
imwrite(zeros(480,640,'uint8') + 128,
        'Gray.jpg', 'jpg');
```

PRETEST

The purpose of the pretest is to check the experiment design for flaws. The pretest allows design problems to be fixed before too much time and money have been invested in the subjective test. Start by viewing the PVSs yourself. The resulting distribution of

quality may show undesirable clusters. If you, the experiment designer, are unable to detect differences between most of your PVSs, the subjects will not either.

A pretest is often performed before some elements of the test are ready (e.g., no automation, no instructions, or an inappropriate environment). The pretest often includes only a subset of the scenes and impairments. A small, biased sampling of subjects (e.g., five to six coworkers or friends) is acceptable, because the goal is to look for design flaws. Example design flaws include problems subjects experienced during the experiment, a narrow range of quality (e.g., no statistically significant conclusions can be reached), and data bunching (e.g., many clips with nearly identical quality). Consider eliminating HRCs with nearly identical quality, when training or testing a model. Zieliński and Rumsey [27] identify potential sources of bias that should be considered during the pretest.

CONDUCTING THE EXPERIMENT

When administering the subjective test, the experimenter should not influence or bias any subject's behavior. To the extent possible, each subject's experience should be identical. The task of choosing an experiment administrator is very important. Good social skills and good communication skills are critical, as part of the administrator's job is to put the subjects at ease. The administrator must be guarded about the test itself and thus less likely to unintentionally influence the results. Questions that may influence the subject's behavior should be answered only after the subject's participation is over.

ETHICS AND INFORMED CONSENT

Awareness of ethical considerations in human testing arose from several infamous psychological and medical experiments. The Belmont Report [28], written in 1979 by the U.S. government, outlines the basic ethical principles in research involving human subjects. In 1991, the U.S. government published the Common Rule [29] for the protection of human subjects. While this policy applies only to U.S. Federal workers, it provides reasonable guidelines for ethical human testing and informed consent.

The first ethical consideration is privacy. Subjects' names must be kept private, and the researcher must ensure that the rating data cannot be used to identify subjects, even accidentally. The easiest and safest way to accomplish this is to identify subjects by number and to never record the number/name association. Second, subjects must be informed of potential risks. Video quality subjective experiments typically have no risk of benefit or harm. Third, subjects should be given an informed consent form to read and sign. The Common Rule contains guidance on appropriate information to include, such as a brief summary of the purpose of the experiment, the method used to keep people's names confidential, any risk or benefit to the subject, notification that participation is voluntary, and who to contact with questions about research subjects' rights or in the event of a research-related injury.

VISION TESTING

Vision testing is traditionally performed before the experiment begins. Unless you are an ophthalmologist, it is inappropriate to tell the subject whether or not they passed the vision test; all people should participate in the experiment, regardless of whether or not their data will be used.

ITU-R BT.500 and ITU-T P.910 require that subjects be screened for normal visual acuity (e.g., with glasses if worn) and normal color vision. Test the subject's distance vision using the Snellen eye chart at a range similar to that used during the experiment. Test color vision with the Ishihara Color Blindness plates under natural lighting (i.e., sunlight). The Ishihara plates should be replaced after about five years, because the colors fade, rendering the test inaccurate. These plates typically only test red–green color blindness, as that is the most common type.

There is some question as to whether there is a difference between ratings from people with normal vision and people who fail the distance vision or color vision test. Pinson et al. [15] found no significant difference in ratings; however, that was not the primary goal of the reported experiment. Moorthy et al. [30] present arguments against the use of vision tests. Regardless, these vision tests convey to the subject that they are participating in an important scientific experiment and should pay attention to their rating task.

SUBJECT DEMOGRAPHICS AND CONVENIENCE SAMPLING

The ideal in psychology is random sampling that perfectly matches the demographics of the population to be studied. This is most easily accomplished by outsourcing subject recruitment to a specialized company that performs market research through focus groups. The drawback is the high cost. Most video quality subjective tests use convenience sampling—i.e., a population of subjects that are easy to obtain. Universities tend to recruit students; large companies tend to recruit employees. To better represent the larger population, consider using a temporary hiring agency or online advertisements.

The problem with convenience sampling is that the research results may not generalize to the larger population. For video quality subjective testing, the relationships between variables will remain correct—but the MOS values will not be absolute (see Pinson et al. [15]). Be careful not to generalize your convenience sampling results into absolute thresholds (e.g., MPEG-2 at this bit rate will result in a quality of 4.0 or better).

INSTRUCTIONS, TRAINING SESSIONS, AND QUESTIONNAIRES

Instructions must be written out and agreed upon before the testing begins. All subjects must receive the same instructions. This eliminates one potential source of subject bias. The instructions should describe the rating cycle, the quality scale, how to record ratings, quirks of your playback system or environment, behavior to be avoided, and the scenario (e.g., watching free video clips on a mobile device and watching a pay-per-view movie). The instructions should include: "Please do not base your opinion on the content of the scene or the quality of the acting." Still, ratings inevitably include both the clip's artistic quality and its technical quality. This is why subjective tests normally include the original video for comparison. After presenting the instructions, ask the subjects if they have any questions.

The training session immediately follows the instructions. The training session serves two purposes. The first is to demonstrate the task. This is easily accomplished with two or three rating cycles. The second purpose is to familiarize the subject with the range of quality and type of impairments in the experiment. This may take much longer (e.g., 5–20 sequences). The SRC used for training session should not appear in the rest of the experiment.

A questionnaire can be used after the rating sessions, to gather additional information. Questionnaires can potentially provide feedback on problems the subject had with the test, whether or not the subject understood the task, and whether the subject noticed a problem with your test setup. A written questionnaire is preferable to asking questions aloud because people are more likely to be blunt and provide additional information. Questionnaires can also be used to understand QoE (see the section "Data Analysis Techniques"). No standard questionnaire exists today.

WRITING THE REPORT

The ultimate goal of a video quality subjective test inevitably includes publishing the results, either internally or externally. The report of results should fully describe

- the goal of the experiment
- the environment (e.g., monitor brand and model, lighting level in lux, viewing distance in screen heights, and picture of environment)
- the test methodology (e.g., ITU recommendation and any departures)
- the rating method
- the SRC (e.g., quantity and sample frames)
- the HRC (e.g., quantity, coding algorithm, bit rate, and transmission error level)
- the experiment design (e.g., full matrix or partial matrix or other, number of PVS)
- the test sessions (e.g., number, duration, playback mechanism, playback compression characteristics, and software used to control the test)
- the subjects (e.g., number of, age and gender distribution)
- the mechanism used to obtain subjects
- the data analysis results.

Always include a picture of the viewing environment in the report. This will provide readers with an improved understanding of the environment (see Figure 8). Brunnström et al. [31] is an example of a superior experiment report.

For privacy reasons, the names of subjects should not be mentioned in any report. Care should be taken when explicitly mentioning vendor names. If the experiment was not designed to directly compare the quality of those vendors' equipment, a comparative analysis might be biased. In such cases, the vendor names should be omitted from external reports.

DATA ANALYSIS TECHNIQUES

Any data analysis is divided into three specific steps: clean the data, choose the correct analytical technique (this step should be done before the subjective study is run), and interpret the results. Each step's method has to fit the problem under investigation. To begin, key statistical concepts will be described.

[FIG8] A picture of the viewing environment should be included in the final report. This sample environment shows a sound isolation booth. A Blu-ray player outside the room plays audiovisual sequences on a broadcast-quality monitor and speakers.

MEASUREMENT SCALES, AGREEMENT, AND ASSOCIATION

The values related to a particular variable (e.g., bit rate) can be measured in different ways [32]. Measurement scales are classified and divided into four types: ratio, interval, ordinal, and nominal. For each type of measurement scale, correct statistical techniques exist and should be used. The ratio measurement scale makes it possible to define a distance between any two values and compute their ratio. For example, the distance between 3 Mbit/s and 6 Mbit/s is 3 Mbit/s, and the second bit rate is two times larger than the first. With the interval measurement scale, the distance between each point is the same but the measured numbers are arbitrary. For example, consider encoder bit rate setup categorized to three values $1 = 3$ Mbit/s, $2 = 5$ Mbit/s, and $3 = 7$ Mbit/s). The measured values 1, 2, and 3 cannot be compared using a ratio (i.e., 3 does not have bit rate three times as high as 1), but the distance between 1 and 2 is the same as the distance between 2 and 3. With an ordinal measurement scale, an order of values can be found but exact distances cannot. For example, bit rate category "medium" has a higher bit rate than "low" and lower than "high." Nevertheless, with an ordinal scale, we cannot determine whether a value "low" has the same distance to "medium" as "medium" to "high." With nominal, each value is different and an order cannot be determined. For example, encoders *A*, *B*, and *C* cannot be ordered without focusing on a specific feature, like price or encoding speed.

Agreement means that subjects should give the same quality ratings, and we only tolerate differences that are caused by a random distribution of measurement noise. On the other hand, association requires only that subjects follow the same pattern. So, if two subjects agree, they also associate, but the opposite is not true [33]. For example, if subject one is always scoring one point lower than subject two (where possible), they do not agree at all but they do associate perfectly. For subjective experiments focused on quality,

agreement is not expected, since a subject can be more or less discriminating. On the other hand, if there is no association it can be explained only if a subject is more tolerant than average for some impairments and less for others. This is not impossible, but an experimenter may choose to consider a subject who does not associate with others to be irrelevant. Perhaps the rating task was too difficult or was not understood properly.

CLEANING THE DATA

Before using any statistical technique, the data has to be cleaned. This involves first detecting irrelevant subjects and second detecting errors in the experiment setup. Detecting an irrelevant subject depends on the way the test was run. It is usual to monitor a subject in a lab; in this case, we can be sure a subject did the test. If this does not happen, the first screening technique is focused on finding out whether a subject actually did the test. Gardlo et al. [34] describe some techniques such as adding content questions, e.g., "Was a car present in the scene?" Such questions reveal whether a subject actually saw a particular sequence. If a test is run in a controlled environment, detecting whether a subject did the test is usually trivial. Nevertheless, the subject could misunderstand the test, the test could be too difficult, or the subject might not pay attention. Therefore, screening is still a necessary and important step.

The screening technique described in the ITU-R Rec. BT-500 is based on the subject agreement and measuring scale being ratios. The basic concept is to measure how often a subject's answers do not fit the confidence interval created by the other subjects' answers. (A MATLAB code for performing this test can be found at http://www.its.bldrdoc.gov/resources/video-quality-research/guides-and-tutorials/subject-screening-overview.aspx.) Subjects do not have to agree among themselves, so we will focus on techniques based on association rather than agreement.

The most popular way to measure association is the Pearson correlation. This technique was used to screen subjects in the Video Quality Experts Group (VQEG) HDTV validation test plan [35]. The Pearson correlation is based on the assumption that the measurement scale is a ratio, and Pearson correlation interpretation and tests generally assume the data have a normal distribution. If the quality scale is short (e.g., ACR), the Pearson correlation should be changed to the Spearman correlation, which is based on the weaker assumption that the measurement scale is ordinal. For a short scale, it is difficult to assume that the answer distribution is close to normal (especially if most answers are close to one of the scale borders) and that the distances between answers are the same. Computing the Pearson or Spearman correlation requires two vectors: u_i (a single subject's ratings) and $\overline{u}_i$ (average of the other subjects' ratings) for all sequences. The MATLAB Statistics Toolbox functions are as follows:

```
a=corr (ui, udashi, 'Type', 'Pearson');
b=corr (ui, udashi, 'Type', 'Spearman');
```

where u_i is denoted by `ui` and $\overline{u}_i$ is denoted by `udashi`.

The question is: At what threshold should a subject be discarded? A correlation that is statistically greater than zero does not guarantee that a subject is relevant. A good example occurred during a VQEG HDTV test [35], the goal of which was to choose the best objective metric for HDTV. For an experiment that used edited sessions on Blu-ray discs, the scoring time was too short for one subject, so he did not see the first few seconds of some sequences. The Pearson correlation was statistically significantly higher than zero (0.633) but much lower than for other subjects' correlation (the lowest being 0.784 and the average 0.877). The VQEG rule is that a subject should be discarded if the correlation is below 0.75 for television (TV) and mobile applications. The disadvantage of using correlation for data cleaning is that the correct threshold must be found experimentally, based on how difficult it is for subjects to evaluate a particular service. The 0.75 threshold cannot simply be used for new services (e.g., 3-D TV, ultraHD) or a different association metric (e.g., Spearman correlation).

These two methodologies (i.e., as described in Rec BT.500 and based on Pearson correlation analysis) are by far the most commonly used to detect irrelevant subjects. They rely on the ratio measuring scale. Since subjective experiments are often performed on a short measuring scale, which is ordinal rather than ratio, some different methods are needed. Adejumo et al. [33] and Gibbons [36] provide descriptions of numerous different agreement and association metrics. A commonly used agreement metric using an ordinal scale is named *Cohen's Kappa coefficient*. It is widely used by the psychology community when subject agreement is especially important. The Kappa coefficient can be computed using the MATLAB function created by Cardillo [37]. No more details are presented in this article since association rather than agreement metrics should be used in case of quality tests. Nevertheless, the interested reader can find details in [33].

Kendall's tau can be an alternative solution when an association metric is needed and the measuring scale is ordinal. Kendall's tau is the difference between the probability of concordance π_c and discordance π_d, given by: $\tau = \pi_c - \pi_d$. Kendall's tau is an easy to interpret parameter, since it refers to concordance and discordance between two vectors. Gibbons [36] provides more details on when Kendall's tau should be used. The MATLAB function is

```
a=corr (ui, udashi, 'Type', 'Kendall');
```

A subject can have a low association with other subjects because the subject did not pay attention to the task or because an error occurred in the experiment setup. While an experiment is being run, many different problems can occur. An intermittent video playback problem in the test interface can cause occasional added visual impairment (e.g., freezes during a sequence that should have played smoothly). Video clips can be played in the incorrect order, so that the written video clip/rating association is incorrect (e.g., an editing error in a prerecorded sequence, a coding bug in automated playback software). Some sequences can be different from the description (e.g., encoded at the wrong bit rate). A time synchronization error can occur between subject

ratings and videos (e.g., a subject scoring on paper rating sheets scores clip N in the box for clip N+1). All of these problems have been observed by the authors of this article.

Specific problems call for specific solutions, so it is impossible to describe a general procedure. Nevertheless, our experience shows that the following three steps are very helpful. First, if prerecorded sequences are used, the results obtained for a PVS should be statistically the same when it appears in different recordings. Second, the quality ratings can be compared to the experimenter's expectations, and sequences with large differences should be checked. Third, the association between users within sessions should be checked. If a user is well associated with others for one session and poorly for another, this indicates a problem with that session.

The most difficult problems to detect are rare interface failures resulting in degradation of the watched video (e.g., the automated test software pauses during video playback for 2% of the sequences). This problem is best detected during the pretest (e.g., the experimenter takes the pretest, and then compares his or her scores to the expected quality). Intermittent playback failures will cause all subjects to score atypically low quality for a few, random clips. One way to find such inconsistencies is to compare a user's association with other users for all sequences with that user's association with other users for a subset of sequences. If the percentage of errors in a subset is higher than in the whole set it is easier to detect them.

If the data cleaning eliminates more than two or three subjects, something may be wrong with your test procedure. To put this into perspective, only one of the 214 subjects in [15] was eliminated for being irrelevant (see [38] to view the individual subject ratings).

DATA ANALYSIS

After screening subjects and ensuring that all of them performed the experiment properly, the final and most important analysis can be run. The section "Goal of Experiment and Design Consequences" described the different reasons to run subjective experiments. These different reasons call for different data analyses.

ANSWERING A QUESTION

An experiment designed to answer a question contains different conditions, which are most often different HRCs but could also be different subjects or different SRCs. Different conditions generate groups of results that can be compared to answer specific question. Therefore, answering a question can be reduced to comparing subsets of subjective experiment results. The most common way to compare two groups is to answer the question of whether the results are statistically the same or not. This question will be answered with a specific significance level. Most often, 5% significance (i.e., 95% confidence) is chosen. This is the default in many MATLAB statistical tests and is specified as 0.05.

The most popular technique for comparing two groups is the Student's t-test. The goal of the Student's t-test is to validate if the difference between the mean values of two groups has a particular value. For example, if our goal is to validate whether the quality obtained for HRC 1 and HRC 2 are the same we should compare vectors $\overline{u}_1$ and $\overline{u}_2$. Each element of the $\overline{u}_i$ vector is a value obtained for the same HRC with different other conditions (e.g., SRCs or repetitions). It is very important to have the same order of conditions in both vectors. The MATLAB function for computing the Student's t-test is: `[h, p]=ttest(u1-u2);` where `h` is one if the difference is statistically different from zero and zero if it is not, and `p` is the corresponding p-value, which has to be larger than 0.05 to conclude that the obtained difference is not statistically significant.

If the goal is to compare multiple groups, then the methodology and significance level must be adjusted to maintain the same significance level for a group as for single comparison. The commonly used methodology for comparing multiple groups is one-way analysis of variance (ANOVA). The MATLAB `anova1` computes one-way ANOVA. A handy way to call this function is to specify a vector of all compared values (e.g., MOSs) and a vector of tags describing groups. For example, `u=[2.3,3.2,1.2,2.4,3.2]` and `g=['A','B','B','A','C']` means that the first and fourth values belong to group A, the second and third to group B and the last one to group C. `anova1` is called using: `p = anova1(u,g);` where `u` is the compared values vector, `g` is the grouping vector, and output `p` is the p-value. Similarly to `ttest`, a value of `p` smaller than 0.05 indicates that at least one group is different.

The disadvantage of both the Student's t-test and ANOVA is the assumption that the data come from a normal distribution, i.e., they follow a specific distribution and can be measured on a ratio or interval scale. This can be validated by the Kolmogorow–Smirnow (small sample) or chi square (large sample) test. If one of those assumptions is not met, different statistical methods should be used. The Student's t-test should be changed to the Mann–Whitney U-test and the one-way ANOVA should be changed to the Kruskal–Wallis test. The Mann-Whitney U-test compares medians not means, and as such it needs only an ordinal measuring scale. The equivalent of this test in MATLAB is `[p, h] = ranksum(u1, u2);` where the `p, h, u1,` and `u2` parameters are the same as for the `ttest` function but ordered differently. The multiple-group comparison version of the Mann–Whitney U-test is the Kruskal–Wallis test, which can be called similarly to the `anova1` function: `p = kruskalwallis(u,g);`

The discrimination powers of the Student's t-test and the ANOVA test are greater than those of the Mann–Whitney U-test and the Kruskal–Wallis test. The cost for this increased discrimination is the requirement of a normal distribution and interval measuring scale. Choosing the Mann–Whitney U-test or Kruskal–Wallis test leads to more conservative conclusions. If a test with the weaker assumption shows differences between groups, then the Student's t-test and ANOVA will show it as well. The opposite is not necessarily true.

A useful MATLAB tool for comparing multiple groups is the `multcompare` function. Both `anova1` and `kruskalwallis` functions can return more than one parameter. After three parameters are returned, `multcompare` can be run by using

```
[p,table,stats] = kruskalwallis(u,g);
multcompare(stats);
```

This function generates a handy interactive plot that makes it easy to compare groups.

TRAINING A METRIC OR ALGORITHM

Subjective experiments maximize measurement accuracy, but also increase cost and time taken. They cannot be used to monitor a service. Therefore, it is a common practice to build a metric that objectively emulates a video quality (i.e., MOS). When training a metric or algorithm, the goal is to find a function that links explanatory variables to a dependent variable. Example explanatory variables are bit rate, packet loss ratio, quality estimation parameters extracted from the video, and subject age. The dependent variable is most often video quality.

The easiest solution is to design a linear model using linear regression. A linear model is a linear function of model parameters, but not necessary a linear function of explanatory variables. Nonlinearities in the explanatory variables are detected and removed (e.g., using square or square root functions), and it is common to use interactions between explanatory variables (e.g., the product of two explanatory variables). According to the above description an example linear model is given by the equation $u = a_0 + a_1 b + a_2 \log ta + a_3 b \log ta$, where u is estimated MOS, a_i are model parameters, b is a bit rate, and ta is the temporal activity of the SRC. A linear model can be estimated in MATLAB by the `glmfit` function, which returns both the estimated values and the p-values of each estimated parameter.

While training the linear model, the researcher examines and understands the relationship between the candidate explanatory variables and the dependent variable. Example techniques include examining the ability of a single explanatory variable to predict the dependent variable (e.g., using the Pearson correlation or root mean square error), and plotting the explanatory variable against the dependent variable to find nonlinearities or outliers. Fox [39] provides instruction on techniques for applying linear regression analysis. The advantage of linear regression is that the resulting linear model is typically easy to explain and understand.

Alternatives to classical linear regression are methods based on machine learning. Many techniques are available. In this article, only three are mentioned. Genetic programming-based symbolic regression analyzes a large number of different models, thus helping to build a model that is similar to a linear model [40]. The advantage of this technique is that the output is easy to interpret. Partial least squares regression is more difficult to interpret but has the advantage of optimizing explanatory variables. Because of the use of principal component analysis, the final output is as simple as possible for a given prediction accuracy using explanatory variables that contain the most significant information [41]. Random neural networks are even more difficult to interpret but can approximate different nonlinear functions [42].

Machine-learning algorithms must be used with care to not over train the model. A typical machine-learning model contains lots of parameters, and relatively little subjective data are typically available to train a video quality metric.

All of the previously presented solutions model MOS (i.e., the average of many ratings), not the actual subjective ratings. If quality is measured on a scale with a small number of levels, each rating level's probability can be predicted using the generalized linear model (GLM). GLM is able to model multinomial distribution. A detail description of using GLM is given in [43].

ANALYZING A METRIC OR ALGORITHM

To analyze a metric, its predictions and subjective results have to be compared. The algorithm that fits the subjective data best should be chosen. This analysis has to address two specific realities of subjective experiments. Previous research shows that two instances of the same subjective experiment repeated in two different laboratories can have high association (measured by correlation) but the results are not identical [3], [15]. Since the results of the two subjective experiments results have high association but not necessarily agreement, a metric should associate with the validation subjective experiment but it does not have to agree (i.e., an offset is possible).

The final conclusion is that metrics should be validated by association rather than by agreement, or agreement should be measured for the metric after the values have been transferred to a common scale. In addition, two metrics can differ due to randomness related to the subjective experiment. Such metrics should be called "the same" even if the agreement or association metric is superior for one of them. The methodology for addressing the problems described above is used by VQEG and is described in ITU-T Rec. P.1401.

THOUGHTS ON QUALITY OF EXPERIENCE

We have provided detailed information for conducting video quality subjective tests. Video quality is one aspect of a larger topic—QoE. Compared to video quality testing, QoE testing is in its infancy, and no step-by-step tutorial is available at this time. Instead, this final section summarizes some QoE definitions and frameworks. This overview points out limitations of video quality subjective tests and identifies areas where QoE issues impact experiment design, monitor selection, subject demographics, post-test questionnaires and, as a consequence of these choices, the strength of the conclusions that can be reached.

Video quality is just one aspect of QoE. According to Le Callet et al. [44], QoE is the degree of delight or annoyance the user receives from an application or service. It results from the fulfillment of the user's expectations (in light of his or her personality and current state) with respect to the utility and/or enjoyment of the application or service. More succinctly, QoE is a measure of how well a service or an application meets the user's expectation of quality (EoQ) [45]. Different artifacts arising along an end-to-end service delivery chain may result in QoE that does not meet a user's EoQ. However, each service provider is expected to aim at the condition QoE = EoQ [46], ensuring revenues, reducing churn, and increasing customer satisfaction.

Today, each of us is a consumer of multimedia services and knows how many variables influence our EoQ. Therefore, a

holistic QoE approach should span the whole telecommunication ecosystem combining user behavior, technical issues, and business models as proposed in [44] and [47].

Batteram et al. [45] propose three dimensions that can be used to express QoE: service availability, service responsiveness, and media quality (i.e., audio and video quality). Service availability is a measure of whether the user can use the desired service, while responsiveness is the time to get the service answer. Media quality relates to all artifacts generated by compression and packet network delivery that deteriorate the user's perception. Audio and video quality subjective tests measure media quality, but fail to quantify the impact of service availability and responsiveness.

On the other hand, Marez and Moor [46] point out that QoE may depend on many service context-of-use factors (i.e., the actual conditions under which an application is used). The service has to be paid for through some provider-defined business mode (e.g., transaction, subscription, and advertisement). The underlying network technology (e.g., wired, wireless, and satellite) impacts QoE, as do other technological factors. Personal, social, cultural, and education issues are influential, and a user's EoQ is modified by the location or device used for service consumption. There is interest in extending video quality subjective testing techniques but as of yet no established solution (e.g., a way to measure MOS that accounts for screen size differences).

The multivariate structure of QoE may suggest initially that, from the QoE analytical modeling point of view, numerous analysis models could be deployed to understand the relationships between variables and their relevance to the actual QoE problem being studied [44]. An expert panel [46] found about 60 multidisciplinary methods (both qualitative and quantitative) suited for QoE investigations.

While the rating scales in the section "Ethics and Informed Consent" are intervals (which define the ratio, interval, ordinal, and nominal measurement scales), QoE variables are often ordinal (e.g., satisfied, neutral, and dissatisfied) or nominal (e.g., gender, user profile, device, and content type). These category variables differ radically from interval variables because distances between categories are not defined and subjects can interpret the categories differently. Thus, most of the techniques from the section "Data Analysis Techniques" are inappropriate. The proper tool for dealing with such unmeasurable variables is categorical data analysis (i.e., multicategory logit models). These techniques are more complicated, and the results derived are a bit more difficult to interpret.

Customer satisfaction surveys solve this problem by using a variety of latent trait models (LTMs). For example, the Item Rasch Theory is the simplest LTM model. The LTM is a powerful approach as it can relate manifest variables (i.e., service features that can be readily judged by a tester) with latent traits (i.e., the tester's experience with a service)—provided that the questionnaire is properly designed. The LTM approach was recently suggested by [48] and [49] as a proper tool for 3-D video quality analysis.

In summary, QoE uses multiple dimensions to measure different users' experiences of service received and relate their experiences to parameters of a service delivery chain and a service context-of-use. A reliable QoE measurement calls for a multidisciplinary approach (e.g., operations research, customer satisfaction surveys, and sociology), because of the different nature of the variates involved. Users' experiences may differ even if they use the application in the same context and under the same network conditions. Therefore, to arrive at a valid QoE assessment, it is necessary to conduct tests with large numbers of subjects. Although such subjective tests are time and resource consuming, emerging crowdsource QoE assessment has recently appeared as a solution [50]. QoE is gaining increasing momentum among researchers, service providers, and network operators; this may eventually result in the implementation of user-centric services.

ACKNOWLEDGMENTS

We would like to thank the many researchers whose questions or advice contributed to this article, as well as Florence Agboma and Jessica Hooker. The work of Lucjan Janowski and Zdzislaw Papir was supported by the AGH University of Science and Technology under contract 15.11.230.075.

AUTHORS

Margaret Pinson (margaret@its.bldrdoc.gov) received her B.S. and M.S. degrees in computer science from the University of Colorado at Boulder in 1988 and 1990. Since 1988, she has been investigating improved methods for assessing video quality at the Institute for Telecommunication Sciences, an office of the National Telecommunication and Information Administration, in Boulder, Colorado. She is a cochair of VQEG and a cochair of the Independent Lab Group in the VQEG, an associate rapporteur of Question 12 in the ITU-T Study Group 9, and the administrator of the Consumer Digital Video Library.

Lucjan Janowski (janowski@kt.agh.edu.pl) is an assistant professor with the Department of Telecommunications, AGH University of Science and Technology. He received his Ph.D. degree in telecommunications in 2006 from the AGH. In 2007, he worked in a postdoctoral position at the Centre National de la Recherche Scientifique (CNRS), LAAS (Laboratory for Analysis and Architecture of Systems of CNRS) in France, where he prepared both malicious traffic analysis and anomaly detection algorithms. In 2010–2011, he spent half a year in a postdoctoral position at the University of Geneva, working on quality of experience (QoE) for health applications. His main interests are statistics and probabilistic modeling of subjects and subjective rates used in QoE evaluation.

Zdzisław Papir (papir@kt.agh.edu.pl) is a professor and a deputy chair with the Department of Telecommunications, AGH University of Science and Technology in Cracow, Poland. From 1994 to 1995 he served as a network design department manager for Polish Cable Television. He was a guest coeditor for *IEEE Communications Magazine* from 1999 to 2006, responsible for the broadband access series. He participates in several R&D projects under the Information Society Technologies Work Program of the European Commission, responsible for network performance evaluation and quality assessment of communication services. He is

also an information and communication technology expert as appointed by the European Commission.

REFERENCES

[1] (2014, Feb. 5). Cisco visual networking index: Global mobile data traffic forecast update, 2013–2018. [Online]. Available: http://www.cisco.com/

[2] M. Kuniavsky, A. Moed, and E. Goodman, *Observing the User Experience: A Practitioner's Guide to User Research*. Waltham, MA: Elsevier Science, 2012.

[3] G. Cermak et al. (2008, Mar. 28). Final report from the video quality experts group on the validation of objective models of multimedia quality assessment, phase I. [Online]. Available: https://www.vqeg.org

[4] A. Younkin and P. Corriveau, "The effects of quantity and location of video playback errors on the average end-user's experience," in *Proc. IEEE Int. Symp. Broadband Multimedia Systems and Broadcasting*, 2008, pp. 1–5.

[5] M. H. Pinson, S. Wolf, and G. Cermak, "HDTV subjective quality of H.264 vs. MPEG-2, with and without Packet Loss," *IEEE Trans. Broadcast.*, vol. 56, no. 1, pp. 86–91, Mar. 2010.

[6] K. Brunnström, L. Nordtström, and B. Andreén, "Visual experience of quality degradations when viewing computer and notebook displays from an oblique angle," *J. Soc. Inform. Display*, vol. 19, no. 5, pp. 387–397, May 2011.

[7] S. Voran and S. Wolf, "The development and evaluation of an objective video quality assessment system that emulates human viewing panels," in *Proc. Int. Broadcasting Convention (IBC)*, July 1992, pp. 504–508.

[8] Q. Huynh-Thu and M. Ghanbari, "Temporal aspect of perceived quality in mobile video broadcasting," *IEEE Trans. Broadcast.*, vol. 54, no. 3, pp. 641–651, Sept. 2008.

[9] S. Voran and A. Catellier, "When should speech coding quality increases be allowed in talk-spurts?" in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, May 2013, pp. 8149–8153.

[10] "Subjective quality evaluation of audio and audiovisual multiparty telemeetings," ITU-T Rec. P.1301, 2012.

[11] G. Berndtsson, M. Folkesson, and V. Kulyk, "Subjective quality assessment of video conferences and telemeetings," in *Proc. Packet Video Workshop (PV)*, May 2012, pp. 25–30.

[12] M. H. Pinson, M. Barkowsky, and P. Le Callet, "Selecting scenes for 2D and 3D subjective video quality tests," *EURASIP J. Image Video Processing*, vol. 2013, no. 1.

[13] Y. Niu and F. Liu, "What makes a professional video? A computational aesthetics approach," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 7, July 2012, pp. 1037–1049.

[14] A. Catellier, M. Pinson, W. Ingram, and A. Webster, "Impact of mobile devices and usage location on perceived multimedia quality," in *Proc. Int. Workshop on Quality of Multimedia Experience* (QoMEX), July 2012, pp. 39–44.

[15] M. H. Pinson, L. Janowski, R. Pépion, Q. Huynh-Thu, Ch. Schmidmer, P. Corriveau, A. Younkin, P. Le Callet, M. Barkowsky, and W. Ingram, "The influence of subjects and environment on audiovisual subjective test: an international study," *IEEE J. Select. Topics Signal Processing*, vol. 6, no. 6, pp. 640–651, Oct. 2012.

[16] T. Tominaga, T. Hayashi, J. Okamoto, and A. Takahashi, "Performance comparisons of subjective quality assessment methods for mobile video," in *Proc. Int. Workshop on Quality of Multimedia Experience* (QoMEX), June 2010, pp. 82–87.

[17] Q. Huynh-Thu and M. Ghanbari, "A comparison of subjective video quality assessment methods for low-bit rate and low-resolution video," in *Proc. Signal and Image Processing*, M. W. Marcellin, Ed., Honolulu, Hawaii, USA, 2005, vol. 479, pp. 70–76.

[18] M. D. Brotherton, Q. Huynh-Thu, D. S. Hands, and K. Brunnström, "Subjective multimedia quality assessment," *IEICE Trans. Fundamentals, Electron. Commun. Comput. Sci.*, vol. E89-A, no. 11, pp. 2920–2932, 2006.

[19] M. Pinson and S. Wolf, "Comparing subjective video quality testing methodologies," in *Proc. SPIE Video Communications and Image Processing Conf.*, Lugano, Switzerland, July 2003, pp. 8–11.

[20] T. Liu, G. Cash, N. Narvekar, and J. Bloom, "Continuous mobile video subjective quality assessment using gaming steering wheel," in *Proc. Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Jan. 2012.

[21] S. Buchinger, W. Robitza, M. Nezveda, M. Sack, P. Hummelbrunner, and H. Hlavacs, "Slider or glove? Proposing an alternative quality rating methodology," in *Proc. 6th Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Jan. 2010.

[22] R. A. Doherty, A. C. Younkin, and P. J. Corriveau, "Paired comparison analysis for frame rate conversion algorithms," in *Proc. Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Jan. 2009.

[23] G. A. Miller. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* [Online]. 63, pp. 81–97. Available: http://www.musanim.com/miller1956/

[24] J. Jonsson and K. Brunnström. (2007). *Getting Started with AcrVQWin*. [Online]. Acreo AB, Kista: Sweden. Available: https://www.acreo.se/acrvqwin

[25] A. Jain, C. Bal, and T. Nguyen, "TALLY: A web-based subjective testing tool," in *Proc. Int. Workshop on Quality of Multimedia Experience* (QoMEX), July 2013, pp. 128–129.

[26] S. Buchinger, W. Robitza, M. Nezveda, P. Hummelbrunner, and H. Hlavacs. (2010, Dec. 15). Towards a comparable and reproducible subjective outdoor multimedia quality assessment, *Third Euro-NF IA.7.5 Workshop on Socio-Economic Issues of Networks of the Future*. [Online]. Available: http://code.google.com/p/subjectiveplayer/

[27] S. Zieliński and F. Rumsey, "On some biases encountered in modern audio quality listening tests—A review," *J. Audio Eng. Soc.*, vol. 56, no. 6, June 2008, pp. 427–451.

[28] U.S. Department of Health & Human Services. (1979, Apr. 18). Ethical principles and guidelines for the protection of human subjects of research. [Online]. Available: http://www.hhs.gov

[29] U.S. Department of Health & Human Services. (1991). Federal policy for the protection of human subjects ('Common Rule'), 45 CFR part 46 [Online]. Available: http://www.hhs.gov

[30] A. K. Moorthy, L. K. Choi, A. C. Bovik, and G. de Veciana, "Video quality assessment on mobile devices: subjective, behavioral and objective studies," *IEEE J. Select. Topics Signal Processing*, vol. 6, no. 6, pp. 652–671, Oct. 2012.

[31] K. Brunnström, L. Nordström, and B. Andrén, "Visual experience of quality degradation when viewing computer and notebooks displays from an oblique angle," *J. Soc. Inform. Display*, vol. 19, no. 5, pp. 387–397, Dec. 2011.

[32] A. Agresti, *Categorical Data Analysis*, 2nd ed. Hoboken, NJ: Wiley, 2002.

[33] A. O. Adejumo, C. Heumann, and H. Toutenburg, "A review of agreement measure as a subset of association measure between raters," Sonderforschungsbereich 386 der Ludwig-Maximilians-Universität München, no. 385, 2004.

[34] B. Gardlo, M. Ries, and T. Hossfeld, "Impact of screening technique on crowdsourcing QoE assessments," Radioelektronika (RADIOELEKTRONIKA), in *Proc. 2012 22nd Int. Conf.*, 17–18 Apr. 2012, pp. 1–4.

[35] M. Pinson et al. (2010). Report on the validation of video quality models for high definition video content, in *VQEG*. [Online]. Available: https://www.vqeg.org

[36] J. D. Gibbons, *Nonparametric Measures of Association*. Thousand Oaks, CA: SAGE, 1993.

[37] G. Cardillo. (2009, Dec.). Cohen's kappa: Compute the Cohen's kappa ratio on a 2x2 matrix [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/15365

[38] M. H. Pinson, L. Janowski, R. Pépion, Q. Huynh-Thu, Ch. Schmidmer, P. Corriveau, A. Younkin, P. Le Callet, M. Barkowsky, and W. Ingram, "Subjective and objective evaluation of an audiovisual subjective dataset for research and development," in *Proc. Int. Workshop on Quality of Multimedia Experience (QoMEX)*, July 2013, pp. 30–31.

[39] J. Fox, *Applied Regression Analysis and Generalized Linear Models*, Thousand Oaks, CA: Sage, 2008.

[40] N. Staelens, D. Deschrijver, E. Vladislavleva, B. Vermeulen, T. Dhaene, and P. Demeester, "Constructing a no-reference H.264/AVC bitstream-based video quality metric using genetic programming-based symbolic regression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 8, pp. 1322–1333, Aug. 2013.

[41] C. Keimel, J. Habigt, M. Klimpke, and K. Diepold, "Design of no-reference video quality metrics with multiway partial least squares regression," in *Proc. Int. Workshop on Quality of Multimedia Experience (QoMEX)*, 7–9 Sept. 2011, pp. 49, 54.

[42] K. D. Singh and G. Rubino, "No-reference quality of Experience monitoring in DVB-H networks," in *Proc. Wireless Telecommunications Symp. (WTS)*, 21–23 Apr. 2010, pp. 1–6.

[43] L. Janowski and Z. Papir, "Modeling subjective tests of quality of experience with a generalized linear model," in *Proc. Int. Workshop on Quality of Multimedia Experience (QoMEX)*, 29–31 July 2009, pp. 35–40.

[44] A. Raake and S. Egger, "Quality and quality of experience," in *Quality of Experience: Advanced Concepts, Applications and Methods*, by S. Möller and A. Raake, Ed. New York: Springer, 2014, ch. 2, pp. 11–33.

[45] H. Batteram, G. Damm, A. Mukhopadhyay, L. Philippart, R. Odysseos, and C. Urrutia-Valdés, "Delivering quality of experience in multimedia networks," *Bell Labs Tech. J.*, vol. 15, no. 1, pp. 175–194, 2010.

[46] L. De Marez and K. De Moor, "The challenge of user- and QoE-centric research and product development in today's ICT-environment," *Observ. J.*, vol. 1, no. 3, 2007, pp. 1–22.

[47] K. Kilkki, "Quality of experience in communications ecosystem," *J. Univ. Comput. Sci.*, vol. 14, no. 5, pp. 615–624, 2008.

[48] M. Grega, L. Janowski, M. Leszczuk, P. Romaniak, and Z. Papir, "Quality of experience evaluation for multimedia services," Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne, vol. 81, no. 4, pp. 142–153, Apr. 2008.

[49] NTT Network Technology Laboratories. (2014, Sept. 30). Communications Service QoE Assessment, Research Portal Site @ NTT [Online]. Available: http://www.ntt.co.jp/qos/qoe/eng/concept/directionality.html

[50] C.-C. Wu, K.-T. Chen, Y.-C. Chang, and C.-L. Lei, "Crowdsourcing multimedia QoE evaluation: A trusted framework," *IEEE Trans. Multimedia*. vol. 15, no. 5, pp. 1121–1137, Aug. 2013.

[51] VQEG eLetter. [Online]. vol. 1, no. 2. Available: http://www.vqeg.org

[SP]

[life SCIENCES]

Dario Farina
and Ales Holobar

Human–Machine Interfacing by Decoding the Surface Electromyogram

Human–machine interfacing (HMI) uses biological signals, such as brain signals, and translates them into control commands for external devices. An exciting application of HMI is the active control of robotic limbs that substitute lost body parts (bionic reconstruction). Despite recent progresses in brain–computer interfacing [1], muscle electrical potentials are still the most important input for clinical HMI. The ensemble electrical activity of a muscle, the electromyogram (EMG), is generated by the neural activation of the motor neurons innervating the muscle and thus contains information on the neural control of movement. Muscle recordings can be performed noninvasively (surface EMG) by electrodes located on the skin to record the electrical activity of the underlying muscle fibers. This recording modality has advantages compared to other HMI modalities, such as those based on implanted electrodes, in terms of long-term stability, biocompatibility, reduced risks of infections, and ethical constraints.

While traditional surface EMG processing exploited the neural origin of the EMG indirectly, in this column we introduce a direct approach for accessing the neural code present in the signal. We discuss a novel concept of HMI that uses the behavior of individual motor neurons extracted from noninvasive muscle signals, rather than from direct nerve interfacing. This interface has broad applicability since it could be used to decode the activity of any nerve, either via its naturally innervated muscle or via reinnervation of new muscles [targeted muscle reinnervation (TMR)]. This HMI offers high potential for neurotechnologies, e.g., active prostheses or rehabilitation robotics, and as a general-purpose HMI, e.g., in sports science or gaming. In the following, we will discuss the models, processing methods, and results that show the feasibility for this novel HMI concept.

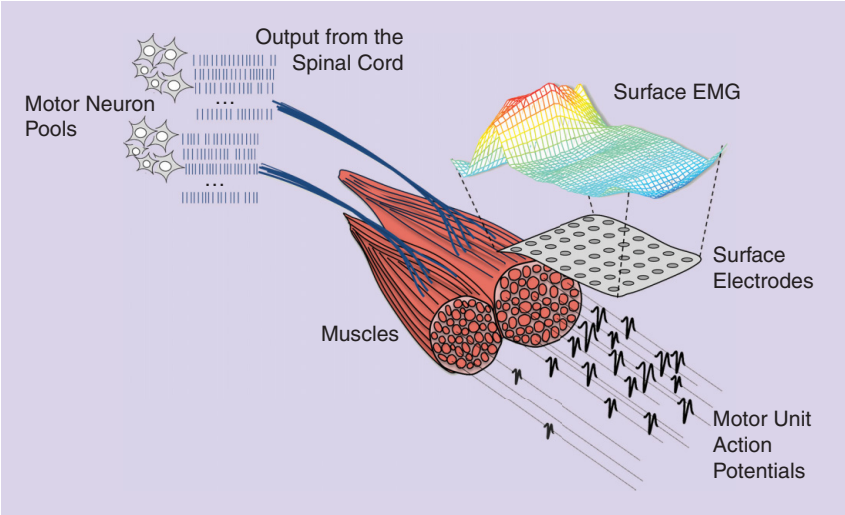
Digital Object Identifier 10.1109/MSP.2014.2359242
Date of publication: 5 December 2014

BACKGROUND

The surface EMG is the electrical manifestation of the activity of muscle fibers. Muscle fibers are not controlled individually but in groups of tens to hundreds that form the muscle units. Each muscle unit is innervated by a single motor neuron [2], and the motor neuron and its innervated muscle fibers constitute the smallest voluntary unit in a movement—the motor unit (MU). Because each action potential discharged by a motor neuron corresponds to an action potential in the innervated muscle fibers, the muscle transforms neural action potentials into compound potentials of the muscle unit. In this transformation, the information on the time instants of motor neuron activations is unaltered. Therefore, the electrical signals recorded over or inside a muscle may be seen as an amplified version of the corresponding neural signals of the efferent nerve fibers. In this regard, the generative model of the EMG signal is the same as the model of the ensemble activity of efferent nerve fibers, i.e., the sum of convolutions of delta functions and deterministic waveforms. The EMG signal can be seen as a neural recording from a peripheral organ—the muscle—that biologically amplifies the signal, as schematically shown in Figure 1. The source of neural information is in the ensemble of activation times of the nerve action potentials, whereas the waveforms of the muscle fiber action potentials reflect peripheral muscle properties. The concept of biological amplification of nerve activity is artificially exploited within the TMR procedure [3]. In TMR, nerves that normally would innervate missing muscles following an amputation are surgically directed into accessory muscles, used exclusively as sources to detect the nerve activity and to form a neural interface with the human body [3].

WHILE TRADITIONAL SURFACE EMG PROCESSING EXPLOITED THE NEURAL ORIGIN OF THE EMG INDIRECTLY, IN THIS COLUMN WE INTRODUCE A DIRECT APPROACH FOR ACCESSING THE NEURAL CODE PRESENT IN THE SIGNAL.

The activation of spinal motor neurons constitutes the ultimate neural representation of human movement because this activation is then transformed into muscle forces and motion. This last neural signal in the chain of neural processing stages for the generation of a voluntary movement is often termed the *neural drive to the muscle*. The link between the EMG and the neural output from motor neurons has been extensively used by extracting properties of the neural code from transformations of the raw EMG signals [4]. For example, the intensity of the neural drive to a muscle is often estimated as the amplitude of the EMG, i.e., the square root of the power, whereas the EMG spectral moments have been used for inferring properties of the active motor neurons [4].



[FIG1] A schematic representation of the generation of the surface EMG signal. In this example, two muscles are represented. The muscles are driven by the neural drive transmitted by the motor neuron pools. Each activation instant of a motor neuron corresponds to the generation of the compound action potential of the innervated muscle fibers, the MU action potential (MUAP). The surface EMG is recorded on the skin overlying the muscles. In this example, a grid of 7 × 6 surface EMG electrodes is shown representatively. The grid provides a spatial sampling of the electric potential generated by the activity of the muscle fibers.

CHALLENGES IN DECODING THE SURFACE EMG

Extracting the neural information from an EMG signal directly corresponds to identifying the trains of delta functions that drive the occurrence of muscle fiber action potentials (Figure 1). Each motor neuron innervates a different group of muscle fibers and, therefore, in principle determines the occurrence of a unique surface action potential waveform. This waveform repeats over time and constitutes an electrical signature of the activity of that specific motor neuron. It is intuitive—and will be evident when discussing EMG models in the following section—that a basic condition for distinguishing sources from the surface EMG is that the surface action potentials discharged by the same MU have a shape of greater similarity than those discharged by different MUs. With respect to other decomposition problems, e.g., spike sorting of neural recordings or of intramuscular EMG [7], the waveforms of the surface action potentials characterizing each source are, however, relatively similar. This problem is related to the bandwidth of the surface EMG signal that is substantially limited by the effect of the tissues interposed between the electrodes and the muscle fibers, i.e., the volume conductor. The effect of the volume conductor is that of a low-pass filter with cutoff frequency inversely related to the distance from the sources. For long distances, a relatively large amount of discriminative information on the shape of the action potentials is filtered out by the volume conductor so that the action potentials generated by different sources appear similar. Accordingly, simulation studies have shown that, out of a population of hundreds of MUs, the majority may have surface action potentials "identical" (within experimental noise variability) to those of at least one other MU [8]. The solution to this problem of nonuniqueness in the electrical signature of different MUs is to record more than one EMG signal by sampling the electrical potential at the surface of the skin at several sites (as schematically represented in Figure 1). The spatial sampling increases the differences between action potential waveforms—in the multichannel domain—so that, for a sufficient number of channels, the representative action potential of each source differs from all others [8]. For the above reasons, the decomposition of the surface EMG is theoretically possible under general conditions, yet only when using multichannel or high-density surface electrode systems [8].

With respect to other spike sorting problems, a second challenge in the decomposition of the surface EMG is the higher level of time superimposition of action potentials. This is due to the larger number of sources that significantly contribute to the signal compared with other recording modalities, such as intramuscular EMG, and thus to the selectivity of the recording. The rate at which the action potentials overlap in time can be estimated as $DR^2 \cdot d \cdot I \cdot (I - 1)$, where DR is the average discharge rate [pulses per second (pps)] of the active sources, d is the average duration (s) of the action potentials, and I is the number of sources contributing to the signal [9]. This equation indicates that the rate of overlap depends quadratically on both the discharge rate and the number of active sources. For example, with all other conditions being equal, the MU action potential (MUAP) overlap rate in a surface EMG signal that comprises 30 significantly contributing MUs is approximately 150 times greater than the rate in a selective intramuscular EMG signal comprising three main sources (these numbers are typical). With these rates of overlap, the procedures based on segmentation of the signal into intervals containing individual waveforms and clustering of these waveforms—these procedures are common for spike sorting of neural recordings [7]—cannot be used for surface EMG decomposition. Indeed, under most conditions, the occurrence of an MUAP waveform isolated in time from all the others in a surface EMG recording is an extremely rare event, much less likely than the superimposition of several waveforms. Therefore, blind source separation methods are preferred as they are minimally influenced by the MUAP overlap rate [5].

SURFACE EMG GENERATION MODEL

In this article, we focus on an EMG decomposition approach based on multichannel surface EMG recordings. These recordings can be modeled by a multiple-input, multiple-output (MIMO) system [5]:

$$\mathbf{EMG}(t) = \mathbf{B}\bar{\mathbf{s}}(t) + \boldsymbol{\omega}(t) \qquad (1)$$

where $\mathbf{EMG}(t) = [\mathrm{EMG}_1(t), \ldots, \mathrm{EMG}_M(t)]^T$ is a vector of M EMG signals, $\boldsymbol{\omega}(t) = [\omega_1(t), \ldots, \omega_M(t)]^T$ is an additive noise vector, and $\bar{\mathbf{s}}(t)$ is the source vector to be estimated that contains the information about the neural drive to the muscle.

The matrix $\mathbf{B}$ represents the volume conductor filters and the model is, thus, convolutive [5]. It contains the action potential waveforms from N MUs as detected by all EMG channels:

$$\mathbf{B} = \begin{bmatrix} \mathbf{m}_{1,1}(t) & \mathbf{m}_{1,2}(t) & \cdots & \mathbf{m}_{1,N}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{m}_{M,1}(t) & \mathbf{m}_{M,2}(t) & \cdots & \mathbf{m}_{M,N}(t) \end{bmatrix}, \qquad (2)$$

where the vector $\mathbf{m}_{i,j}(t)$ denotes the L samples long action potential of the jth MU, as measured by the ith uptake electrode at the time instant t.

The vector $\bar{\mathbf{s}}(t) = [s_1(t), s_1(t-1), \ldots s_1(t-L+1), s_2(t), \ldots, s_R(t), s_{\mathrm{NR}}(t-1), \ldots s_{\mathrm{NR}}(t-L+1)]^T$ contains blocks of L consecutive samples from N motor neuron spike trains [5]. Each spike train $s_j(t)$ is modeled as a binary sequence of zeros and ones (Figure 1), describing the discharge pattern of the jth motor neuron.

Although sparse, the data model (1) is typically undercomplete, with more sources than measurements, and thus difficult to invert. The number of EMG channels depends on the acquisition system, but it is typically limited to several tens, whereas there are potentially hundreds of MUs active during a motor task. However, the sources that contribute action potentials of relatively small energy can be included into the noise term so that the number of sources is reduced, with a concomitant decrease in the signal-to-noise ratio (SNR). Surface EMG decomposition consists of estimating the source vector $\bar{\mathbf{s}}(t)$ given the measurements $\mathbf{EMG}(t)$.

BLIND SOURCE IDENTIFICATION FROM THE SURFACE EMG

A few algorithms have been proposed for inversion of the data model (1) and, thus, estimation of the neural drive to the muscle from surface EMG signals. Among them, the convolution kernel compensation (CKC) method [5] performs a computationally efficient compensation of the shapes of the MUAP and directly estimates the neural information from the surface EMG (Figure 2). The method works sequentially for each MU, computing the following linear minimum mean-square estimator (LMMSE) of a MU spike train:

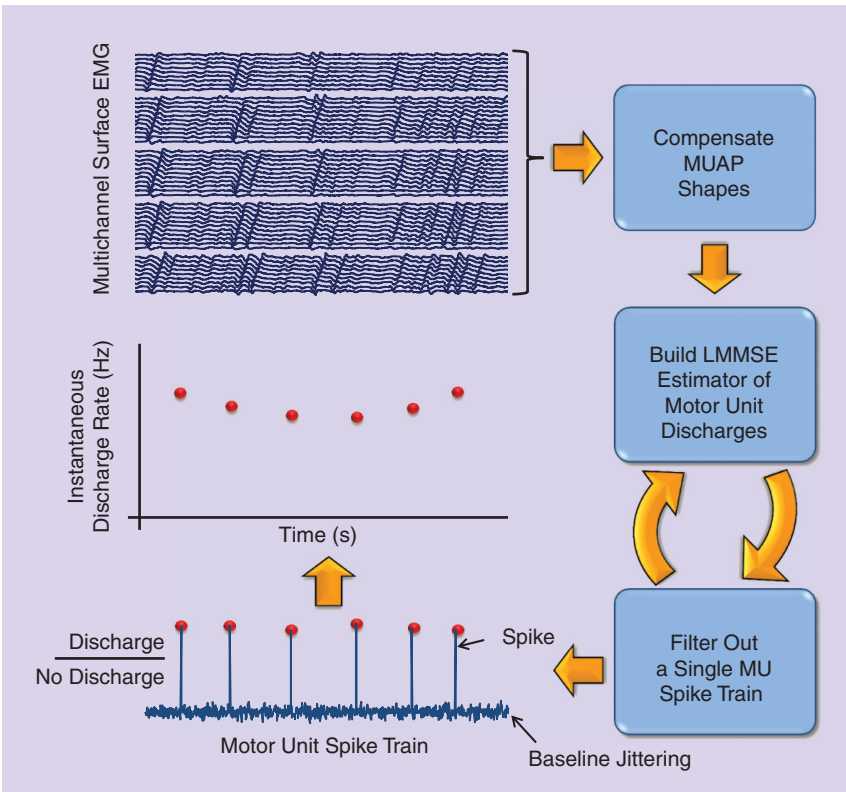
$$\hat{s}_j(t) = \mathbf{c}_{\hat{s}_j\mathrm{EMG}}^T \mathbf{C}_{\mathrm{EMG}}^{-1} \mathbf{EMG}(t), \qquad (3)$$

where $\mathbf{C}_{\mathrm{EMG}} = E(\mathbf{EMG}(t)\mathbf{EMG}^T(t))$ is the correlation matrix of the EMG signals, $\mathbf{c}_{s_j\mathrm{EMG}} = E(s_j(t)\mathbf{EMG}^T(t))$ is the cross-correlation vector between the jth MU spike train and the vector $\mathbf{EMG}(t)$, and $E(\cdot)$ stands for statistical expectation. The cross-correlation vector $\mathbf{c}_{s_j\mathrm{EMG}}$ can be estimated either by the sequential probabilistic approach introduced in [5] or by the gradient optimization technique presented in [13]. The resulting estimate of the MU spike train is exemplified in Figure 2 where the output spikes are estimates of the MU discharges, whereas the baseline jittering combines the impact of physiological and instrumental noises.

The CKC technique also allows for an efficient and reliable assessment of accuracy in the identification of each MU spike train. As demonstrated in [10], the ratio between the energies of the spikes and the baseline jittering in the spike train estimate (Figure 2) strongly correlates with the accuracy in identification of the MU discharges. Therefore, this energy ratio enables a signal-based validation of EMG decomposition and, thus, an automatic selection of the most reliable information.

HMI

Interfacing muscles with noninvasive procedures for HMI does not suffer from limitations due to biocompatibility, chronic stability or risks of infections typical for



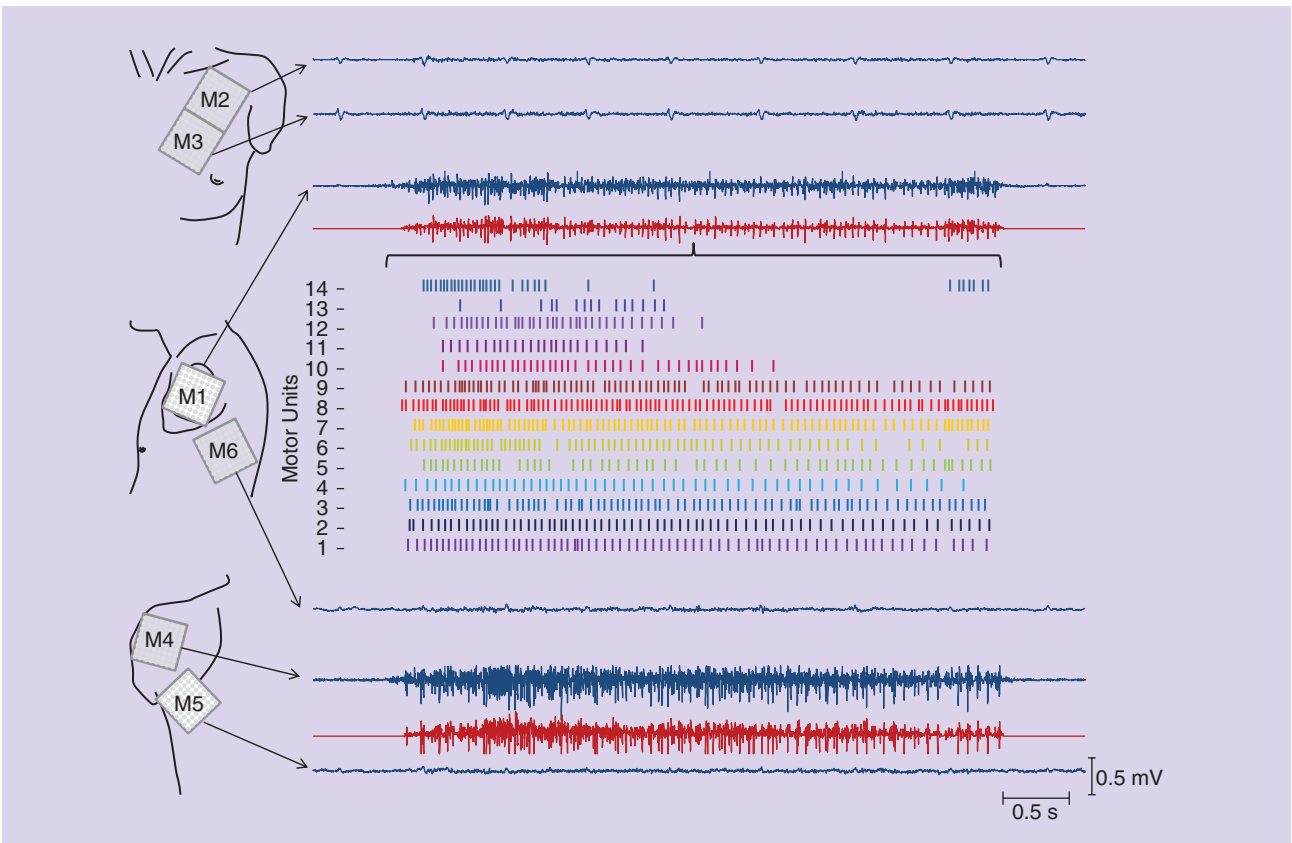
[FIG2] **A schematic representation of the working principle of the CKC algorithm for surface EMG decomposition [5]. The algorithm first compensates the MUAP shapes as detected by the grid of surface electrodes and then iteratively estimates the spike trains of individual MUs. For each identified MU, the ratio between the energy of spikes and the baseline jittering is used to efficiently estimate the accuracy in the identification of MU discharges [10].**

implanted devices and can be used for both short interventions (e.g., therapy) and chronic applications. For these reasons, with respect to other types of neural interfacing, such as direct nerve or brain recordings, the use of muscles as sources of biological activity for HMI is currently feasible in large-scale clinical applications, as, e.g., in active prostheses [11].

The classic approach for HMI with surface EMG is based on extracting global variables of the signal. For example, time-domain features from remnant muscles in amputees can be classified into motions for multifunctional advanced prosthetic devices [11]. With respect to these approaches, we previously described the possibility of decoding EMG signals into individual motor neuron activities where the spike trains of motor neurons are directly estimated. In an HMI, these spike trains can be translated into commands of external devices.

THE CLASSIC APPROACH FOR HMI WITH SURFACE EMG IS BASED ON EXTRACTING GLOBAL VARIABLES OF THE SIGNAL.

The extraction of spinal motor neuron behaviors by means of muscle recordings has several advantages over alternative approaches of interfacing the peripheral nervous system for decoding motor neuron output. For example, the SNR and quality of the EMG signal is superior to current direct nerve recordings. Moreover, the number of motor neurons whose activities can be concurrently decoded is usually larger from muscle than from nerve recordings. Although muscles may not always be present for direct interfacing, the TMR procedure allows for the decoding of any efferent nerve activity by redirecting the nerve to appropriate accessory muscles or specifically transferred muscle tissue [3]. Thus, decoding the EMG is a very general neural interface. On the whole, with respect to nerve interfacing, decoding the efferent nerve activity can be



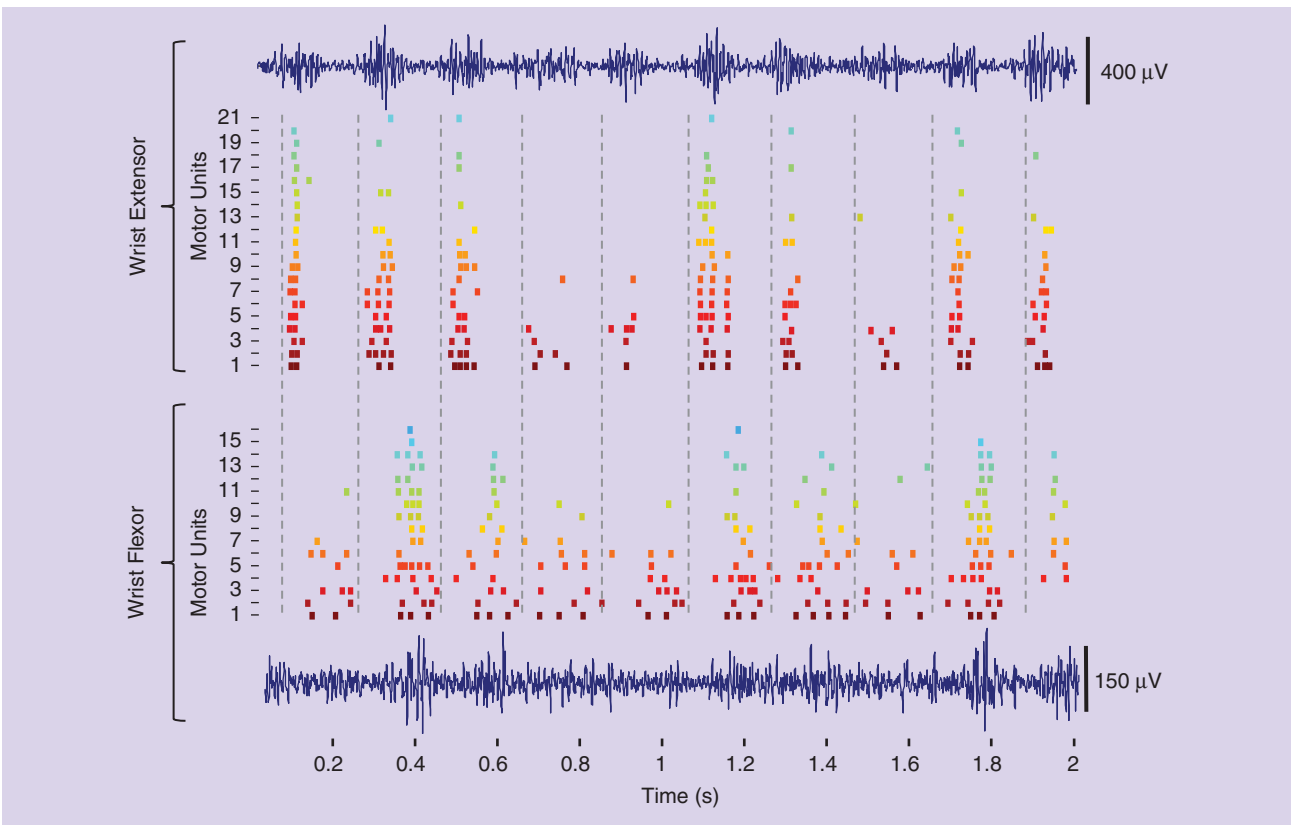
[FIG3] The decomposition of multichannel surface EMG signals recorded from a patient with a high-level amputation of the right arm due to shoulder disarticulation. The patient underwent a TMR procedure approximately one year prior to these measurements. In this representative task, the patient attempted execution of the elbow flexion with the missing limb. Surface EMG signals were recorded in single differential derivations with six grids of 64 electrodes each (8 × 8 electrodes, 10-mm interelectrode distance), labeled as M1 to M6. For reasons of clarity, only one EMG signal per grid is depicted in blue. During the task, significant muscle activity was recorded only in two of the six grids used for recording (indicated as M1 and M4). An ECG artefact is evident in the signals from grids M2 and M3. The signals from the two active grids M1 and M4 were independently decomposed by the CKC method. For clarity, the decomposition results are reported only for M1 that yielded the activity of 14 MUs (the identified MU discharges are depicted by vertical bars). The sums of MUAP trains, identified from the presented channels, are depicted in red to show qualitatively the difference between the original EMG signal and the one obtained by decomposition. (Figure reproduced from and used courtesy of [12].)

performed from muscle recordings in an easier, more robust, and noninvasive way. Of course, nerve interfacing involves the possibility of stimulating nerve afferent fibers that cannot be done directly from muscles and is beyond the scope of this article.

Figure 3 represents an example of decoding single motor neuron activity from a patient who underwent a TMR procedure. The decoded neural drive indicates the efferent signal traveling from nerves to newly innervated muscles and represents the neural code for the attempted movement—elbow flexion. The accessory muscles are used only for amplifying the nerve activity and detecting EMG signals that are then decomposed into the constituent spike trains. The decoded spike trains correspond to the natural neural commands sent by nerves to the missing muscles of the amputated limb. Therefore, the decoding of the activity from the main nerves contributing to the task allows the extraction of the last neural code for the movement, even in the case of absence of all relevant muscles for the task. This neural code can then be translated into actions or commands to run an external device, such as a robotic limb [12].

The example shown in Figure 3 represents a potential use of the proposed HMI for the control of active prostheses. It underlines the generality of the approach that can be applied to nonstationary signals extracted by muscles with different anatomy and structure [12]. A second example of use of this novel HMI will be presented next to underline the generality of the technique to conditions under which the neural information has properties different from the normal physiological activation of motor neurons.

Under pathological tremor conditions, the voluntary muscle activation is superimposed to a pathological oscillatory activity of the motor neurons. This pathological contribution is due to a combination of central oscillatory activity, reflex loops, and limb properties, generating a typical oscillatory behavior of the limb at 4–12 Hz. The behavior of motor neurons in tremulous contractions is characterized by high levels of short-term synchronization—implying a high correlation between the sources—and the presence of irregular discharges. These characteristics are detrimental for general source separation methods that often assume that the sources are independent or uncorrelated, and for which the regularity in MU discharges can be used as a priori information. Despite these complex statistical properties of the sources, accurate decomposition of the surface



[FIG4] The decomposition of multichannel surface EMG recorded with two-dimensional grids of electrodes (each comprising 12 × 5 electrodes, with 8-mm interelectrode distance) from wrist extensors and wrist flexors in an essential tremor patient during the arms outstretched task. For clarity, only one representative EMG signal per grid is depicted in blue. In total, 21 and 16 MUs were identified by the CKC algorithm from the extensors and flexors, respectively. Their discharges (denoted by colored vertical bars) are highly synchronized and demonstrate rather dissimilar bursts of muscle activity (separated by vertical dashed lines). The bursts in extensors and flexors are out of phase, causing the mechanical oscillations of the wrist. Interestingly, in flexors this pathological tremor pattern is not clearly evident from the raw EMG signals, but is revealed by surface EMG decomposition.

EMG during a pathological tremor is possible with the CKC method, as recently demonstrated [13].

Figure 4 shows an example of this approach and the detection of a relatively large population of motor neurons from signals recorded in a tremor patient. The possibility of decoding motor neuron behavior in tremor patients has opened novel perspectives in understanding the neural factors responsible for pathological tremor as well as in the diagnosis of tremor conditions. Moreover, based on the results representatively shown in Figure 4, an HMI can be built that uses the behavior of individual MUs to extract the phase and frequency of the pathological tremor oscillations online and employs this information to trigger peripheral electrical stimulation for tremor suppression (neuromodulation) in a wearable neuroprosthesis.

The two HMI examples just provided are representative of the type of information that can be extracted from muscle signals and used for controlling external devices for function substitution or neuromodulation. Several other applications of the proposed HMI concept are being explored, such as functional electrical therapy and biofeedback.

OUTLOOK

To develop more robust muscle interfaces suited for a broader range of clinical applications, there is the need for improved electrode technology. Textile systems are promising and would allow an easy placement of high-density electrode grids; other solutions, however, may be more suitable for specific applications, e.g., silicon-based electrodes to be embedded into prosthesis sockets. For chronic applications, the methods that have been described could be applied through epimysial implants of high-density electrode systems, with the advantage of minimal change in the relative position between electrodes and sources. Further developments of nonobtrusive and noninvasive acquisition systems, when combined with the signal processing techniques presented in this article, offer a considerable potential in numerous applications, such as ergonomics, rehabilitation, training of athletes, or gaming. In these applications, the information extracted from muscles can, for example, be used for the assessment of muscle activation, force, and central and peripheral fatigue.

Unlike the use of global EMG features, the decoding of groups of MUs has the advantage that the properties of the group are largely invariant with the set of detected neurons. The motor neurons receive, in fact, a large amount of common synaptic input that constitutes the effective drive to the muscle. The common input determines relative high degrees of correlation between the activation of motor neurons so that the number of motor neurons that need to be identified for understanding the behavior of the full motor neuron pool is relatively small. The exact number depends on the level of synaptic noise and on the time resolution needed in the HMI. Despite these considerations, there are no sufficient studies for estimating the number of motor neurons that are needed for a stable interface.

THE FIRST STEPS TOWARD A SOLUTION HAVE ALREADY BEEN TAKEN WITH ALGORITHMS ALLOWING FOR ONLINE DECOMPOSITION OF THE SURFACE EMG.

Finally, the mapping of the motor neuron behavior into the commands to an external device for HMI is still an open area of research. This regression problem should be addressed with the HMI user in the loop, in online studies where both the subject and the mapping adapt to each other. The first steps toward a solution have already been taken with algorithms allowing for online decomposition of the surface EMG [6]. However, open challenges remain that could be addressed with adaptive machine-learning methods supporting the bidirectional adaptation of the system and the user.

ACKNOWLEDGMENTS

The methods and perspectives described in this contribution are based on our research, which is partly funded by the ERC Advanced Grant DEMOVE (contract 267888) (DF), Slovenian Research Agency (contract L5-5550), (AH) and the EU Project NeuroTREMOR (grant agreement number ICT-2011.5.1-287739) (DF and AH).

AUTHORS

Dario Farina (dario.farina@bccn.uni-goettingen.de) is a full professor and the founding chair of the Department of Neurorehabilitation Engineering, University Medical Center Göttingen, Georg-August University, Göttingen, Germany.

Ales Holobar (ales.holobar@um.si) is an associate professor at the Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia.

REFERENCES

[1] N. V. Thakor, "Translating the brain-machine interface," *Sci. Transl. Med.,* vol. 5, no. 210, pp. 210–217, 2013.

[2] C. J. Heckman and R. M. Enoka, "Motor unit," *Compr. Physiol*., vol. 2, no. 4, pp. 2629–2682, 2012.

[3] T. A. Kuiken, L. A. Miller, R. D. Lipschutz, B. A. Lock, K. Stubblefield, P. D. Marasco, P. Zhou, and G. A. Dumanian, "Targeted reinnervation for enhanced prosthetic arm function in a woman with a proximal amputation: A case study," *Lancet*, vol. 369, pp. 371–380, Feb. 2007.

[4] D. Farina, R. Merletti, and R. M. Enoka,"The extraction of neural strategies from the surface EMG," *J. Appl. Physiol.*, vol. 96, pp. 1486–1495, Apr. 2004.

[5] A. Holobar and D. Zazula, "Multichannel blind source separation using convolution kernel compensation," *IEEE Trans. Signal Processing*, vol. 55, no. 9, pp. 4487–4496, 2007.

[6] V. Glaser, A. Holobar, and D. Zazula, "Real-time motor unit identification from high-density surface EMG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, pp. 949–958, Nov. 2013.

[7] M. S. Lewicki, "A review of methods for spike sorting: The detection and classification of neural action potentials," *Network: Computat. Neural Syst.*, vol. 9, pp. R53–R78, Nov. 1998.

[8] D. Farina, F. Negro, M. Gazzoni, and R. M. Enoka, "Detecting the unique representation of motor-unit action potentials in the surface electromyogram," *J. Neurophysiol*., vol. 100, pp. 1223–1233, Sept. 2008.

[9] I. Bankman, K. Johnson, and K. Schneider, "Optimal detection, classification, and superposition resolution in neural waveform recordings," *IEEE Trans. Biomed. Eng.*, vol. 8, pp. 836–841, Aug. 1993.

[10] A. Holobar, M. A. Minetto, and D. Farina, "Accurate identification of motor unit discharge patterns from high-density surface EMG and validation with a novel signal-based performance metric," *J. Neural Eng.*, vol. 11, p. 016008, Feb. 2014.

[11] N. Jiang, S. Dosen, K. R. Müller, and D. Farina, "Myoelectric control of artificial limbs: Is there the need for a change of focus?" [In the Spotlight] *IEEE Signal Processing Mag.*, vol. 29, no. 5, pp. 148–152, 2012.

[12] D. Farina, H. Rehbaum, A. Holobar, I. Vujaklija, N. Jiang, C. Hofer, S. Salminger, H. W. van Vliet, and O. Aszmann, "Noninvasive, accurate assessment of the behavior of representative populations of motor units in targeted reinnervated muscles," *IEEE Trans. Neural Syst. Rehabil. Eng.,* vol. 22, no. 4, pp. 810–819, July 2014.

[13] A. Holobar, V. Glaser, J. A. Gallego, J. L. Dideriksen, and D. Farina, "Non-invasive characterization of motor unit behaviour in pathological tremor," *J. Neural Eng.*, vol. 9, p. 056011, Oct. 2012.

[SP]

[social SCIENCES]

Daniel Gatica-Perez

Signal Processing in the Workplace

According to the U.S. Bureau of Labor Statistics, during 2013 employed Americans "worked an average of 7.6 hours on the days they worked," and "83% did some or all of their work at their workplace" [1]. Understanding processes in the workplace has been the subject of disciplines like organizational psychology and management for decades. In particular, the study of nonverbal communication at work is fundamental as "face-to-face interaction with superiors, subordinates, and peers consumes much of our time and energy" [2] and a variety of phenomena including job stress, rapport, and leadership can be revealed by and perceived from the tone of voice, gaze, facial expressions, and body cues of coworkers and managers [2].

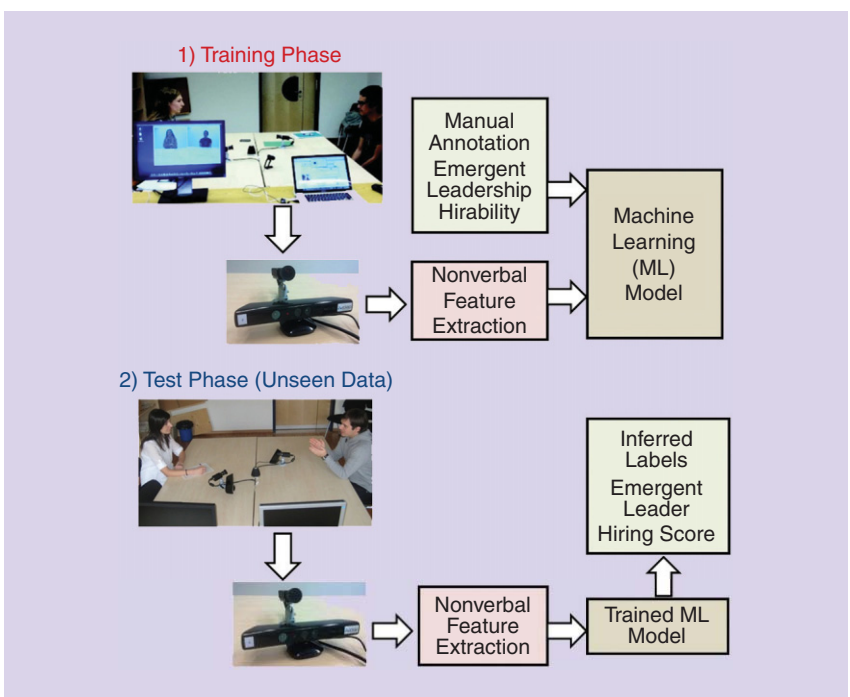
In parallel to these developments, progress in audio-visual sensing and machine perception is making the extraction of several of these nonverbal cues feasible and scalable. This trend creates opportunities toward improving the scientific understanding of phenomena in organizations and to develop technology that supports individuals and groups at work. Furthermore, it defines a domain where signal processing researchers can find new problems while working with social scientists.

In this column, a framework developed with collaborators in organizational psychology is described, aimed at inferring high-level constructs of interest in the workplace from nonverbal behavior. We summarize our experience tackling two tasks: identifying emergent leaders in small groups and assessing the hirability of candidates in employment interviews. The examples discussed in this column have been recorded in a standard lab setting [9], in which sensors are fixed in a specific environment that volunteer participants have to visit, but also in moderately in-the-wild settings, where a portable sensing solution has been used to bring participants to quiet indoor environments for recordings [5], which gives flexibility for volunteer recruits. Sensors have included Webcams and commercial microphone arrays for the portable case and high-resolution cameras and Microsoft Kinect for the lab case. As the interactions take place around a table in real workplaces, we have exploited this setting for sensor placement. One specific goal of our work with psychologists has been the deployment of the sensing lab in their institution, with the goal of promoting a wider and more frequent use of the technology in their discipline.

This column's material is adapted from [5] and [9] (refer to the original papers for details).

A FRAMEWORK FOR SOCIAL INFERENCE FROM NONVERBAL BEHAVIOR

The computational framework we have developed is shown as a diagram in Figure 1 [5], [9]. It follows a supervised machine-learning approach, where training and test phases are defined to automatically infer variables of interest (hirability in job interviews or emergent leadership in small groups) from dyadic or group interactions. At the onset, experiments are designed jointly by psychologists and engineers and



[FIG1] The computational framework to study work-related tasks.

Digital Object Identifier 10.1109/MSP.2014.2359247
Date of publication: 5 December 2014

involve the selection and deployment of sensing technology, the design of the specific interaction to be recorded, a battery of questionnaires to be completed by study participants, and human coding tasks to be completed by external observers.

Questionnaire data completed by participants and additional coding data provided by external observers are used both for psychology research and as ground-truth data for computational analysis. Questionnaires, designed and validated by psychologists, are often adapted from previous literature and administered to participants in the experiments. Additional coding data can be produced by trained psychology students or experts. The manual annotation process in Figure 1 involves the postprocessing of the above data to define ground truth in amenable form for machine-learning tasks. Concretely, hirability scores in job interviews provided by trained coders can be used to define a regression task (e.g. estimate the actual score) or a classification task (e.g., high versus low score levels); furthermore, questionnaire data provided by the participants in a group discussion about the perceived leadership of each team member can be aggregated to define the ground truth in a task whose goal is to identify one person in each group.

The nonverbal feature extraction process has involved both the development of new techniques to extract cues from audio and video and the use of existing modules. Cues related to speaking activity, prosody, body and head activity, and gaze have been used in the work described here (facial expressions have been used in other instances of our work.) The bidisciplinary approach has influenced our choices regarding the extraction of behavioral cues previously documented in psychology research with respect to their predictive value for the variables of interest (hirability or emergent leadership.) This has facilitated placing the results of our studies in the context of previous literature. At the same time, machine learning gives the possibility to extract new features, some of which might not be readily interpretable but effective for automatic inference. Moreover, the use of machine-learning methods [e.g., support vector machines (SVMs)] can spur constructive dialog with psychologists, who are less familiar with these methods and in contrast are more acquainted with classical statistical methods and especially interested in interpretable approaches.

EMERGENT LEADERSHIP IN SMALL GROUPS

In the context of groups, the so-called vertical dimension of social relations includes constructs like dominance, status, and leadership, all referring to the position that members occupy in a group [3]. In particular, research on leadership in organizational psychology and management has characterized leadership styles used to direct groups as well as emerging phenomena. Emergent leaders are individuals who rise among the members of a group and gain power from the group members themselves, instead of doing so from external entities (e.g., upper management) [4]. As much work today is done in groups, identifying emergent leaders is relevant in practice for recruitment, training, and development in organizations.

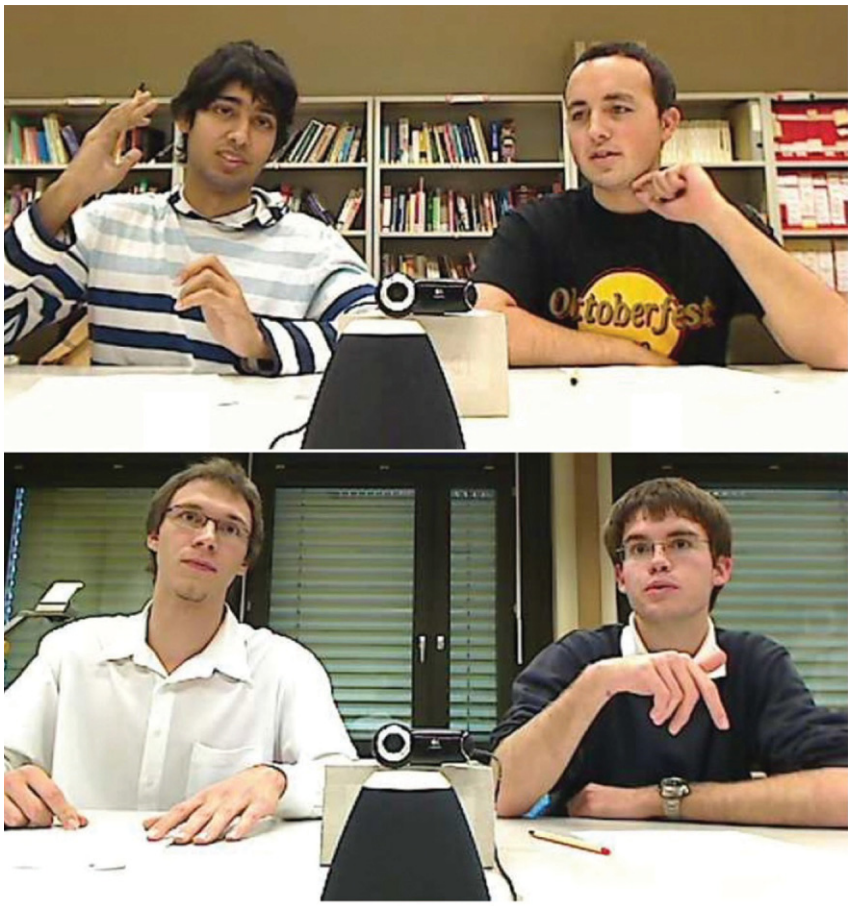
THE NONVERBAL FEATURE EXTRACTION PROCESS HAS INVOLVED BOTH THE DEVELOPMENT OF NEW TECHNIQUES TO EXTRACT CUES FROM AUDIO AND VIDEO AND THE USE OF EXISTING MODULES.

Connections between nonverbal behavior and emergent leadership have been studied for several decades [4]. While an extensive discussion is not provided here, different studies have found connections between ratings of perceived emergent leadership and manually coded cues like speaking time, arm movements, and gaze (including given and received gaze and joint patterns of looking/speaking.) Some of these cues have also been linked to dominance, a similar but not identical concept related to a tendency to control others via observable acts [3].

In [5], we followed the approach described in Figure 1 to identify the emergent leader in three- to four-person groups. We used two Webcams and a Dev-Audio Microcone microphone array as sensors. Each camera covers two people, and the Microcone provides audio for prosody feature extraction while generating a segmentation of the speech of each person (Figure 2). Groups of unacquainted people were asked to play the "winter survival task," a commonly used exercise to study group decision making and performance. In the task, participants need to rank a list of items according to their relevance for survival in a hypothetical plane crash in winter. Individuals first generate their own rankings and then discuss and collectively agree on a final list, the interaction eliciting the possible emergence of a leader. After concluding the list, participants were asked to fill out questionnaires to characterize the other group members, including variables like perceived leadership, perceived dominance, and perceived competence. The resulting emergent leadership (ELEA) corpus includes audio, video, and questionnaire data for 40 groups (148 individuals) and is publicly available for academic research.

Standard speech processing and computer vision methods were used to extract a variety of nonverbal cues. From the audio track for each participant, this included the amount of speaking time, number and average length of speaking turns, number of interruptions, speech spectral flatness, energy variation, and pitch variation. From video, features included a head activity measure obtained from a head tracker and optical flow estimates and a body activity measure based on an improvement of classic motion templates (motion energy images). Details can be found in [5]. In subsequent work [6], head pose (as a proxy for gaze) and joint looking/speaking patterns were also extracted using visual trackers based on particle filtering.

A correlation analysis of the perceived variables from the questionnaires first showed that the emergent leader was significantly perceived as a dominant person, with a second, less strong correlation effect between perceived leadership and competence. This is an interesting finding that relates different organizational constructs with one another. Furthermore, a



[FIG2] A photo from the ELEA corpus. (Photo taken from and used courtesy of [5].)

[FIG3] An interviewer and a candidate in a job interview. (Photo taken from and used courtesy of [9].)

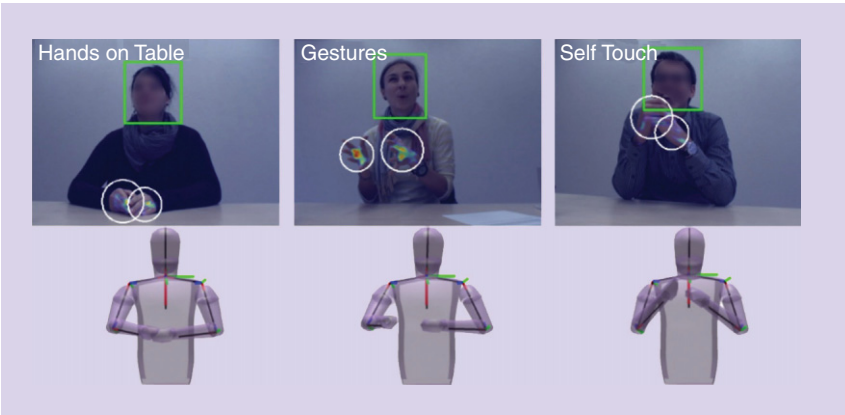
correlation analysis between the perceived questionnaire variables and the nonverbal features showed that emergent leadership is linked to participants who talk more, take more turns, interrupt more, and move their body more. This motivated the automatic recognition approach from these cues. Using standard classification techniques (SVMs or ranked feature fusion), the method identified the emergent leader in a group with accuracy between 70 and 85% depending on the modalities and classifiers used. Two results relevant for signal processing are that the cues derived from the audio track were more discriminant than the visual cues, and when combined, visual cues can bring a slight performance improvement.

HIRABILITY IN JOB INTERVIEWS

Interviews are an integral part of the recruitment process and, as such, they have been extensively studied in organizational psychology and management [7], [8]. From the social computing perspective, employment interviews are an important subject of study because of their impact on a person's life, their expressiveness, and the volume with which they are generated. Automatic analysis could be used to provide feedback to candidates, to support training programs, or to summarize large volumes of data in big organizations.

Previous literature on nonverbal communication has studied links between a number of features and job interview perceptions and outcomes. Interviewers most often do not meet the applicants in person before the interview; they interact on the basis of previous information provided by the applicant (resume, reference letters, LinkedIn profiles) and the behavior during the interview itself. Interviewers form impressions of a number of attributes of the candidate, hirability being one of them, and use these impressions and other available information to make decisions. Studies based on manually coded cues have found that candidates who are perceived as more hirable and competent (or who are actually hired) display an array of cues including smiling, eye contact, nodding, reduced interpersonal distance, body posture (oriented toward the interviewer), and specific speaking patterns [7], [8]. Taken together, this so-called immediacy behavior might convey a sense of larger availability or closeness, which as some literature suggests can lead to positive impressions on interviewers and, as a consequence, more positive assessments of candidates.

In [9], we analyzed job interviews following the approach in Figure 1. We first collected a corpus of 62 interviews where candidates applied for a real (albeit short) paid job, related to recruiting volunteers on the street for future psychology experiments. The job itself had connections to a sales position. We used Microsoft Kinect and high-resolution cameras to collect video, and the Microcone to collect audio (Figure 3). The interviews were structured (i.e., they consisted of a fixed number of questions, asked in the same order to each candidate) and behavioral (i.e., the questions were designed to elicit behavioral



[FIG4] **Interview frames with face and hands detection outputs, recognized activity, and estimated 3-D upper body pose. (Photo taken from and used courtesy of [11].)**

responses from the candidates). Interviews lasted 11 min on average. Among a variety of questionnaire and manual coding data that were collected, a hirability measure was provided by a psychology student who watched the interview using audio and video from both the candidate and the interviewer and who was trained at the task.

Regarding nonverbal cues, in addition to audio features similar to the ones described in the previous section, we developed methods to extract head nods [10] and body cues [11]. In [10], we demonstrated the advantages in terms of performance of a multimodal approach for nodding recognition, in which the observation of the (self-) speaking state of a person (speaking or silent) is used to learn two separate nodding/non-nodding classifiers, one for each speaking state. In [11], we developed a method to extract body cues from RGB video, by first detecting a person's face and hands, then inferring an approximation of the three-dimensional (3-D) pose of the upper body, and finally using this representation to do recognition of basic conversational cues like self-touch and gestures (see Figure 4). These approaches were later extended to use the depth information from Kinect.

The suite of nonverbal cues was used in [9] both for correlation analysis and a regression task, where the hirability measure provided by the trained student was the variable to be predicted. Regarding correlation, the results showed that candidates who spoke longer and faster and who took longer speaking turns received higher hirability scores. Visual features related to the amount of head motion also showed positive effects with hirability. For the regression task, using the coefficient of determination (R^2) as performance measure, the approach achieved a best result of $R^2 = 0.36$ using ridge regression and all features extracted from a candidate. This initial result shows promise, but overall the problem is challenging. As in the case of emergent leadership, cues from the audio track were more discriminative compared to video cues. Finally, some of the cues of the interviewer turned out to be predictive of hirability, which suggests that the behavior of the interacting partner can also be informative about the self, and highlights the importance to think about this problem in contextual terms.

PERSPECTIVES

This column summarized our experience studying two research problems in organizational psychology using automatically measured nonverbal behavior and machine learning. More generally, how can research at the boundaries between signal processing and organizational psychology be expanded? Three possible directions are the following.

First, we need to communicate the possibilities of multimodal signal processing and machine-learning methods within the social and organizational psychology communities, creating further partnerships where common goals can be defined and pursued. In their discipline, our collaborators have advocated for the benefits of this approach in their specific research and have shared experiences on how similar work could be incorporated into other research lines [12]. As with other examples of multidisciplinary work, there are important issues of language, methodology, expectations, and practices that need to be sorted out. Should engineers only be service providers for psychology labs? What is the level at which automation should stop? What is the value (and the place) of computational approaches for recognition that are high performing but less interpretable? What is the level of experimental control that a discipline is willing to lose to conduct experiments in the wild? These are a few questions that we have encountered in our own work.

Second, from the perspective of ubiquitous applications, interactivity is key. Some aspects of the methodology presented here could be embedded in real-time awareness tools to support sectors in industry where privacy-sensitive feedback at work would be positive. This includes hospitality, sales, and public communication. Another relevant dimension is training [13]. In addition to smartphones, the current surge of wearable devices including wristbands, smart watches, and glasses are opening new ways to sense and interact. Ethics and privacy need to be a fundamental part of future designs.

Finally, as new studies from the lab toward real workplaces become possible, computational models to handle longitudinal and relational data are needed. While lab studies are intrinsically localized in time, future work that aims at understanding teams in the workplace over days, weeks, or months require thinking about time and relations in a different way (for example, dynamic graphs with multidimensional attributes at multiple time scales.) This is a direction where signal processing methods could be especially useful, both via adaptation of existing techniques and through the development of new frameworks.

ACKNOWLEDGMENTS

The research discussed here is joint work with colleagues at the University of Lausanne, Switzerland (Marianne Schmid

Mast and Denise Frauendorfer), Idiap (Dairazalia Sanchez-Cortes, Oya Aran, Laurent Nguyen, Alvaro Marcos, Dinesh Babu Jayagopi, and Jean-Marc Odobez), and other institutions (Tanzeem Choudhury, Cornell University; Marta Marron, University of Alcala, Spain; and Daniel Pizarro, University of Auvergne, France.) I thank all of them, and acknowledge the support by the Swiss National Science Foundation (SONVB and UBImpressed projects) and the European Commission (NOVICOM project).

AUTHOR

Daniel Gatica-Perez (gatica@idiap.ch) is the head of the Social Computing Group at Idiap Research Institute and Maître d'Enseignement et de Recherche at the École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland.

REFERENCES

[1] (2014, June 18). American time use survey summary. [Online]. Available: http://www.bls.gov/news.release/atus.nr0.htm

[2] M. Remland, "Uses and consequences of nonverbal communication in the context of organizational life," in *The SAGE Handbook of Nonverbal Communication*, V. Manusov and M. Patterson, Eds. Thousand Oaks, CA: Sage Publications, 2006, pp. 501–521.

[3] J. A. Hall, E. J. Coats, and L. Smith, "Nonverbal behavior and the vertical dimension of social relations: A meta-analysis," *Psychol. Bull.*, vol. 131, no. 6, pp. 898–924, 2005.

[4] R. T. Stein, "Identifying emergent leaders from verbal and nonverbal communications," *Pers. Social Psychol.*, vol. 32, no. 1, pp. 125–135, 1975.

[5] D. Sanchez Cortes, O. Aran, M. Schmid Mast, and D. Gatica-Perez, "A nonverbal behavior approach to identify emergent leaders in small groups," *IEEE Trans. Multimedia*, vol. 14, nos. 2–3, pp. 816–832, June 2012.

[6] D. Sanchez Cortes, O. Aran, D. Jayagopi, M. Schmid Mast, and D. Gatica-Perez, "Emergent leaders through looking and speaking: From audio-visual data to multimodal recognition," *J. Multimodal User Interfaces* (Special Issue on Multimodal Corpora), vol. 7, nos. 1–2, pp. 39–53, Mar. 2013.

[7] A. S. Imada and M. D. Hakel, "Influence of nonverbal communication and rater proximity on impressions and decisions in simulated employment interviews," *Appl. Psychol.*, vol. 62, no. 3, pp. 295–300, 1977.

[8] R. J. Forbes and P. R. Jackson, "Non-verbal behaviour and the outcome of selection interviews," *Occupational Psychol.*, vol. 53, no. 1, pp. 65–72, 1980.

[9] L. S. Nguyen, D. Frauendorfer, M. Schmid Mast, and D. Gatica-Perez, "Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior," *IEEE Trans. Multimedia*, vol. 16, no. 4, pp. 1018–1031, June 2014.

[10] L. S. Nguyen, J.-M. Odobez, and D. Gatica-Perez, "Using self-context for multimodal detection of head nods in face-to-face interaction," in *Proc. ACM Int. Conf. Multimodal Interaction (ICMI)*, Santa Monica, Oct. 2012, pp. 289–292.

[11] A. Marcos-Ramiro, D. Pizarro-Perez, M. Marron-Romera, L. S. Nguyen, and D. Gatica-Perez, "Body communicative cue extraction for conversational analysis," in *Proc. IEEE Int. Conf. Face and Gesture Recognition (FG)*, Shanghai, Apr. 2013, pp. 1–8.

[12] D. Frauendorfer, M. Schmid Mast, L. S. Nguyen, and D. Gatica-Perez, "Nonverbal social sensing in action: Unobtrusive recording and extracting of nonverbal behavior in social interactions illustrated with a research example," *J. Nonverb. Behav.* (Special Issue on Contemporary Perspectives in Nonverbal Research), vol. 38, no. 2, pp. 231–245, June 2014.

[13] M. E. Hoque and R. W. Picard, "Rich nonverbal sensing to enable new possibilities in social skills training," *IEEE Comput.*, vol. 47, no. 4, pp. 28–35, Apr. 2014.

[SP]

[special REPORTS] (continued from page 14)

embeddings. Then, the challenge is how to effectively embed the full structure in the appropriate semantic space. If this is done well, the speech recognition component of the overall system will have powerful constraints to exploit, leading to the reduction of its language model's perplexity and improvement of its recognition accuracy.

Sejnoha: I think that the signal acquisition, making sense out of a very noisy world, is a very important challenge and something we have to continue working on. The fundamental modeling and modeling language—I think we're making good progress in these areas. When it comes to extraction and knowing what to do, that borders on AI. How do you define the goal of an interaction with a user in a way that it is efficient and where unexpected intelligent things happen? I think that's still a fairly novel area. You will see a lot of progress there.

The big challenge is connecting to the myriad of forms of content and services that people want to interact with, and part of that is an engineering issue and part of it is the fundamental problem of the promise of the semantic web. We have lots of stuff out there, but it is siloed, it's opaque. It doesn't advertise its capabilities, or describe its knowledge in machine understandable terms. As we get closer to the real Internet of Things, we will do better on that front. When you tell your virtual assistant to turn down your thermostat, they can talk to each other.

***IEEE SPM**: What qualifications would be needed for engineers interested in specializing in speech technology? What skill sets would be most helpful?*

Sejnoha: The field has huge multidisciplinary demands. Some background in digital signal processing and modeling is important. Of course, AI and machine learning. Also, software development. And linguistics.

RESEARCHERS INTERVIEWED

Li Deng is the principal researcher and manager of research of the Deep Learning Technology Center at Microsoft Research.

Vlad Sejnoha is the chief technology officer of Nuance Communications.

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

[SP]

advertisers INDEX

The Advertisers Index contained in this issue is compiled as a service to our readers and advertisers: the publisher is not liable for errors or omissions although every effort is made to ensure its accuracy. Be sure to let our advertisers know you found them through *IEEE Signal Processing Magazine*.

ADVERTISER	PAGE	URL	PHONE
ICIP 2015	PAGE 3	www.icip2015.org	
Mathworks	CVR 4	www.mathworks.com/ltc	+1 508 647 7040

advertising SALES OFFICES

James A. Vick
Sr. Director, Advertising
Phone: +1 212 419 7767;
Fax: +1 212 419 7589
jv.ieeemedia@ieee.org

Marion Delaney
Advertising Sales Director
Phone: +1 415 863 4717;
Fax: +1 415 863 4717
md.ieeemedia@ieee.org

Product Advertising

MIDATLANTIC
Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, PA, DE, MD, DC, KY, WV

NEW ENGLAND/SOUTH CENTRAL/ EASTERN CANADA
Jody Estabrook
Phone: +1 774 283 4528;
Fax: +1 774 283 4527
je.ieeemedia@ieee.org
ME, VT, NH, MA, RI, CT, AR, LA, OK, TX
Canada: Quebec, Nova Scotia, Newfoundland, Prince Edward Island, New Brunswick

SOUTHEAST
Thomas Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
tf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

MIDWEST/CENTRAL CANADA
Dave Jones
Phone: +1 708 442 5633;
Fax: +1 708 442 7620
dj.ieeemedia@ieee.org
IL, IA, KS, MN, MO, NE, ND, SD, WI, OH
Canada: Manitoba, Saskatchewan, Alberta

MIDWEST/ ONTARIO, CANADA
Will Hamilton
Phone: +1 269 381 2156;
Fax: +1 269 381 2556
wh.ieeemedia@ieee.org
IN, MI. Canada: Ontario

WEST COAST/MOUNTAIN STATES/ WESTERN CANADA
Marshall Rubin
Phone: +1 818 888 2407;
Fax: +1 818 888 4907
mr.ieeemedia@ieee.org
AZ, CO, HI, NM, NV, UT, AK, ID, MT, WY, OR, WA, CA. Canada: British Columbia

EUROPE/AFRICA/MIDDLE EAST ASIA/FAR EAST/PACIFIC RIM
Louise Smith
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
les.ieeemedia@ieee.org
Europe, Africa, Middle East
Asia, Far East, Pacific Rim, Australia, New Zealand

Recruitment Advertising

MIDATLANTIC
Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, CT, PA, DE, MD, DC, KY, WV

NEW ENGLAND/EASTERN CANADA
Liza Reich
Phone: +1 212 419 7578;
Fax: +1 212 419 7589
e.reich@ieee.org
ME, VT, NH, MA, RI. Canada: Quebec, Nova Scotia, Prince Edward Island, Newfoundland, New Brunswick

SOUTHEAST
Cathy Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
cf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

MIDWEST/SOUTH CENTRAL/ CENTRAL CANADA
Darcy Giovingo
Phone: +224 616 3034;
Fax: +1 847 729 4269
dg.ieeemedia@ieee.org;
AR, IL, IN, IA, KS, LA, MI, MN, MO, NE, ND, SD, OH, OK, TX, WI. Canada: Ontario, Manitoba, Saskatchewan, Alberta

WEST COAST/SOUTHWEST/ MOUNTAIN STATES/ASIA
Tim Matteson
Phone: +1 310 836 4064;
Fax: +1 310 836 4067
tm.ieeemedia@ieee.org
AZ, CO, HI, NV, NM, UT, CA, AK, ID, MT, WY, OR, WA. Canada: British Columbia

EUROPE/AFRICA/MIDDLE EAST
Louise Smith
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
les.ieeemedia@ieee.org
Europe, Africa, Middle East

Digital Object Identifier 10.1109/MSP.2014.2359761

dates AHEAD

Please send calendar submissions to:
Dates Ahead, c/o Jessica Barragué
IEEE Signal Processing Magazine
445 Hoes Lane
Piscataway, NJ 08855 USA
e-mail: j.barrague@ieee.org
(Colored conference title indicates SP-sponsored conference.)

2015

APRIL

Data Compression Conference (DCC)
7–9 April, Snowbird, Utah, United States.
URL: http://www.cs.brandeis.edu/~dcc/index.html

14th IEEE International Conference on Information Processing in Sensor Networks (IPSN)
13–17 April, Seattle, Washington, United States.
General Chair: Suman Nath
URL: http://ipsn.acm.org/2015

First IEEE Conference on Network Softwarization (NetSoft)
13–17 April, London, United Kingdom.
General Cochairs: Prosper Chemouil and George Pavlou
URL: http://sites.ieee.org/netsoft/

12th IEEE International Symposium on Biomedical Imaging (ISBI)
16–19 April, Brookyln, New York, United States.
General Chairs: Elsa Angelini and Jelena Kovacevic
URL: http://biomedicalimaging.org/2015/

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
19–24 April, Brisbane, Australia.
General Cochairs: Vaughan Clarkson and Jonathan Manton
URL: http://icassp2015.org/

MAY

31st Picture Coding Symposium (PCS)
31 May–3 June, Cairns, Australia.
General Chairs: David Taubman and Mark Pickering
URL: http://www.pcs2015.org

JUNE

Third IEEE International Workshop on Compressed Sensing Theory and Its Applications to Radar, Sonar, and Remote Sensing (CoSeRa)
22–24 June, Pisa, Italy.
General Chairs: Fulvio Gini and Joachim Ender
URL: http://www.cosera2015.iet.unipi.it/

16th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)
28 June–1 July, Stockholm, Sweden.
General Chairs: Joakim Jaldén and Björn Ottersten
URL: http://www.spawc2015.org/

IEEE International Conference on Multimedia and Expo (ICME)
29 June–3 July, Turin, Italy.
General Chairs: Enrico Magli, Stefano Tubaro, and Anthony Vetro
URL: http://www.icme2015.ieee-icme.org/index.php

JULY

Third IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)
12-15 July, Chengdu, China.
General Chairs: Yingbo Hua and Dezhong Yao
URL: http://www.chinasip2015.org/

AUGUST

12th IEEE International Conference on Advanced Video- and Signal-Based Surveillance (AVSS)
25–28 August, Karlsruhe, Germany.
General Chairs: Jürgen Beyerer and Rainer Stiefelhagen
URL: http://avss2015.org

SEPTEMBER

Sensor Signal Processing for Defence (SSPD)
9–10 September, Edinburgh, United Kingdom.
http://www.see.ed.ac.uk/drupal/udrc/sspd/

IEEE International Conference on Image Processing (ICIP)
28 September–1 October, Quebec City, Quebec, Canada.
URL: http://www.icip2015.org/

OCTOBER

IEEE International Workshop on Multimedia Signal Processing (MMSP)
19–21 October, Xiamen, China.
General Chairs: Xiao-Ping Zhang, Oscar C. Au, and Jonathan Li
URL: http://www.mmsp2015.org/

DECEMBER

IEEE 6th International Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP)
13 –16 December, Cancun, Mexico.

IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)
13–17 December, Scottsdale, Arizona, United States.
URL: http://www.asru2015.org/

IEEE Global Conference on Signal and Information Processing (GlobalSIP)
14–16 December, Orlando, Florida, United States.
General Chairs: José M. F. Moura and Dapeng Oliver Wu

2016

MARCH

41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
21–25 March, Shanghai, China.
URL: http://dmlab.sjtu.edu.cn/icassp/icassp2016.html

APRIL

Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)
10–13 July, Rio de Janeiro, Brazil.
General Chairs: Rodrigo C. de Lamare and Martin Haardt
URL: http://delamare.cetuc.puc-rio.br/sam2016/index.html

SP

Digital Object Identifier 10.1109/MSP.2014.2361039
Date of publication: 5 December 2014

IEEE SIGNAL PROCESSING SOCIETY

CONTENT GAZETTE

[ISSN 2167-5023]

JANUARY 2015

IEEE Signal Processing Society

IEEE

General Chairs
Joakim Jaldén
Björn Ottersten

Technical Program Chairs
Visa Koivunen
Mats Bengtsson

TPC
Biao Chen
Chong-Yun Chi
Philippe Ciblat
Tim Davidson
Min Dong
Feifei Gao
Kaibin Huang
Shi Jin
Erik Larsson
Amir Leshem
David Love
Matthew McKay
Chandra Murthy
Michael Rabbat
Ahmed Sadek
Mathini Sellathurai
Milica Stojanovic
Weifeng Su
Cihan Tependelenlioglu
Zhi Tian
Wolfgang Utschick
Sergiy Vorobyov
Pengfei Xia
Shengli Zhou

Special Session Chairs
Ana Peréz Neira
Mikael Skoglund

Plenary Chair
Lee Swindlehurst

Tutorial Chair
Daniel Palomar

Financial Chair
Lars Kildehøj Rasmussen

Publicity Chairs
Ragnar Thobaben
Rasmus Brandt

Local Arrangements
Magnus Jansson
Saikat Chatterjee
KTH Event

The 16th IEEE International Workshop on Signal Processing Advances in Wireless Communications, June 28 – July 1, Stockholm, Sweden

The SPAWC series of yearly workshops is dedicated to the latest advances in signal processing for wireless communications, and is sponsored by the IEEE Signal Processing Society. SPAWC 2015 will be held at the Main Campus of KTH Royal Institute of Technology in central Stockholm, Sweden, from June 28 – July 1, 2015. Carrying on the tradition of prior years, SPAWC 2015 will feature a combination of plenary talks, tutorials, invited as well as contributed poster sessions, presented in an environment that is conducive to discussions and the exchange of ideas.

Prospective authors are invited to submit papers in the following areas:

- Smart antennas and MIMO systems, large scale MIMO
- 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
- Energy efficient, green, radios
- Interference management, dynamic spectrum management
- Heterogeneous networks, small cells
- Millimeter wave, 60GHz communications
- Modeling, estimation and equalization of wireless channels
- Acquisition, synchronization, and tracking
- Compensation of transceiver front end non-idealities
- Full duplex systems
- Physical layer security
- Feedback in wireless networks
- Cognitive radio and networks
- Cooperative sensing, compressed sensing, and sparse signal processing
- Ultra-wideband radio, localization, RFID
- Cross-layer issues, joint source-channel coding, delay-limited communication
- Signal processing for optical and satellite communications
- Signal processing for nano- and molecular communications
- Emerging techniques and technologies, 5G

Full papers of up to a five-page limit should be submitted via EDAS.

Deadlines:
- **March 2, 2015 Paper submission**
- **April 20, 2015 Acceptance notification**
- **May 10, 2015 Final paper due**

IEEE reserves the right to exclude a paper from distribution after the conference (e.g. removal from IEEE Xplore) if the paper is not presented at the conference.

IEEE

www.spawc2015.org

IEEE Signal Processing Society

IEEE TRANSACTIONS ON
SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

NOVEMBER 15, 2014 VOLUME 62 NUMBER 22 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

IEEE TRANSACTIONS ON
SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

DECEMBER 1, 2014 VOLUME 62 NUMBER 23 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

IEEE

IEEE/ACM TRANSACTIONS ON

AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society | acm Association for Computing Machinery | www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed | MEDLINE U.S. National Library of Medicine

NOVEMBER 2014 | VOLUME 22 | NUMBER 11 | ITASFA | (ISSN 2329-9290)

REGULAR PAPERS

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 2329-9290) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $307.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

The Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop

10th-13th July 2016,Rio de Janeiro, Brazil

IEEE Signal Processing Society

Call for Papers

General Chairs
Rodrigo C. de Lamare,
PUC-Rio, Brazil and University of York, United Kingdom

Martin Haardt,
TU Ilmenau, Germany

Technical Chairs
Aleksandar Dogandzic,
Iowa State University, USA

Vítor Nascimento,
University of São Paulo, Brazil

Special Sessions Chair
Cédric Richard,
University of Nice, France

Publicity Chair
Maria Sabrina Greco,
Universityof Pisa, Italy

Important Dates

Special Session Proposals
29th January , 2016

Submission of Papers
26th February, 2016

Notification of Acceptance
29th April , 2016

Final Manuscript Submission
16th May, 2016

Advance Registration
16th May, 2016

Technical Program
The SAM Workshop is an important IEEE Signal Processing Society event dedicated to sensor array and multichannel signal processing. The organizing committee invites the international community to contribute with state-of-the-art developments in the field. SAM 2016 will feature plenary talks by leading researchers in the field as well as poster and oral sessions with presentations by the participants.

Welcome to Rio de Janeiro! – The workshop will be held at the Pontifical Catholic University of Rio de Janeiro, located in Gávea, in a superb area surrounded by beaches, mountains and the Tijuca National Forest, the world's largest urban forest. Rio de Janeiro is a world renowned city for its culture, beautiful landscapes, numerous tourist attractions and international cuisine. The workshop will take place during the first half of July about a month before the 2016 Summer Olympic Games when Rio will offer plenty of cultural activities and festivities, which will make SAM 2016 a memorable experience.

Research Areas
Authors are invited to submit contributions in the following areas:
- Adaptive beamforming
- Array processing for biomedical applications
- Array processing for communications
- Blind source separation and channel identification
- Computational and optimization techniques
- Compressive sensing and sparsity-based signal processing
- Detection and estimation
- Direction-of-arrival estimation
- Distributed and adaptive signal processing
- Intelligentsystems and knowledge-based signal processing
- Microphone and loudspeaker array applications
- MIMO radar
- Multi-antenna systems: multiuser MMO, massive MIMO and space-time coding
- Multi-channel imaging and hyperspectral processing
- Multi-sensor processing for smart grid and energy
- Non-Gaussian, nonlinear, and non-stationary models
- Performance evaluations with experimental data
- Radar and sonar array processing
- Sensor networks
- Source Localization, Classification and Tracking
- Synthetic aperture techniques
- Space-time adaptive processing
- Statistical modelling for sensor arrays
- Waveform diverse sensors and systems

Submission of papers – Full-length four-page papers will be accepted only electronically.

Special session proposals – They should be submitted by e-mail to the Technical Program Chairs and the Special Sessions Chair and include a topical title, rationale, session outline, contact information, and list of invited speakers.

IEEE TRANSACTIONS ON

IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

OCTOBER 2014 VOLUME 23 NUMBER 10 IIPRE4 (ISSN 1057-7149)

PAPERS

IEEE

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING

NEW!

The new IEEE Transactions on Computational Imaging seeks original manuscripts for publication. This new journal will publish research results where computation plays an integral role in the image formation process. All areas of computational imaging are appropriate, ranging from the principles and theory of computational imaging, to modeling paradigms for computational imaging, to image formation methods, to the latest innovative computational imaging system designs. Topics of interest include, but are not limited to the following:

Imaging Models and Representation

- Statistical-model based methods
- System and image prior models
- Noise models
- Graphical and tree-based models
- Perceptual models

Computational Sensing

- Coded source methods
- Structured light
- Coded aperture methods
- Compressed sensing
- Light-field sensing
- Plenoptic imaging
- Hardware and software systems

Computational Image Creation

- Sparsity-based methods
- Statistically-based inversion methods, Bayesian regularization
- Super-resolution, multi-image fusion
- Learning-based methods, Dictionary-based methods
- Optimization-based methods; proximal iterative methods, ADMM

Computational Photography

- Non-classical image capture, Generalized illumination
- Time-of-flight imaging
- High dynamic range imaging
- Focal stacks

Computational Consumer Imaging

- Cell phone imaging
- Camera-array systems
- Depth cameras

Computational Acoustic Imaging

- Multi-static ultrasound imaging
- Photo-acoustic imaging
- Acoustic tomography

Computational Microscopic Imaging

- Holographic microscopy
- Quantitative phase imaging
- Multi-illumination microscopy
- Lensless microscopy

Tomographic Imaging

- X-ray CT
- PET
- SPECT

Magnetic Resonance Imaging

- Diffusion tensor imaging
- Fast acquisition

Radar Imaging

- Synthetic aperture imaging
- Inverse synthetic imaging
- Terahertz imaging

Geophysical Imaging

- Multi-spectral imaging
- Ground penetrating radar
- Seismic tomography

Multi-spectral Imaging

- Multi-spectral imaging
- Hyper-spectral imaging
- Spectroscopic imaging

Editor-in-Chief: W. Clem Karl, Boston University.
To submit a paper go to: https://mc.manuscriptcentral.com/tci-ieee

IEEE Signal Processing Society | EMB | GRSS

IEEE TRANSACTIONS ON

INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

NOVEMBER 2014 VOLUME 9 NUMBER 11 (ISSN 1556-6013)

PAPERS

IEEE

2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'15)

The 2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'15) 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 and 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
- Source localization
- 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

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, guitar amplifier modeling
- Upmixing and downmixing

Audio and Speech Coding
- Waveform coding and parameter 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
- Spatial hearing
- Speech and audio quality assessment

www.waspaa.com
Mohonk Mountain House
New Paltz, New York
October 18-21, 2015

Workshop Committee

General Chairs
Laurent Daudet
Université Paris Diderot

Gaël Richard
Telecom ParisTech

Technical Program Chair
Bryan Pardo
Northwestern University

Finance Chair
Dorothea Kolossa
Ruhr-Universität Bochum

Far East Liaison
Nobutaka Ono
National Institute of Informatics (Japan)

Publ. Chair & Industry Liaison
John Hershey
Mitsubishi Electric Research Laboratories

Local Arrangements Chair
Juan Bello
New York University

Registration Chair
Bob L. Sturm
Queen Mary University of London

Important Dates

Submission of papers
April 10, 2015

Notification of acceptance
June 26, 2015

Early registration until
August 14, 2015

Workshop
October 18-21, 2015

IEEE

IEEE Signal Processing Society

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/

NOVEMBER 2014 VOLUME 16 NUMBER 7 ITMUF8 (ISSN 1520-9210)

PAPERS

Call For Papers
Special Issue of IEEE Transactions on Multimedia
"Multimedia: The Biggest Big Data"

SUMMARY

Multimedia is increasingly becoming the "biggest big data" as the most important and valuable source for insights and information. It covers from everyone's experiences to everything happening in the world. There will be lots of multimedia big data --- surveillance video, entertainment and social media, medical images, consumer images, voice and video, to name a few, only if their volumes grow to the extent that the traditional multimedia processing and analysis systems cannot handle effectively. As such, multimedia big data will emerge as the next "must have" competency in our society, and is spurring on tremendous amounts of research and development of related technologies and applications. As an active and inter-disciplinary research field, multimedia big data also presents a great opportunity for multimedia computing in the big data era. The challenges and opportunities highlighted in this field will foster some interesting future developments in the multimedia research and applications.

SCOPE

The goal of this special issue is to provide a premier forum for researchers working on the aforementioned multimedia big data aspects to present their recent research results. It also provides an important opportunity for multidisciplinary works connecting big data to multimedia computing. Topics of interest include, but are not limited to

- New theory and models for multimedia big data computing
- Ultra-high efficiency compression, coding and transmission for multimedia big data
- Content analysis and mining for multimedia big data
- Semantic retrieval of multimedia big data
- Deep learning and cloud computing for multimedia big data
- Green computing for multimedia big data (e.g., high efficiency storage)
- Security and privacy in multimedia big data
- Interaction, access, visualization of multimedia big data
- Multimedia big data systems
- Novel and incentive applications of multimedia big data in various fields (e.g., search, healthcare, transportation, and retail)

IMPORTANT DATES

Submission deadline: February 28, 2015
Revision due: May 31, 2015
Camera-ready manuscript due: July 21, 2015
First notification: April 28, 2015
Final notification of acceptance: July 5, 2015
Tentative publication date: August 2015

SUBMISSION PROCEDURE

Papers should be formatted according to the IEEE Transactions on Multimedia guidelines for authors (see: http://www.signalprocessingsociety.org/tmm/tmm-author-info/). By submitting/resubmitting your manuscript to this transactions, you are acknowledging that you accept the rules established for publication of manuscripts, including agreement to pay all over-length page charges, color charges, and any other charges and fees associated with publication of the manuscript. Manuscripts (both 1-column and 2-column versions are required) should be submitted electronically through the online IEEE manuscript submission system at http://mc.manuscriptcentral.com/tmm-ieee. When selecting a manuscript type, the authors must click on BigMM Special Issue. All the submitted papers will go through the same review process as that for the regular TMM paper submissions. Referees will consider originality, significance, technical soundness, clarity of exposition, and relevance to the special issue topics above.

GUEST EDITORS

Shu-Ching Chen, Florida International University, USA (chens@cs.fiu.edu)
Ramesh Jain, University of California, Irvine, USA (jain@ics.uci.edu)
Yonghong Tian, Peking University, China (yhtian@pku.edu.cn)
Haohong Wang, TCL Research America, USA (haohongwang@gmail.com)

IEEE JOURNAL OF
SELECTED TOPICS IN SIGNAL PROCESSING

IEEE Signal Processing Society

www.ieee.org/sp/index.html

OCTOBER 2014 VOLUME 8 NUMBER 5 IJSTGY (ISSN 1932-4553)

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

DECEMBER 2014 VOLUME 21 NUMBER 12 ISPLEM (ISSN 1070-9908)

LETTERS

SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags

2015 IEEE International Workshop on Multimedia Signal Processing

Xiamen, China, October 19 – October 21, 2015

http://www.mmsp2015.org

IEEE

General Chairs

Xiao-Ping Zhang – *Ryerson U, Canada*

Oscar C. Au – *HKUST, Hong Kong*

Jonathan Li – *Xiamen U, China*

Technical Chairs

Tao Mei – *Microsoft Research Asia*

Gene Cheung – *NII, Japan*

Special Session Chairs

John Paisley – *Columbia U, USA*

Yap-Peng Tan – *NTU, Singapore*

Overview Chairs

Homer Chen–*NTU, Taiwan*

Anthony Vetro – *MERL, USA*

Local Arrangement Chair

Xinghao Ding –*Xiamen U, China*

Rongrong Ji – *Xiamen U, China*

Finance Chairs

Chia-Wen Lin – *NTHU, Taiwan*

Yue Huang – *Xiamen U, China*

Publications Chairs

Vicky Zhao – *U. Alberta, Canada*

Delu Zeng – *Xiamen U, China*

Publicity Chairs

Lina Stankovic – *U. Strathclyde, UK*

Ivan Bajic – *Simon Fraser U,. Canada*

Registration Chair

Liujuan Cao – *Xiamen U, China*

Demo Chair

Wenxin Hong – *Xiamen U, China*

Industry Liaison

Alexander Loui – *Kodak, USA*

North America Liaison

Antonio Ortega, *USC, USA*

Asia Liason

Feng Wu – *USTC, China*

Europe Liaison

Fernando Pereira – *IST-IT, Portugal*

IEEE Signal Processing Society

Tentative Call for Papers

MMSP 2015 is the 17th 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 ***Heterogeneous Big Data Analytics in Multimedia*** theme. The workshop will bring together researchers and developers in multimedia signal processing and applications to share their latest achievements and explore future directions and synergies in these exciting areas.

Papers are solicited in (but not limited to) the following topics, covering this year's theme and the general scope of multimedia signal processing:

- Theories and applications for heterogeneous big media data analytics
- Semantic extraction and knowledge mining from heterogeneous big media data
- Massive-scale media detection and recognition
- Content-based analysis, retrieval and annotation for big media data
- Feature learning for heterogeneous big media data representation
- Multimedia security, forensic, privacy for big data
- Multimedia quality assessment and enhancement
- Affective computing and cross-media sentiment analysis
- Media algorithm optimization and complexity analysis
- Multimedia in economics, finance, business analytics
- Multimedia signals in geomatics
- Image/video coding and processing
- Speech/audio recognition and processing
- Multimedia communications and interactions

Top 10% Paper Award

This award is granted to as many as 10% of the total paper submissions, and is open to all accepted papers. Papers will be evaluated based on originality, technical contribution, and presentation quality during the workshop.

Paper Submission

Prospective authors should submit full-length papers of 6 pages in two-column IEEE format, including author affiliation and address, figures, tables and references, to the submission website. Only electronic submissions are accepted. Paper submission implies the intent of at least one of the authors to register and present the paper, if accepted.

Important Dates

Proposals for Special Sessions:	March 20, 2015
Submission of Paper:	May 28, 2015
Notification of acceptance:	July 6, 2015

IEEE SignalProcessing MAGAZINE

[VOLUME 32 NUMBER 1 JANUARY 2015]

QUANTITATIVE BIOIMAGING

SIGNAL PROCESSING IN LIGHT MICROSCOPY

VIDEO QUALITY ASSESSMENT

HUMAN–MACHINE INTERFACING BY DECODING THE SURFACE ELECTROMYOGRAM

ACCURACY, APPS ADVANCE SPEECH RECOGNITION

IEEE Signal Processing Society

IEEE

CONTENTS

VOLUME 32 NUMBER 1

SPECIAL SECTION—QUANTITATIVE BIOIMAGING

FEATURE

COLUMNS

DEPARTMENT

Digital Object Identifier 10.1109/MSP.2014.2364658

Call for Papers

www.icip2015.org

The International Conference on Image Processing (ICIP), sponsored by the IEEE Signal Processing Society, is the premier forum for the presentation of technological advances and research results in the fields of theoretical, experimental, and applied image and video processing. ICIP 2015, the twenty second in the series that has been held annually since 1994, brings together leading engineers and scientists in image and video processing from around the world. Research frontiers in fields ranging from traditional image processing applications to evolving multimedia and video technologies are regularly advanced by results first reported in ICIP technical sessions.

Topics include, but are not limited to:

- **Image/video coding and transmission:** Still-image and video coding, stereoscopic and 3-D coding, distributed source coding, source/channel coding, image/video transmission over wireless networks;
- **Image/video processing:** Image and video filtering, restoration and enhancement, image segmentation, video segmentation and tracking, morphological processing, stereoscopic and 3-D processing, feature extraction and analysis, interpolation and super-resolution, motion detection and estimation, color and multispectral processing, biometrics;
- **Image formation:** Biomedical imaging, remote sensing, geophysical and seismic imaging, optimal imaging, synthetic-natural hybrid image systems;
- **Image scanning, display, and printing**: Scanning and sampling, quantization and half toning, color reproduction, image representation and rendering, display and printing systems, image-quality assessment;
- **Image/video storage, retrieval, and authentication:** Image and video databases, image and video search and retrieval, multimodality image/ video indexing and retrieval, authentication and watermarking;
- **Applications:** Biomedical sciences, mobile imaging, geosciences & remote sensing, astronomy & space exploration, document image processing and analysis, other applications.

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.

Call for Tutorials: Tutorials will be held on Sunday, September 27, 2015. Proposals should be submitted by January 15, 2015 to tutorials@icip2015.org and must include title, outline of the tutorial and its motivation, short description, contact information and credentials for each presenter including name, affiliation, email, mailing address, and a two-page resume.

Call for Special Sessions: Proposals should be submitted by November 27, 2014 in a single PDF document sent to specialsessions@icip2015.org. Please include title, motivation for the special session topic, potential authors and titles of papers, as well as contact information and credentials for each organizer including name, affiliation, email, mailing address, and a short resume.

Important Dates

- Special Sessions Proposals: **27 November 2014**
- Regular Papers Submission: **15 January 2015**
- Tutorials Proposal: **15 January 2015**

Visit icip2015.org for details on paper submission, social events, no-show policy, and more.

Organizing Committee

General Co-Chairs
Jean-Luc DUGELAY
André MORIN

Technical Program Chairs
Fabrice LABEAU
Jean-Philippe THIRAN

Finances
Jean FORTIN

Plenary Sessions
Stéphane COULOMBE
Kenneth ROSE

Special Sessions
Oscar C. AU
Éric DUBOIS

Tutorials
Janusz KONRAD
André ZACCARIN

Local Arrangements
Paul FORTIER

Registration
Xavier MALDAGUE

Exhibit/Industry
Khaled EL-MALEH
Branislav KISACANIN

Publicity
Aishy AMER

Publications
Mireille BOUTIN

Electronic Media
Abdulmotaleb EL SADDIK
Benoît HUET

International Liaison
Carlo S. REGAZZONI
Wan-Chi SIU

Student Activities
Sylvie DANIEL
Guoliang FAN

Awards
Phil CHOU

Québec
Office du tourisme de Québec

Subject: IEEE Signal Processing Cup 2015 at ICASSP2015

Call for Participation: IEEE Signal Processing Cup 2015
http://icassp2015.org/signal-processing-cup-2015/

Challenge: **Heart Rate Monitoring During Physical Exercise Using Wrist-Type Photoplethysmographic (PPG) Signals**
For details of the competition project, please visit: www.zhilinzhang.com/spcup2015/

The IEEE Signal Processing Society organizes the SP Cup competition for undergraduates at ICASSP2015. This competition aims to provide undergraduate students with the opportunity to form teams and work together to solve a challenging and interesting real-world problem using signal-processing techniques. Three teams will be selected to present their work, and the prizes will be awarded at ICASSP 2015.

You are very welcome to participating in the competition. Please also help us to circulate this email to other colleagues or students you know who may be interested in this competition.

Participation in the Competition:

Each team participating in the competition is to be composed of one faculty member (whose role is the supervisor of the team members), at most one graduate student (who will assist the supervisor in supervising the undergraduate team members), and at least 3 but no more than 10 undergraduates. At least three of the undergraduate team members must be either IEEE SP members or student members.

Participating teams must submit their project by February 6, 2015. Each submission should include a report, in the form of an IEEE conference paper, on the technical details of the methods used and the results, as well as the programs developed (MATLAB is preferred). Participating teams must register to join the competition by January 16, 2015. The online registration system will be open in September 2014. By February 27, 2015, the best 3 teams will be identified to participate in the final competition at ICASSP2015.

Important Dates:

January 16, 2015 (Friday): Team registration to join the SP Cup competition
February 6, 2015 (Friday): Submission deadline for participating teams
February 27, 2015 (Friday): Announcement of the best 3 teams
April 20, 2015: Final competition at ICASSP 2015

Team Prizes:

The champion:	$5,000
The first runner-up:	$2,500
The second runner-up:	$1,500

Each team invited to ICASSP2015 will have their travel expenses supported by the SP Society. Each team member is offered up to $1200 for continental travel, or $1,700 for intercontinental travel, and at most 3 people from each team will be supported.

Enquiries:

Technical problems: zhilinzhang@ieee.org
General enquiry: sp-enq-spcup@ieee.org

Organizers:

Bio Imaging and Signal Processing Technical Committee (BISP TC)
IEEE SPS Student Services Committee

12TH IEEE INTERNATIONAL CONFERENCE ON

ADVANCED VIDEO AND SIGNAL-BASED SURVEILLANCE.

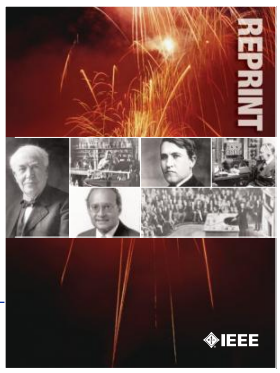
AUGUST 25-28 2015

KARLSRUHE GERMANY.

GENERAL CHAIR	JÜRGEN BEYERER, FRAUNHOFER IOSB / KIT, GERMANY.
	RAINER STIEFELHAGEN, KIT, GERMANY.
ORGANIZATION CHAIR	EDUARDO MONARI, FRAUNHOFER IOSB, GERMANY.
PROGRAM CHAIRS	ROMAN PFLUGFELDER, TU GRAZ/AIT, AUSTRIA.
	PETER TU, GE GLOBAL RESEARCH, USA.
WORKSHOPS CHAIR	RITA CUCCHIARA, UNIV. OF MODENA/REGGIO EMILIA, ITALY.
	JORDI GONZALES, UA DE BARCELONA, SPAIN.
CHALLENGE CHAIR	JAMES FERRYMAN, UNIV. OF READING, UK.
PUBLICITY CHAIRS	MARION STAUB, FRAUNHOFER IOSB, GERMANY.
PUBLICATIONS CHAIR	FRANCOIS BREMOND, INRIA, FRANCE.
	JEAN-MARC ODOBEZ, IDIAP/EPFL, SWITZERLAND.
LOCAL ARRANGEMENTS	UWE BÖHM, KIT, GERMANY.
& FINANCIAL CHAIR	MARION STAUB, FRAUNHOFER IOSB, GERMANY.
AREA CHAIRS	FATIH PORIKLI, NICTA, AUSTRALIA.
	MICHAEL ARENS, FRAUNHOFER IOSB, GERMANY.
	BIR BHANU, UC RIVERSIDE, USA.
	RAINER STIEFELHAGEN, KIT, GERMANY.
	JÖRG HÄHNER, UNIV. OF AUGSBURG, GERMANY.
	JEAN-LUC DUGELAY, EURECOM, FRANCE.
STEERING COMMITTEE	FATIH PORIKLI, NICTA, AUSTRALIA.
	CARLO REGAZZONI, UNIV. OF GENUA, ITALY.
	ANDREA CAVALLARO, QUEEN MARY UNIV. OF LONDON, UK.
	MASSIMO PICCARDI, UNIV. OF TECH. SYDNEY, AUSTRALIA.
	KOSTAS PLATANIOTIS, UNIV. OF TORONTO, CANADA.
	BERNHARD RINNER, UNIV. OF KLAGENFURT, AUSTRIA.
	PETER TU, GE GLOBAL RESEARCH, USA.
	STEFANO TUBARO, POLITECNICO DIE MILANO, ITALY.

INFO@AVSS2015.ORG

WWW.AVSS2015.ORG



IEEE ORDER FORM FOR REPRINTS

Purchasing IEEE Papers in Print is easy, cost-effective and quick.
Complete this form, send via our secure fax (24 hours a day) to 732-981-8062 or mail it back to us.

PLEASE FILL OUT THE FOLLOWING

Author: _____

Publication Title: _____

Paper Title: _____

_____

RETURN THIS FORM TO:
IEEE Publishing Services
445 Hoes Lane
Piscataway, NJ 08855-1331

Email the Reprint Department at reprints@ieee.org for questions regarding this form

PLEASE SEND ME

■ 50 ■ 100 ■ 200 ■ 300 ■ 400 ■ 500 or _____(in multiples of 50) reprints.
■ YES ■ NO Self-covering/title page required. COVER PRICE: $74 per 100, $39 per 50.
■ $58.00 Air Freight must be added for all orders being shipped outside the U.S.
■ $21.50 must be added for all USA shipments to cover the cost of UPS shipping and handling.

PAYMENT

■ Check enclosed. Payable on a bank in the USA.
■ Charge my: ■ Visa ■ Mastercard ■ Amex ■ Diners Club

Account # _____ Exp. date _____

Cardholder's Name (please print): _____

_____

■ Bill me (you must attach a purchase order) Purchase Order Number _____

Send Reprints to:

_____

_____

_____

_____

Bill to address, if different:

_____

_____

_____

_____

Because information and papers are gathered from various sources, there may be a delay in receiving your reprint request. This is especially true with postconference publications. Please provide us with contact information if you would like notification of a delay of more than 12 weeks.

Telephone: _____ Fax: _____ Email Address: _____

2012 REPRINT PRICES (without covers)

Number of Text Pages

	1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44	45-48
50	$129	$213	$245	$248	$288	$340	$371	$408	$440	$477	$510	$543
100	$245	$425	$479	$495	$573	$680	$742	$817	$885	$953	$1021	$1088

Larger quantities can be ordered. Email reprints@ieee.org with specific details.

Tax Applies on shipments of regular reprints to CA, DC, FL, MI, NJ, NY, OH and Canada (GST Registration no. 12534188).
Prices are based on black & white printing. Please call us for full color price quote, if applicable.

2015 IEEE MEMBERSHIP APPLICATION

(students and graduate students must apply online)

IEEE
Advancing Technology for Humanity

Start your membership immediately: Join online www.ieee.org/join

Please complete both sides of this form, typing or **printing in capital letters**. Use only English characters and abbreviate only if more than 40 characters and spaces per line. We regret that incomplete applications cannot be processed.

1 Name & Contact Information

Please PRINT your name as you want it to appear on your membership card and IEEE correspondence. As a key identifier for the IEEE database, circle your last/surname.

☐ Male ☐ Female Date of birth (Day/Month/Year) ____/____/____

Title First/Given Name Middle Last/Family Surname

▼ Primary Address ☐ Home ☐ Business (All IEEE mail sent here)

Street Address

City State/Province

Postal Code Country

Primary Phone

Primary E-mail

▼ Secondary Address ☐ Home ☐ Business

Company Name Department/Division

Street Address City State/Province

Postal Code Country

Secondary Phone

Secondary E-mail

To better serve our members and supplement member dues, your postal mailing address is made available to carefully selected organizations to provide you with information on technical services, continuing education, and conferences. Your e-mail address is not rented by IEEE. Please check box only if you do not want to receive these postal mailings to the selected address. ☐

2 Attestation

I have graduated from a three- to five-year academic program with a university-level degree.
☐ Yes ☐ No

This program is in one of the following fields of study:
☐ Engineering
☐ Computer Sciences and Information Technologies
☐ Physical Sciences
☐ Biological and Medical Sciences
☐ Mathematics
☐ Technical Communications, Education, Management, Law and Policy
☐ Other (please specify): _____

This academic institution or program is accredited in the country where the institution is located. ☐ Yes ☐ No ☐ Do not know

I have _____ years of professional experience in teaching, creating, developing, practicing, or managing within the following field:

☐ Engineering
☐ Computer Sciences and Information Technologies
☐ Physical Sciences
☐ Biological and Medical Sciences
☐ Mathematics
☐ Technical Communications, Education, Management, Law and Policy
☐ Other (please specify): _____

3 Please Tell Us About Yourself

Select the numbered option that best describes yourself. This information is used by IEEE magazines to verify their annual circulation. Please enter numbered selections in the boxes provided.

A. Primary line of business → ☐

1. Computers
2. Computer peripheral equipment
3. Software
4. Office and business machines
5. Test, measurement and instrumentation equipment
6. Communications systems and equipment
7. Navigation and guidance systems and equipment
8. Consumer electronics/appliances
9. Industrial equipment, controls and systems
10. ICs and microprocessors
11. Semiconductors, components, sub-assemblies, materials and supplies
12. Aircraft, missiles, space and ground support equipment
13. Oceanography and support equipment
14. Medical electronic equipment
15. OEM incorporating electronics in their end product (not elsewhere classified)
16. Independent and university research, test and design laboratories and consultants (not connected with a mfg. co.)
17. Government agencies and armed forces
18. Companies using and/or incorporating any electronic products in their manufacturing, processing, research or development activities
19. Telecommunications services, telephone (including cellular)
20. Broadcast services (TV, cable, radio)
21. Transportation services (airline, railroad, etc.)
22. Computer and communications and data processing services
23. Power production, generation, transmission and distribution
24. Other commercial users of electrical, electronic equipment and services (not elsewhere classified)
25. Distributor (reseller, wholesaler, retailer)
26. University, college/other educational institutions, libraries
27. Retired
28. Other_____

B. Principal job function → ☐

1. General and corporate management
2. Engineering management
3. Project engineering management
4. Research and development management
5. Design engineering management —analog
6. Design engineering management —digital
7. Research and development engineering
8. Design/development engineering —analog
9. Design/development engineering—digital
10. Hardware engineering
11. Software design/development
12. Computer science
13. Science/physics/mathematics
14. Engineering (not elsewhere specified)
15. Marketing/sales/purchasing
16. Consulting
17. Education/teaching
18. Retired
19. Other_____

C. Principal responsibility → ☐

1. Engineering and scientific management
2. Management other than engineering
3. Engineering design
4. Engineering
5. Software: science/mngmnt/engineering
6. Education/teaching
7. Consulting
8. Retired
9. Other_____

D. Title → ☐

1. Chairman of the Board/President/CEO
2. Owner/Partner
3. General Manager
4. VP Operations
5. VP Engineering/Dir. Engineering
6. Chief Engineer/Chief Scientist
7. Engineering Management
8. Scientific Management
9. Member of Technical Staff
10. Design Engineering Manager
11. Design Engineer
12. Hardware Engineer
13. Software Engineer
14. Computer Scientist
15. Dean/Professor/Instructor
16. Consultant
17. Retired
18. Other_____

Are you now or were you ever a member of IEEE?
☐ Yes ☐ No If yes, provide, if known:

Membership Number Grade Year Expired

4 Please Sign Your Application

I hereby apply for IEEE membership and agree to be governed by the IEEE Constitution, Bylaws, and Code of Ethics. I understand that IEEE will communicate with me regarding my individual membership and all related benefits. **Application must be signed.**

Signature Date

Over Please

Information for Authors

(Updated/Effective September 17, 2014)

The IEEE TRANSACTIONS ON SIGNAL PROCESSING is published online twice per month (semimonthly) covering advances in the theory and application of signal processing. The scope is reflected in the EDICS: the Editor's Information and Classification Scheme. Please consider the journal with the most appropriate scope for your submission.

Authors are encouraged to submit manuscripts of Regular papers (papers which provide a complete disclosure of a technical premise), or Comment Correspondences (brief items that provide comment on a paper previously published in the TRANSACTIONS). Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere.

Every manuscript must (a) provide a clearly defined statement of the problem being addressed, (b) state why it is important to solve the problem, and (c) give an indication as to how the current solution fits into the history of the problem, including bibliographic references to related work rather than restating established algorithms and scientific principles.

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, no new experimental data or new application. Experimental contributions will be rejected without review if there is insufficient experimental data. The TRANSACTIONS are published in English. Papers that have a large number of typographical and/or grammatical errors will also be rejected without review.

By submission/resubmission of your manuscript to this TRANSACTIONS, you are acknowledging that you accept the rules established for publication of manuscripts, including agreement to pay all overlength page charges, color charges, and any other charges and fees associated with publication of the manuscript. Such charges are not negotiable and cannot be suspended.

New and revised manuscripts should be prepared following the "Manuscript Submission" guidelines below, and submitted to the online manuscript system ScholarOne Manuscripts. After acceptance, finalized manuscripts should be prepared following the "Final Manuscript Submission Guidelines" below. Do not send original submissions or revisions directly to the Editor-in-Chief or Associate Editors; they will only access your manuscript electronically via the ScholarOne Manuscripts 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. For a regular paper, the manuscript may not exceed 13 double-column pages, including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references.

   Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOneManuscripts site. You are encouraged to also submit a single-column, double-spaced version (11 point font or larger), but page length restrictions will be determined by the double-column version.

   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.

   Proofread your submission, confirming that all figures and equations are visible in your document before you "SUBMIT" your manuscript. Proofreading is critical; once you submit your manuscript, the manuscript cannot be changed in any way. You may also submit your manuscript as a PostScript or MS Word file. The system has the capability of converting your files to PDF, however it is your responsibility to confirm that the conversion is correct and there are no font or graphics issues prior to completing the submission process.
3. *Additional Documents for Review*. Please upload pdf versions of all items in the reference list which are not publicly available, such as unpublished (submitted) papers. Other materials for review such as supplementary tables and figures may be uploaded as well. Reviewers will be able to view these files only if they have the appropriate software on their computers. Use short filenames without spaces or special characters. When the upload of each file is completed, you will be asked to provide a description of that file.
4. *Multimedia Materials*. IEEE Xplore can publish multimedia files (audio, images, video) and Matlab code along with your paper. Alternatively, you can provide the links to such files in a README file that appears on Xplore along with your paper. For details, please see http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect6 under "Multimedia." To make your work reproducible by others, the TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper. Files that are to be included with the final paper must be uploaded for consideration in the review process.
5. *Submission.* After uploading all files and proofreading them, submit your manuscript by clicking "Submit." A confirmation of the successful submission will open on screen containing the manuscript tracking number and will be followed with an e-mail confirmation to the corresponding and all contributing authors. Once you click "Submit," your manuscript cannot be changed in any way.
6. *Copyright Form and Consent Form.* By policy, IEEE owns the copyright to the technical contributions it publishes on behalf of the interests of the IEEE, its authors, and their employers; and to facilitate the appropriate reuse of this material by others. To comply with the IEEE copyright policies, authors are required to sign and submit a completed "IEEE Copyright and Consent Form" prior to publication by the IEEE.

   The IEEE recommends authors to use an effective electronic copyright form (eCF) tool within the ScholarOne Manuscripts system. You will be redirected to the "IEEE Electronic Copyright Form" wizard at the end of your original submission; please simply sign the eCF by typing your name at the proper location and click on the "Submit" button.

Comment Correspondence. Comment Correspondences provide brief comments on material previously published in the TRANSACTIONS. A comment correspondence may not exceed 2 pages in double-column, single double-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. Comment Correspondences are submitted in the same way as regular manuscripts (see "Manuscript Submission" above for instructions).

Manuscript Length. Papers published on or after 1 January 2007 can now be up to 10 pages, and any paper in excess of 10 pages will be subject to over length page charges. The IEEE Signal Processing Society has determined that the standard manuscript length shall be no more than 10 published pages (double-column format, 10 point type) for a regular submission. Manuscripts that exceed these limits will incur mandatory over length page charges, as discussed below. Since changes recommended as a result of peer review may require additions to the manuscript, it is strongly recommended that you practice economy in preparing original submissions.

Exceptions to 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 over length 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 related to a manuscript previously rejected by any journal, you are expected to submit supporting documents identifying the previous submission and detailing how issues raised in the previous reviews have been addressed. 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.

Full details of the resubmission process can be found in the Signal Processing Society "Policy and Procedures Manual" at http://www.signalprocessingsociety.org/about/governance/policy-procedure/.

Author Misconduct.

Author Misconduct Policy: Plagiarism includes copying someone else's work without appropriate credit, using someone else's work without clear delineation of citation, and the uncited reuse of an authors previously published work that also involves other authors. Plagiarism is unacceptable.

Self-plagiarism involves the verbatim copying or reuse of an authors own prior work without appropriate citation; it is also unacceptable. Self-plagiarism includes duplicate submission of a single journal manuscript to two different journals, and submission of two different journal manuscripts which overlap substantially in language or technical contribution.

Authors may only submit original work that has not appeared elsewhere in a journal publication, nor is under review for another journal publication. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper. If authors have used their own previously published work as a basis for a new submission, they are required to cite the previous work and very briefly indicate how the new submission offers substantively novel contributions beyond those of the previously published work.

It is acceptable for conference papers to be used as the basis for a more fully developed journal submission. Still, authors are required to cite related prior work; the papers cannot be identical; and the journal publication must include novel aspects.

Digital Object Identifier 10.1109/TSP.2014.2358875

Author Misconduct Procedures: The procedures that will be used by the Signal Processing Society in the investigation of author misconduct allegations are described in the IEEE SPS Policies and Procedures Manual.

Author Misconduct Sanctions: The IEEE Signal Processing Society will apply the following sanctions in any case of plagiarism, or in cases of self-plagiarism that involve an overlap of more than 25% with another journal manuscript:

1) immediate rejection of the manuscript in question;
2) immediate withdrawal of all other submitted manuscripts by any of the authors, submitted to any of the Society's publications (journals, conferences, workshops), except for manuscripts that also involve innocent co-authors; immediate withdrawal of all other submitted manuscripts by any of the authors, submitted to any of the Society's publications (journals, conferences, workshops), except for manuscripts that also involve innocent co-authors;
3) prohibition against each of the authors for any new submissions, either individually, in combination with the authors of the plagiarizing manuscript, or in combination with new co-authors, to all of the Society's publications (journals, conferences, workshops). The prohibition shall continue for one year from notice of suspension.

Further, plagiarism and self-plagiarism may also be actionable by the IEEE under the rules of Member Conduct.

Submission Format.

Authors are encouraged to prepare manuscripts employing the on-line style files developed by IEEE. All manuscripts accepted for publication will require the authors to make final submission employing these style files. The style files are available on the web at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2 under "Template for all TRANSACTIONS." (LaTeX and MS Word).

Authors using LaTeX: the two PDF versions of the manuscript needed for submission can both be produced by the IEEEtran.cls style file. A double-spaced document is generated by including \documentclass[11pt,draftcls,onecolumn]{IEEEtran} as the first line of the manuscript source file, and a single-spaced double-column document for estimating the publication page charges via \documentclass[10pt,twocolumn,twoside]{IEEEtran} for a regular submission, or \documentclass[9pt,twocolumn,twoside]{IEEEtran} for a Correspondence item.

- *Title page and abstract:* The first page of the manuscript shall contain the title, names and contact information for all authors (full mailing address, institutional affiliations, phone, fax, and e-mail), the abstract, and the EDICS. An asterisk * should be placed next to the name of the Corresponding Author who will serve as the main point of contact for the manuscript during the review and publication processes.

  An abstract must be a well-written stand-alone paragraph 150-250 words long, with no displayed equations, footnotes, references or tabular material. The abstract should indicate the scope of the paper and summarize the author's conclusions, making it a useful tool for information retrieval. Visit http://www.signalprocessingsociety.org/publications/periodicals/tsp/tsp-author-info/ for specifications and description.
- *EDICS:* All submissions must be classified by the author with an EDICS (Editors' Information Classification Scheme) selected from the list of EDICS published online at http://www.signalprocessingsociety.org/publications/periodicals/tsp/TSP-EDICS/
- NOTE: EDICS are necessary to begin the peer review process. Upon submission of a new manuscript, please choose the EDICS categories that best suit your manuscript. Failure to do so will likely result in a delay of the peer review process.
- The EDICS category should appear on the first page—i.e., the title and abstract page—of the manuscript.
- *Illustrations and tables:* Each figure and table should have a caption that is intelligible without requiring reference to the text. Illustrations/tables may be worked into the text of a newly-submitted manuscript, or placed at the end of the manuscript. (However, for the final submission, illustrations/tables must be submitted separately and not interwoven with the text.)

  Illustrations in color may be used but, unless the final publishing will be in color, the author is responsible that the corresponding grayscale figure is understandable.

  In preparing your illustrations, note that in the printing process, most illustrations are reduced to single-column width to conserve space. This may result in as much as a 4:1 reduction from the original. Therefore, make sure that all words are in a type size that will reduce to a minimum of 9 points or 3/16 inch high in the printed version. Only the major grid lines on graphs should be indicated.
- *Abbreviations:* This TRANSACTIONS follows the practices of the IEEE on units and abbreviations, as outlined in the Institute's published standards. See http://www.ieee.org/portal/cms_docs_iportals/iportals/publications/authors/transjnl/auinfo07.pdf for details.
- *Mathematics:* All mathematical expressions must be legible. Do not give derivations that are easily found in the literature; merely cite the reference.

Final Manuscript Submission Guidelines.

Upon formal acceptance of a manuscript for publication, instructions for providing the final materials required for publication will be sent to the Corresponding Author. Finalized manuscripts should be prepared in LaTeX or MS Word, and are required to use the style files established by IEEE, available at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2.

Instructions for preparing files for electronic submission are as follows:

- For regular papers, the final 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. Without expressed approval from the Editor-in-Chief, papers that exceed 16 pages in length will not publish.
- Files must be self-contained; that is, there can be no pointers to your system setup.
- Include a header to identify the name of the TRANSACTIONS, the name of the author, and the software used to format the manuscript.
- Do not import graphics files into the text file of your finalized manuscript (although this is acceptable for your initial submission). If submitting on disk, use a separate disk for graphics files.
- Do not create special macros.
- Do not send PostScript files of the text.
- File names should be lower case.
- Graphics files should be separate from the text, and not contain the caption text, but include callouts like "(a)," "(b)."
- Graphics file names should be lower case and named fig1.eps, fig2.tif, etc.
- Supported graphics types are EPS, PS, TIFF, or graphics created using Word, Powerpoint, Excel or PDF. Not acceptable is GIF, JPEG, WMF, PNG, BMP or any other format (JPEG is accepted for author photographs only). The provided resolution needs to be at least 600 dpi (400 dpi for color).
- Please indicate explicitly if certain illustrations should be printed in color; note that this will be at the expense of the author. Without other indications, color graphics will appear in color in the online version, but will be converted to grayscale in the print version.

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, please visit the IEEE Author Digital Tool Box at the following URL: http://www.ieee.org/publications_standards/publications/authors/auth_names_native_lang.pdf

Additional instructions for preparing, verifying the quality, and submitting graphics and multimedia files are available via http://www.ieee.org/publications_standards/publications/authors/authors_journals.html.

Open Access.

This 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 over-length 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 (six pages, in the case of Technical Correspondences until their publication will be discontinued).

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 technical correspondence until their publication will be discontinued. These are mandatory page charges and the author(s) will be held responsible for them. They are not negotiable or voluntary. The author(s) signifies his willingness to pay these charges simply by submitting his/her/their manuscript to the TRANSACTIONS. The Publisher holds the right to withhold publication under any circumstance, as well as publication of the current or future submissions of authors who have outstanding mandatory page charge debt.

Color Charges. Color figures which appear in color only in the electronic (Xplore) version can be used free of charge. In this case, the figure will be printed in the hardcopy version in grayscale, and the author is responsible that the corresponding grayscale figure is intelligible. Color reproduction in print is expensive, and all charges for color are the responsibility of the author. The estimated costs are as follows. There will be a charge of $62.50 for each figure; this charge may be subject to change without notification. In addition, there are printing preparation charges which may be estimated as follows: color reproductions on four or fewer pages of the manuscript: a total of approximately $1045; color reproductions on five pages through eight pages: a total of approximately $2090; color reproductions on nine through 12 pages: a total of approximately $3135, and so on. Payment of fees on color reproduction is not negotiable or voluntary, and the author's agreement to publish the manuscript in the TRANSACTIONS is considered acceptance of this requirement.

Information for Authors

(Updated/Effective September 2014)

The IEEE SIGNAL PROCESSING LETTERS is a monthly, archival publication designed to provide rapid dissemination of original, cutting-edge ideas and timely, significant contributions in signal, image, speech, language and audio processing.

The scope of the journal is reflected in the EDICS, listed in the page: http://www.signalprocessingsociety.org/publications/periodicals/letters/letters-edics/. These EDICS are used as the Editor's Information and Classification Scheme. Please consider the journal with the most appropriate scope for your submission.

Every manuscript must (a) provide a clearly defined statement of the problem being addressed, (b) state why it is important to solve the problem, and (c) give an indication as to how the current solution fits into the history of the problem, including bibliographic references to related work rather than restating established algorithms and scientific principles.

All submissions are prescreened prior to the peer review process to verify if they meet the publication criteria explained next. The Editor in Chief determines if the paper will be assigned to an Associate Editor, who will handle the peer review process and make a final decision, or if it becomes a candidate for an Immediate Rejection (IR), which is the case if the paper is: (a) out of scope for the journal (a paper with few or no citations to recent signal processing publications is a good suspect); (b) limited in novelty (i.e. there is no new theory, algorithms, or experimental data); (c) incomplete, and based on insufficient numerical (simulation/experiment) data and comparisons, inadequate bibliography, or poor description of the context of the work; (d) incomprehensible, poorly written (e.g. has a large number of grammatical errors, illegible figures, etc.).

Furthermore, a paper that is a resubmission (revised version) of a previously rejected manuscript from any journal must include a supporting document indicating the previous submission and summarizing substantial changes made to address reviewer comments in order to be considered for publication. Otherwise, it is a candidate for an Immediate Reject exclusively on the basis of this omission.

The Editor in Chief decisions regarding the paper are supported by recommendations by the Senior Area Editor who prescreens the paper, or by recommendations coming from the Associate Editor selected to handle the peer review. By submission/resubmission of your manuscript to the LETTERS, you are acknowledging that you accept the rules established for publication of manuscripts.

New and revised manuscripts should be prepared following the New Manuscript Submission guidelines below, and submitted to the online submission system ScholarOne Manuscripts. Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere. After acceptance, finalized manuscripts should be prepared following the Final Manuscript Submission Guidelines below. Do not send original submissions or revisions directly to the Editor-in-Chief or Associate Editors; they will only access your manuscript electronically via the ScholarOne Manuscripts 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. All co-authors must have valid private ScholarOne accounts, and manuscripts lacking this will be summarily rejected.
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. For a LETTERS submission, the main body of the manuscript may not exceed 4 double-column single-spaced pages, including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs. A fifth page containing references is permitted. Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOneManuscripts site. You are encouraged to also submit a single-column, double-spaced version (11 point font or larger), but page length restrictions will be determined by the double-column version. The section below entitled "Submission Format" provides further details and support for generating a paper in the correct format.

   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 Post-Script 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. *Additional Documents for Review:* Please upload pdf versions of all items in the reference list which are not publicly available, such as unpublished (submitted) papers. Other materials for review such as supplementary tables and figures may be uploaded as well. Reviewers will be able to view these files only if they have the appropriate software on their computers. Use short filenames without spaces or special characters. When the upload of each file is completed, you will be asked to provide a description of that file.
4. *Multimedia Materials:* IEEE Xplore can publish multimedia files (audio, images, video) and Matlab code along with your paper. Alternatively, you can provide the links to such files in a README file that appears on Xplore along with your paper. For details, please see IEEE Author Digital Toolbox under "Multimedia." To make your work reproducible by others, the TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper. Files that are to be included with the final paper must be uploaded for consideration in the review process.
5. *Submission:* After uploading all files and proofreading them, submit your manuscript by clicking "Submit." A confirmation of the successful submission will open on screen containing the manuscript tracking number and will be followed with an e-mail confirmation to the corresponding and all contributing authors. Once you click "Submit," your manuscript cannot be changed in any way.
6. *Copyright 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.

Manuscript Length. No manuscript can be more than 5 double-column single-spaced published pages, 10 point type, with the fifth page containing no material save references. We will immediately request that longer submissions be shortened and they will not be given an official reception date; this applies even to papers that have been accepted as part of the review process but whose length turns out to be excessive for publication.

Resubmission of Previously Rejected Manuscripts.

Authors of rejected manuscripts are allowed to resubmit their manuscripts only once. The Signal Processing Society strongly discourages resubmission of rejected manuscripts more than once. At the time of submission, you will be asked whether you consider your manuscript to be a new submission or a resubmission of an earlier rejected manuscript. If you choose to submit a new version of your manuscript, you will be asked to submit supporting documents identifying the previous submission number and detailing how your new version addresses all of the reviewers' comments.

Full details of the resubmission process can be found in the Signal Processing Society "Policy and Procedures Manual".

Author Misconduct.

Author Misconduct Policy: Plagiarism includes copying someone else's work without appropriate credit, using someone elses work without clear delineation of citation, and the uncited reuse of an authors previously published work that also involves other authors. Plagiarism is unacceptable.

Self-plagiarism involves the verbatim copying or reuse of an author own prior work without appropriate citation; it is also unacceptable. Self-plagiarism includes duplicate submission of a single journal manuscript to two different journals, and submission of two different journal manuscripts which overlap substantially in language or technical contribution.

Authors may only submit original work that has not appeared elsewhere in a journal publication, nor is under review for another journal publication. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper. If authors have used their own previously published work as a basis for a new submission, they are required to cite

Digital Object Identifier 10.1109/LSP.2014.2361592

the previous work and very briefly indicate how the new submission offers substantively novel contributions beyond those of the previously published work.

It is acceptable for conference papers to be used as the basis for a more fully developed journal submission. Still, authors are required to cite related prior work; the papers cannot be identical; and the journal publication must include novel aspects.

Author Misconduct Procedures: The procedures that will be used by the Signal Processing Society in the investigation of author misconduct allegations are described in the IEEE SPS Policies and Procedures Manual.

Author Misconduct Sanctions: The IEEE Signal Processing Society will apply the following sanctions in any case of plagiarism, or in cases of self- plagiarism that involve an overlap of more than 25% with another journal manuscript:

1. immediate rejection of the manuscript in question;
2. immediate withdrawal of all other submitted manuscripts by any of the authors, submitted to any of the Society's publications (journals, conferences, workshops), except for manuscripts that also involve innocent co-authors;
3. prohibition against each of the authors for any new submissions, either individually, in combination with the authors of the plagiarizing manuscript, or in combination with new co-authors, to all of the Society's publications (journals, conferences, workshops). The prohibition shall continue for one year from notice of suspension.

Further, plagiarism and self-plagiarism may also be actionable by the IEEE under the rules of Member Conduct.

Submission Format. Authors are encouraged to prepare manuscripts employing the on-line style files developed by IEEE. All manuscripts accepted for publication will require the authors to make final submission employing these style files. The style files are available on the web at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2 under Template for all TRANSACTIONS. (LaTeX and MS Word), with the caveat that the paper submitted should fit into 4 double column pages and that the Author's Biographies are omitted in the LETTERS.

Authors using LaTeX: the two PDF versions of the manuscript needed for submission can both be produced by the IEEEtran.cls style file. A double-spaced document is generated by including /documentclass[11pt,draftcls,onecolumn]{IEEEtran} as the first line of the manuscript source file, and a single-spaced double-column document for estimating the publication page charges via \documentclass[10pt,twocolumn,twoside]{IEEEtran}.

- *Title page and abstract:* The first page of the manuscript shall contain the title, names and contact information for all authors (full mailing address, institutional affiliations, phone, fax, and e-mail), the abstract, and the EDICS. An asterisk $*$ should be placed next to the name of the Corresponding Author who will serve as the main point of contact for the manuscript during the review and publication processes. An abstract limited to 100 words is required. The abstract should not only indicate the scope of the paper but should also summarize the author's conclusions so that the abstract by itself may be useful in information-retrieval systems.
- *EDICS:* All submissions must be classified by the author with an EDICS (Editors Information Classification Scheme) selected from the list of EDICS published online at: http://www.signalprocessingsociety.org/publications/periodicals/LETTERS/LETTERS-edics/.
- Note: EDICS are necessary to begin the peer review process. Upon submission of a new manuscripts, please choose the EDICS categories that best suit your manuscript. Failure to do so will likely result in a delay of the peer review process.
- The EDICS category should appear on the first page — i.e., the title and abstract page of the manuscript.
- *Illustrations and tables:* Each figure and table should have a caption that is intelligible without requiring reference to the text. Illustrations/tables may be worked into the text of a newly-submitted manuscript, or placed at the end of the manuscript. (However, for the final submission, illustrations/tables must be submitted separately and not interwoven with the text.) Illustrations in color may be used but, unless the final publishing will be in color, the author is responsible that the corresponding grayscale figure is understandable.

  In preparing your illustrations, note that in the printing process, most illustrations are reduced to single-column width to conserve space. This may result in as much as a 4:1 reduction from the original. Therefore, make sure that all words are in a type size that will reduce to a minimum of 9 points or 3/16 inch high in the printed version. Only the major grid lines on graphs should be indicated.
- *Abbreviations:* The LETTERS follows the practices of the IEEE on units and abbreviations, as outlined in the Institutes published standards. See http://www.ieee.org/documents/auinfo07.pdf for details.
- *Mathematics:* All mathematical expressions must be legible. Do not give derivations that are easily found in the literature; merely cite the reference.
- *Native Language Author Names:* IEEE supports the publication of Chinese, Japanese, and Korean (CJK) author names in the native language alongside the English versions of the names in the author list of an article. Chinese authors may use either Simplified or Traditional characters. Authors must provide the native language name in unicode characters to be displayed in the byline of the article, in parentheses, after the English version of the name. The manuscript can be prepared using the "Insert Symbols" list in Microsoft Word or the CJK Ascii Unicode for LaTeX. For more information, see http://www.ieee.org/publications_standards/publications/authors/auth_names_native_lang.pdf.

Final Manuscript Submission Guidelines. Upon formal acceptance of a manuscript for publication, instructions for providing the final materials required for publication will be sent to the Corresponding Author. Finalized manuscripts should be prepared in LaTeX or MS Word, and are required to use the style files established by IEEE, available at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2 under Template for all LETTERS. Instructions for preparing files for electronic submission are as follows:

- Files must be self-contained; that is, there can be no pointers to your system setup.
- Include a header to identify the name of the LETTERS, the name of the author, and the software used to format the manuscript.
- Do not create special macros.
- Do not send PostScript files of the text.
- File names should be lower case.
- Graphics files should be separate from the text, and not contain the caption text, but include callouts like (a), (b).
- Graphics file names should be lower case and named fig1.eps, fig2.tif, etc.
- Supported graphics types are EPS, PS, TIFF, or graphics created using Word, PowerPoint, Excel or PDF. Not acceptable is GIF, JPEG, WMF, PNG, BMP or any other format (JPEG is accepted for author photographs only). The provided resolution needs to be at least 600 dpi (400 dpi for color). Additional instructions for preparing, verifying the quality, and submitting graphics are available via http://www.ieee.org/publications_standards/publications/authors/authors_journals.html.

Open Access.

This 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 over- length 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 Charge: 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 four pages that comprise the standard length.

2015 IEEE SIGNAL PROCESSING SOCIETY MEMBERSHIP APPLICATION

Mail to: **IEEE OPERATIONS CENTER, ATTN: Louis Curcio, 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

IEEE

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

8.

2015 IEEE MEMBER RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15 Pay Full Year	1 Mar -15 Aug 15 Pay Half Year**
United States	$193.00 ☐	$96.50 ☐
Canada (incl. GST)	$171.25 ☐	$85.63 ☐
Canada (incl. HST for PEI)	$184.30 ☐	$92.15 ☐
Canada (incl. HST for Nova Scotia)	$185.75 ☐	$92.88 ☐
Canada (incl. HST for NB, NF and ON)	$182.85 ☐	$91.43 ☐
Canada (incl. GST and QST Quebec)	$185.71 ☐	$92.86 ☐
Africa, Europe, Middle East	$158.00 ☐	$79.00 ☐
Latin America	$149.00 ☐	$74.50 ☐
Asia, Pacific	$150.00 ☐	$75.00 ☐

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 13%, 14% 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.

IEEE Signal Processing Society

2015 SPS MEMBER RATES

		16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*		$ 20.00 ☐	$ 10.00 ☐
Fee includes: **IEEE Signal Processing Magazine** (electronic and digital), **Inside Signal Proc. eNewsletter** (electronic) and **IEEE Signal Processing Society Content Gazette** (electronic).			
Add $15 to enhance SPS Membership and also receive:		$15.00 ☐	$ 7.50 ☐
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.			
Publications available only with SPS membership:			
Signal Processing, IEEE Transactions on:	Print	$190.00 ☐	$ 95.00 ☐
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print	$145.00 ☐	$ 72.50 ☐
Image Processing, IEEE Transactions on:	Print	$188.00 ☐	$ 94.00 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$163.00 ☐	$ 81.50 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$160.00 ☐	$ 80.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 35.00 ☐	$ 17.50 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 55.00 ☐	$ 27.50 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$ 65.00 ☐	$ 32.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
***New!* IEEE Trans. on Cognitive Comm. &Networking**	Electronic	$ 26.00 ☐	$ 13.00 ☐
***New!* IEEE Trans. on Computational Imaging**	Electronic	$ 28.00 ☐	$ 14.00 ☐
***New!* IEEE Trans. on Big Data**	Electronic	$ 25.00 ☐	$ 12.50 ☐
***New!* IEEE Trans. on Molecular, Biological, & Multi-scale Communications**	Electronic	$ 24.00 ☐	$ 12.00 ☐
IEEE Internet of Things Journal	Electronic	$ 26.00 ☐	$ 13.00 ☐
IEEE Trans. on Cloud Computing	Electronic	$ 42.00 ☐	$ 21.00 ☐
IEEE Trans. on Computational Social Systems	Electronic	$ 30.00 ☐	$ 15.00 ☐
***New!* IEEE Trans. on Signal & Info Proc. Over Networks**	Electronic	$ 28.00 ☐	$ 14.00 ☐
IEEE Biometrics Compendium:	Online	$ 30.00 ☐	$ 15.00 ☐
Computing in Science & Engrg. Mag.:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 149.00 ☐	$ 74.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 74.00 ☐	$ 37.00 ☐
	Electronic	$ 53.00 ☐	$ 26.50 ☐
	Print & Electronic	$ 89.00 ☐	$ 44.50 ☐
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM	$ 40.00 ☐	$ 20.00 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 42.00 ☐	$ 21.00 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$149.00 ☐	$ 74.50 ☐
Network Science and Engrg., IEEE Trans. on:	Electronic	$ 33.00 ☐	$ 16.50 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 25.00 ☐	$ 12.50 ☐
	Electronic	$ 25.00 ☐	$ 12.50 ☐
	Print & Electronic	$ 40.00 ☐	$ 20.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$149.00 ☐	$ 74.50 ☐
IEEE Sensors Journal:	Print	$150.00 ☐	$ 75.00 ☐
	Electronic	$ 50.00 ☐	$ 25.00 ☐
Smart Grid, IEEE Transactions on:	Print	$100.00 ☐	$ 50.00 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$120.00 ☐	$ 60.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$120.00 ☐	$ 60.00 ☐
	Electronic	$ 48.00 ☐	$ 24.00 ☐
	Print & Electronic	$120.00 ☐	$ 60.00 ☐
IEEE Wireless Communications Letters:	Print	$ 80.00 ☐	$ 40.00 ☐
	Electronic	$ 18.00 ☐	$ 9.00 ☐
	Print & Electronic	$ 95.00 ☐	$ 47.50 ☐
***New!* IEEE Life Sciences Letters** (Open Access Pub)	Electronic	No Fee	

*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 follow information:
Member Recruiter Name: _____
IEEE Recruiter's Member Number (Required): _____

2015 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: Louis Curcio, 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

IEEE

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 _____
Day Month Year

☐ Male ☐ Female

7. CONTACT INFORMATION

Office Phone/Office Fax | Home Phone/Home Fax

Office E-Mail | Home E-Mail

8.

2015 IEEE STUDENT MEMBER RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15 Pay Full Year	1 Mar -15 Aug 15 Pay Half Year**
United States	$32.00 ☐	$16.00 ☐
Canada (incl. GST)	$33.60 ☐	$16.80 ☐
Canada (incl. HST for NB, NF, and ON)	$36.16 ☐	$18.08 ☐
Canada (incl. HST for Nova Scotia)	$36.80 ☐	$18.40 ☐
Canada (incl. HST for PEI)	$36.48 ☐	$18.24 ☐
Canada (incl. GST and QST Quebec)	$36.79 ☐	$18.40 ☐
Africa, Europe, Middle East, Latin America, Asia, Pacific	$27.00 ☐	$13.50 ☐

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 13%,14% 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.

IEEE Signal Processing Society

2015 SPS STUDENT MEMBER RATES

	16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*	$10.00 ☐	$ 5.00 ☐

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: $ 8.00 ☐ $ 4.00 ☐
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.

Publications available only with SPS membership:

Signal Processing, IEEE Transactions on:	Print	$ 95.00 ☐	$ 47.50 ☐
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print	$ 73.00 ☐	$ 36.50 ☐
Image Processing, IEEE Transactions on:	Print	$ 94.00 ☐	$ 47.00 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$ 82.00 ☐	$ 41.00 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$ 80.00 ☐	$ 40.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 18.00 ☐	$ 9.00 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 28.00 ☐	$ 14.00 ☐
	Electronic	$ 20.00 ☐	$ 10.00 ☐
	Print & Electronic	$ 65.00 ☐	$ 32.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
***New!* IEEE Trans. on Cognitive Comm. &Networking**	Electronic	$ 13.00 ☐	$ 6.50 ☐
***New!* IEEE Trans. on Computational Imaging**	Electronic	$ 14.00 ☐	$ 7.00 ☐
***New!* IEEE Trans. on Big Data**	Electronic	$ 13.00 ☐	$ 6.50 ☐
***New!* IEEE Trans. on Molecular, Biological, & Multi-Scale Communications**	Electronic	$ 12.00 ☐	$ 6.00 ☐
IEEE Internet of Things Journal	Electronic	$ 13.00 ☐	$ 6.50 ☐
IEEE Trans. on Cloud Computing	Electronic	$ 21.00 ☐	$ 10.50 ☐
IEEE Trans. on Computational Social Systems	Electronic	$ 15.00 ☐	$ 7.50 ☐
***New!* IEEE Trans. on Signal & Info Proc. Over Networks**	Electronic	$ 14.00 ☐	$ 7.00 ☐
IEEE Biometrics Compendium:	Online	$ 15.00 ☐	$ 7.50 ☐
Computing in Science & Engrg. Mag.:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 75.00 ☐	$ 37.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 37.00 ☐	$ 18.50 ☐
	Electronic	$ 27.00 ☐	$ 13.50 ☐
	Print & Electronic	$ 45.00 ☐	$ 22.50 ☐
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM	$ 20.00 ☐	$ 10.00 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 21.00 ☐	$ 10.50 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 75.00 ☐	$ 37.50 ☐
Network Science and Engrg., IEEE Trans. on:	Electronic	$ 17.00 ☐	$ 8.50 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 13.00 ☐	$ 6.50 ☐
	Electronic	$ 13.00 ☐	$ 6.50 ☐
	Print & Electronic	$ 20.00 ☐	$ 10.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 75.00 ☐	$ 37.50 ☐
IEEE Sensors Journal:	Print	$150.00 ☐	$ 75.00 ☐
	Electronic	$ 28.00 ☐	$ 14.00 ☐
Smart Grid, IEEE Transactions on:	Print	$ 50.00 ☐	$ 25.00 ☐
	Electronic	$ 20.00 ☐	$ 10.00 ☐
	Print & Electronic	$ 60.00 ☐	$ 30.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$ 60.00 ☐	$ 30.00 ☐
	Electronic	$ 24.00 ☐	$ 12.00 ☐
	Print & Electronic	$ 60.00 ☐	$ 30.00 ☐
IEEE Wireless Communications Letters:	Print	$ 40.00 ☐	$ 20.00 ☐
	Electronic	$ 9.00 ☐	$ 4.50 ☐
	Print & Electronic	$ 48.00 ☐	$ 24.00 ☐
***New!* IEEE Life Sciences Letters** (Open Access Pub)	Electronic	No Fee	

*IEEE membership required or requested
Affiliate application to join SP Society only.

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): _____