Newsjournal of the Society for Industrial and Applied Mathematics sinews.siam.org

siam news

Volume 50/ Issue 8
October 2017

Imaging Sciences Special Issue

Submit an abstract and register for the 2018 SIAM Conference on Imaging Science, to be held June 5-8, 2018 in Bologna, Italy.

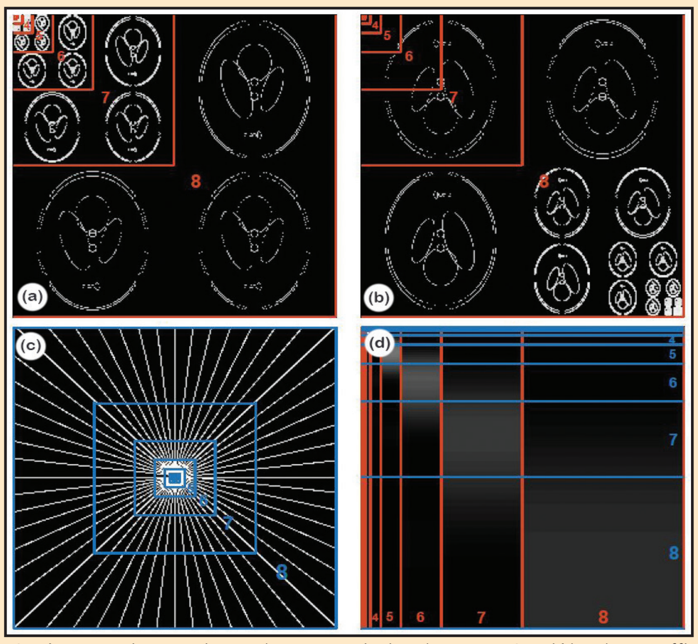


Figure 2. *Asymptotic sparsity and asymptotic incoherence.* **2a.** *Wavelet coefficients of the Shepp-Logan phantom, arranged according to increasing scale.* **2b.** *Coefficients of the flipped phantom.* **2c.** *Radial sampling map in k-space. The square annulus regions denote the essential frequency concentration of wavelets at a given scale.* **2d.** *Absolute values of the matrix $U = F\Phi$, with larger and smaller values corresponding to lighter or darker colours respectively. Vertical lines indicate the wavelet scales and horizontal lines indicate the annular frequency bands. Image credit: Alexander Bastounis, Ben Adcock, and Anders C. Hansen.*

In an article on page 5, Alexander Bastounis, Ben Adcock, and Anders C. Hansen describe recent compressed sensing applications that yield significant benefits to imaging.

Unifying Different Perspectives: From Cubism to Convolution

By Rujie Yin

Visual perception as a fundamental sensation that shapes our understanding of the world has long been of interest to both science and art. Neural activities related to various stages of visual perception are associated with different areas of the visual cortex. Objects are first "observed" by cells tuned to elementary stimulus. In subsequent stages, specific regions of the brain—which handle more complex structures—are activated depending on what one is looking at; faces and Chinese characters are very different! Neuroscientists still do not sufficiently understand the integration of elements detected in early stages to create the concept of an object. On the other end of the spectrum, artists are experimenting with the same subject. Cubism is one avant-garde art movement exploring the relation between concept formation and perception. Cubist paintings usually depict objects in parts, from multiple viewpoints simultaneously (see Figure 1, on page 4). It is difficult (but not impossible) for spectators to "picture" the objects in these paintings by unifying visually-observed pieces in their mind, like solving a virtual puzzle.

Image processing, during which an image is represented as a mathematical object whose properties reflect its characterization, presents a similar puzzle. Different representations typically provide varied yet complementary interpretations of an image. A variety of proposed image models have successfully generated high-performance algorithms, but uniting these models poses a challenge — not unlike creating a cubist painting. I will subsequently discuss how convolution, a well-known mathematical operation, can effectively combine two distinct classes of image models.

Many image models fall into two types of representations: local and nonlocal, which offer intrinsically different viewpoints. Local image representations focus on the characterization of local features present in images. Wavelet decomposition, during which wavelets serve as the elementary stimuli in our visual cortex, is perhaps the most classical local image representation. It is widely observed that, given a wavelet basis, an image can often be well-approximated by only a few basis elements (wavelets). Furthermore, the wavelets are both locally-supported and shift- and scale-

See **Cubism** *on page 4*

Processing Manifold-Valued Images

By Ronny Bergmann, Friederike Laus, Johannes Persch, and Gabriele Steidl

The mathematical notion of a manifold dates back to 1828, when Carl Friedrich Gauss established an important invariance property of surfaces while proving his Theorema Egregium. In his habilitation lecture in 1854, Bernhard Riemann intrinsically extended Gauss's theory to manifolds of arbitrary dimension, such that they are not dependent upon the embedding in higher dimensional spaces. This is now called a Riemannian manifold. Modern image acquisition methods are able to capture information that is no longer restricted to Euclidean spaces but can be manifold-valued. Such imaging methods include the following:

– Interferometric Synthetic Aperture Radar (InSAR), used in geodesy and remote sensing where each image pixel is on the circle $\mathbb{S}^1$

– Diffusion-tensor magnetic resonance imaging (DT-MRI), which produces images with values in the manifold of symmetric, positive, definite 3×3 matrices SPD(3), thus mapping directional diffusion processes of molecules—mainly water—in biological tissues

– Electron Backscatter Diffraction (EBSD), which analyzes crystalline materials where the images have pixels in the group of rotations SO(3) modulo the crystal's symmetry group.

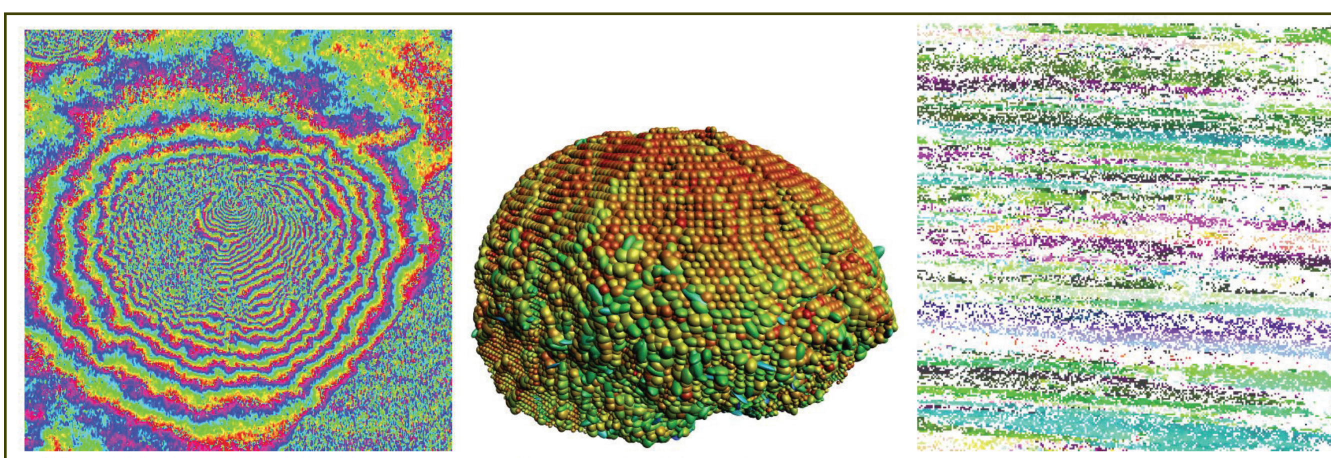
Images produced by these techniques are depicted in Figure 1. Furthermore, manifold-valued images arise when working in color spaces that are more adapted to human color perception than the RGB space, such as HSI (hue, saturation, intensity) or CB (chromaticity, brightness). Here, we deal with the (product)-manifolds $\mathbb{S}^1 \times \mathbb{R} \times \mathbb{R}$ and $\mathbb{S}^2 \times \mathbb{R}$, respectively.

Processing manifold-valued signals and images proposes new challenges in image processing that affect classical tasks like denoising, inpainting, and segmentation. In the real-valued case, variational methods—convex optimization in particular—are well adapted to such large-scale problems. Other successful approaches include nonlocal patch-based methods, which rely on an image self-similarity assumption, and recently-developed learning methods. What follows offers a brief overview of recent results in manifold-valued image processing, obtained by attempting to generalize results from the real to manifold-valued setting.

Tiny inaccuracies that result in noisy data are inevitable when measurements are taken, regardless of whether the measurements are manifold-valued. Modeling an image as a realization of some random variable, whose distribution reflects the circumstances under which the image is taken, is a common method to account for this randomness. In the real-valued case, the standard approach assumes that the noise is additive, white, and Gaussian, which is asymptotically justified by the Central Limit Theorem. Much effort has been spent denoising real-valued images corrupted with white Gaussian noise, but the situation completely changes with manifold-valued images. The definition of a Gaussian distribution is not canonical on a manifold, and different attempts—generalizing different characterizing properties of the real-valued normal distribution—appear in the literature.

Current state-of-the-art denoising methods for real-valued images include nonlocal

See **Manifold-Valued Images** *on page 3*



(a) InSAR data from Mt. Vesuvius

(b) DT-MRI from the Camino project

(c) EBSD data of an aluminum sample

Figure 1. *Manifold-valued images acquired with different devices, yielding data values given on the following:* **1a.** *The circle $\mathbb{S}^1$, colored using the hue [11].* **1b.** *The manifold of symmetric positive definite 3×3 matrices, illustrated using their eigenvalues and eigenvectors to draw an ellipsoid [3].* **1c.** *Orientations SO(3), where the orientation modulo the phase is mapped onto a colorized sphere. 1c courtesy of the Institute of Materials Science and Engineering at the University of Kaiserslautern.*

Nonprofit Org
U.S. Postage
PAID
Permit No 360
Bellmawr, NJ

SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

siam news
Volume 50/ Issue 8/ October 2017

Issue Acknowledgments

SIAM News would like to thank Eric Miller (Tufts University) for his help in putting together this Imaging Sciences Special Issue.

A Multiprecision World

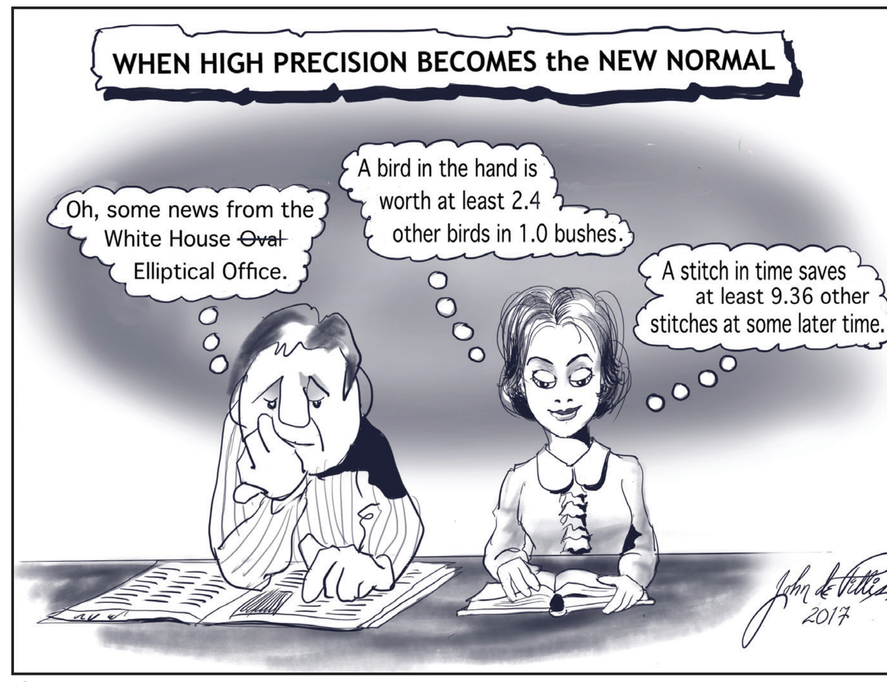
Traditionally, floating-point arithmetic has come in two precisions: single and double. But with the introduction of support for other precisions, thanks in part to the influence of applications, the floating-point landscape has become much richer in recent years.

To see how today's multiprecision world came about, we need to start with two important events from the 1980s. The IEEE standard for binary floating-point arithmetic was published in 1985. It defined single precision (32-bit) and double precision (64-bit) floating-point formats, which carry the equivalent of about eight and 16 significant decimal digits, respectively. This led to the relatively homogeneous world of floating-point arithmetic that we enjoy today, which contrasts starkly with the 1970s, when different computer manufacturers used different floating-point formats and even different bases (hexadecimal in the case of some IBM machines). The second important event in the 1980s was the introduction of the Intel 8087 coprocessor, which carried out floating-point computations in hardware (in conjunction with an 8086 processor) and enabled much faster scientific computations on desktop machines. Intel went on to incorporate the coprocessor into the main processor in the Pentium and subsequent series of processors.

Throughout the 1990s, we had the choice of working in single or double precision arithmetic in most computing environments. Single precision did not intrinsically run faster than double precision on Intel chips, but its lower storage requirement could lead to speed benefits due to better use of cache memory.

FROM THE SIAM PRESIDENT
By Nicholas Higham

The picture started to change in 1999 when Intel introduced streaming single instruction, multiple data (SIMD) extensions (SSE), which allowed single precision arithmetic to execute up to twice as fast as double. A few years later, the Cell processor, designed by Sony, Toshiba, and IBM for use in the Sony PlayStation 3 gaming system, offered single precision arithmetic running up to 14 times faster than double precision, thus presenting interesting opportunities for scientific computing. These developments directed efforts towards algorithms with the ability to exploit two precisions to solve a problem faster or more accurately than just one precision. The concept of such algorithms is not new. Up until the 1970s, many computers could accumulate inner products at twice the working precision and no extra cost, and the method of iterative refinement for linear systems—first programmed by James Hardy Wilkinson on the Pilot ACE in 1948—exploited this capability to improve the accuracy of an initial solution computed with LU factorization. A new form of iterative refinement that employs single precision to accelerate the double precision solution process was developed in [3].



Cartoon created by mathematician John de Pillis.

In the last few years, the advent of half precision arithmetic (16 bits) has enriched the floating-point landscape. Although the 2008 revision of the IEEE standard originally defined it only as a storage format, manufacturers have started to offer half precision floating-point arithmetic in accelerators such as graphics processing units (GPUs). Half precision offers both speed benefits (it operates up to twice as fast as single precision, though only the top-end GPUs attain the factor 2) and lower energy consumption. The main application driver for half precision is machine learning (and in particular, deep learning), where algorithms have been found empirically to perform satisfactorily in low precision.

I am not aware of any rigorous analysis that explains the success of machine learning algorithms run in half—or even lower—precision. One possible explanation is that we are solving the wrong optimization problem (as the correct one is too difficult to solve) and thus do not need to solve it accurately. Another is that low precision has a beneficial regularizing effect. Yet from the traditional numerical analysis point of view, half precision is dubious. The usual rounding error bound for the inner product of two n-vectors contains the constant nu, where u is the unit roundoff, so in half precision (which has $n \approx 5\times10^{-4}$), we cannot guarantee even one correct significant digit in the computed inner product once n exceeds 2,048. Indeed, the set of half precision numbers is small: there are only 61,441 normalized numbers, and the spacing between 32,768 and the largest number, 65,504, is 32.

People will be tempted to use half precision as it becomes more accessible in hardware, potentially with serious consequences if relative errors of order 1 are obtained in critical applications. The limitation that half precision has a range of only $10^{\pm5}$ means that in many problems, one is just as likely to obtain NaNs as output (resulting from overflow) as completely incorrect numbers. This presents work for our community to better understand the behavior of algorithms in low precision, perhaps through a statistical approach to rounding error analysis instead of the usual approach of proving worst-case bounds.

The precision landscape has been getting more interesting at the higher end as well. The 2008 IEEE standard revision introduced a quadruple precision floating-point format, which is available almost exclusively in software (the IBM z13 processor being a rare exception), perhaps as a compiler option. *Arbitrary* precision arithmetic is available in several environments, including Maple, Mathematica, Sage, Julia through its BigFloat data type, and MATLAB with the Symbolic Math Toolbox or the Multiprecision Computing Toolbox (Advanpix). Several of these systems utilize the GNU MPFR Library, an open source C library for multiple precision floating-point computations. Having arbitrary precision floating-point arithmetic at our fingertips is not something many of us are accustomed to. I first became intrigued with the possibility during a visit to the University of Toronto (U of T) in the 1980s, when Tom Hull introduced me to Numerical Turing. Turing was a Pascal-like language developed in U of T's Department of Computer Science for teaching, and Hull's Numerical Turing augmented it with variable precision decimal floating-point arithmetic.

Field-programmable gate arrays, which have always been configurable for different precisions of fixed-point arithmetic but now can additionally support floating-point arithmetic, also have a role to play. These low-power devices offer the possibility of customizing the floating-point format in hardware to meet the precision requirements of an application.

Once arithmetic of several precisions is available (half, single, double, quadruple), we want to harness it to compute results of the desired accuracy as efficiently as possible, bearing in mind the relative costs of the precisions.[1] A natural scenario is iterative methods such as Newton's method, where there may be no point in computing iterates accurately in the early stages of an iteration when far from the solution; increasing the precision during the iteration may reduce execution time. We can also ask whether using just a little extra precision in certain key parts of an algorithm can bring benefits to the speed or accuracy, and whether it can stabilize a potentially unstable algorithm. See [2, 5] for some recent work along these lines.

See **Multiprecision World** *on page 3*

[1] https://nickhigham.wordpress.com/2017/08/31/how-fast-is-quadruple-precision-arithmetic/

siam news

ISSN 1557–9573.

Manifold-Valued Images

Continued from page 1

patch-based approaches, in particular the nonlocal Bayes algorithm of Marc Lebrun, Antoni Buades, and Jean-Michel Morel [9], whose idea was reinterpreted and generalized to manifolds in [8]. The proposed estimation procedure relies heavily on the computation of Riemannian centers of mass — counterparts of the classical mean or expectation. Numerical examples demonstrate the excellent denoising performance of the proposed estimator for different manifolds, including $\mathbb{S}^1$, the sphere $\mathbb{S}^2$, and SPD(3). Figure 2 illustrates a result obtained for an orientation field, i.e., an $\mathbb{S}^2$-valued image. However, these are academic examples, and part of ongoing research is to examine the acquisition-dependent noise models appearing in applications.

Variational approaches that are not restricted to denoising generate a restored image as a minimizer of some functional of the form

$$\mathcal{J}(u) = \mathcal{D}(u, f) + \lambda\mathcal{R}(u), \quad \lambda > 0,$$

where $\mathcal{D}$ is a data-fitting term measuring the distance to the given data f, $\mathcal{R}$ is a regularization term—also called a prior—reflecting the assumed properties of an ideal (clean) image, and $\lambda > 0$ is a factor that balances the influence of the regularizer. A usual data-fitting term for real-valued images is the squared Euclidean distance between f and u, which is naturally replaced by the squared geodesic distance for manifold-valued images. Choosing appropriate regularization terms is usually more involved; selections should ensure a reduction of noise in the minimizer and the preservation of the image structure. In the Euclidean case, the total variation (TV)—proposed by Leonid Rudin, Stanley Osher, and Emad Fatemi [12]—is demonstrably a powerful, edge-preserving, and convex nonsmooth regularizer. Mariano Giaquinta and Domenico Mucci [7] used Cartesian currents to investigate the notation of TV of functions with values on a manifold, while Evgeny Strekalovskiy and Daniel Cremers [13] first applied the technique to phase-valued images.

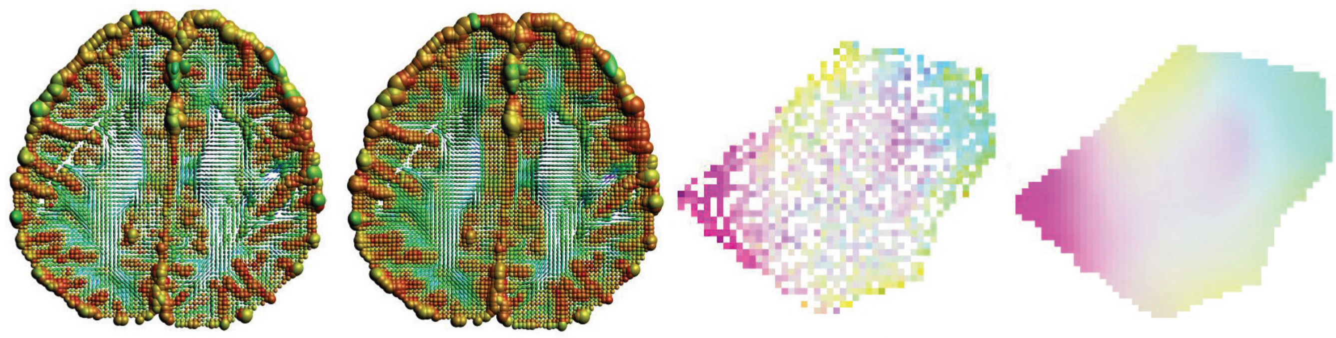
In addition to first-order derivatives that occur in the classical TV approach, the incorporation of higher-order derivatives into the model to reduce the staircasing effect caused by TV regularization and adapt to specific applications is also desirable. The definition of second-order spatial differences is not straightforward for manifold-valued images. Following the idea that the second-order difference term in the Euclidean setting reads $\|x-2y+z\| = 2\|\frac{1}{2}(x+z)-y\|$, [2] provides a definition on manifolds that uses the geodesic distance from the midpoints of the geodesics connecting x with z to y.

One can apply Riemannian optimization methods to compute a minimizer of the resulting functionals $\mathcal{J}$. These intrinsic methods are often very efficient, since they exploit the underlying geometric structure of the manifold [10]. Various methods have been proposed for smooth functions $\mathcal{J}$, ranging from simple gradient descents on manifolds to more sophisticated trust region methods. Recently, so-called half-quadratic minimization methods, belonging to the group of quasi-Newton methods and covering (for example) iteratively re-weighted least squares methods, were generalized to manifold-valued images [3].

Unfortunately, the TV regularization term is not differentiable. However, in the Euclidean setting it is convex, meaning that convex analysis tools—including powerful algorithms based on duality theory—are applicable. A prominent example is the alternating direction method of multipliers (ADMM), which is equivalent to the Douglas-Rachford algorithm. Proximal mappings, which one can efficiently compute for special priors appearing in Euclidean image processing tasks, are a central ingredient of these algorithms.

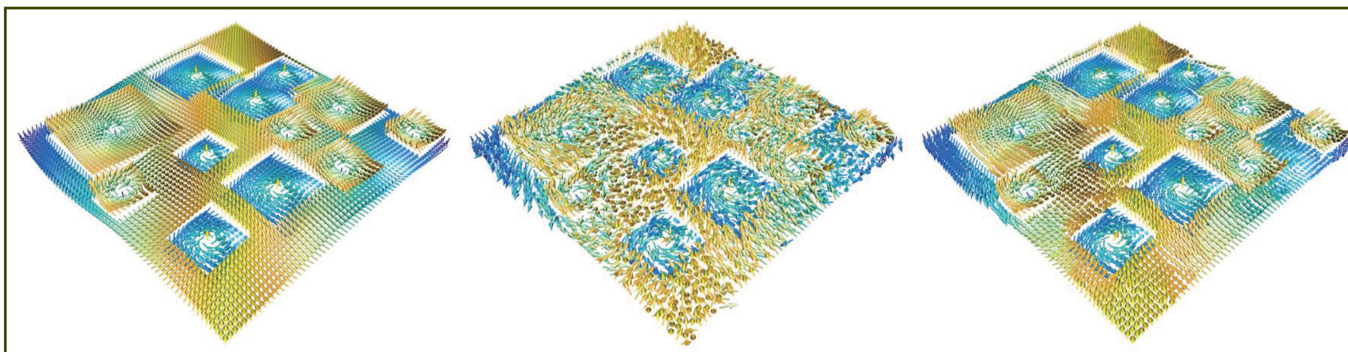
Several efforts have recently attempted to translate concepts from convex analysis to manifolds. One can establish a certain theory of convex functions on Hadamard manifolds, i.e., complete, simply-connected Riemannian manifolds of nonpositive sectional curvature, as (for example) $\mathrm{SPD}(n)$ or hyperbolic spaces. In particular, one can introduce the (inexact) cyclic proximal point algorithm on these manifolds [1], a method that was also used to minimize the functional with first- and second-order differences [2]. Efficient computation of the proximal mapping of second-order differences utilizes the machinery of Jacobi fields and is restricted to symmetric spaces. Figure 3 shows two denoising results employing first- and second-order differences. Symmetric spaces are characterized by the property that geodesic reflections at points are isometries. Since the classical Douglas-Rachford algorithm relies on point reflections, it was natural to extend this algorithm to symmetric Hadamard manifolds [4]. However, Hadamard spaces do not embody *all* the nice properties of convex analysis. For example, reflections at convex sets are not nonexpansive in general, and although the parallel Douglas-Rachford algorithm shows a very good numerical performance, existing theoretical convergence results remain limited to manifolds with constant nonpositive curvature.

Finally, an important aspect when working with real data is the practical implementability of the developed methods and algorithms. In the spirit of reproducible research, several groups provide their software and toolboxes, e.g., the manifold-valued image restoration toolbox MVIRT,[1] the manifold optimization Manopt package[2] [5], and toolboxes focusing on specific manifolds, such as the MTEX toolbox[3] for EBSD images. All of these packages are available in MATLAB, making it possible to test one's own ideas and enter this active field of research.



(a) Slice of the Camino data set (b) Denoised slice (c) Noisy and incomplete EBSD data (d) Denoised and inpainted data

Figure 3. *Denoising and inpainting of manifold-valued images by a variational method. 3a and 3b courtesy of [2], 3c and 3d courtesy of [3].*



(a) Original image (b) Noisy image (c) Denoised image

Figure 2. *Denoising of an artificial $\mathbb{S}^2$-valued image by a patch-based method. Image courtesy of [8].*

1 Available open source at www.mathematik.uni-kl.de/imagepro/members/bergmann/mvirt/

2 Available open source at manopt.org

3 Avaiable open source at mtex-toolbox.github.io/

References

[1] Bačák, M. (2014). *Convex Analysis and Optimization in Hadamard Spaces*. Vol. 22. *Series in Nonlinear Analysis and Applications*. Berlin, Germany: De Gruyter.

[2] Bačák, M., Bergmann, R., Steidl, G., & Weinmann, A. (2016). A second order non-smooth variational model for restoring manifold-valued images. *SIAM Journal on Scientific Computing, 38*(1), 567-597.

[3] Bergmann, R., Chan, R.H., Hielscher, R., Persch, J., & Steidl, G. (2016). Restoration of manifold-valued images by half-quadratic minimization. *Inverse Problems and Imaging, 10*(2), 281-304.

[4] Bergmann, R., Persch, J., & Steidl, G. (2016). A parallel Douglas-Rachford algorithm for restoring images with values in symmetric Hadamard manifolds. *SIAM Journal on Imaging Sciences, 9*(3), 901-937.

[5] Boumal, N., Mishra, B., Absil, P.A., & Sepulchre, R. (2014). Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research, 15*, 1455-1459.

[6] Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J., & Alexander, D.C. (2006). Camino: Open-source diffusion-MRI reconstruction and processing. In *Proceedings of the International Society for Magnetic Resonance in Medicine Scientific Meeting and Exhibition*. (pp. 2759). Seattle, WA: International Society for Magnetic Resonance in Medicine.

[7] Giaquinta, M., & Mucci, D. (2007). Maps of bounded variation with values into a manifold: total variation and relaxed energy. *Pure Applied Mathematics Quarterly, 3*(2), 513-538.

[8] Laus, F., Nikolova, M., Persch, J., & Steidl, G. (2007). A nonlocal denoising algorithm for manifold-valued images using second order statistics. *SIAM Journal on Imaging Sciences, 10*(1), 416-448.

[9] Lebrun, M., Buades, A., & Morel, J.-M. (2013). A nonlocal Bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences, 6*(3), 1665-1688.

[10] Ring W., & Wirth, B. (2012). Optimization methods on Riemannian manifolds and their application to shape space. *SIAM Journal on Optimization, 22(2)*, 596-627.

[11] Rocca, F., Prati, C., & Guarnieri, A.M. (1997). Possibilities and limits of SAR interferometry. *Image Processing Techniques*, 15-26.

[12] Rudin, L.I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena, 60*(1), 259-268.

[13] Strekalovskiy, E. & Cremers, D. (2013). Total cyclic variation and generalizations. *Journal of Mathematical Imaging and Vision, 47(3), 258-277.*

Ronny Bergmann received his Ph.D. in mathematics from the University of Lübeck in 2013. He is currently a postdoctoral research fellow at the University of Kaiserslautern. Friederike Laus and Johannes Persch received their M.Sc. in mathematics from the University of Kaiserslautern in 2015 and are currently Ph.D. students under the supervision of Gabriele Steidl. Gabriele Steidl earned her Ph.D. and habilitation in mathematics from the University of Rostock. She is currently head of the Mathematical Image Processing Group at the University of Kaiserslautern and a consultant at the Fraunhofer Institute for Industrial Mathematics.

Multiprecision World

Continued from page 2

If we aim to achieve a given fairly low level of accuracy or residual with an iterative method, say t bits, we can ask what the best choice of precision ($p > t$ bits) is in which to run the computations. It turns out that for Krylov methods (for example), the number of iterations can depend strongly on the precision [4], meaning that the fastest computation might not result from the lowest precision that achieves the desired accuracy.

SIAM News readers may remember "A Hundred-dollar, Hundred-digit Challenge" announced by Nick Trefethen in January 2002. That challenge asked for 10 problems to be solved to 10-digit accuracy. Although high precision arithmetic could be used in the solutions as part of a brute force attack, it turned out to be generally not necessary [1]. This example serves as a reminder that mathematical ingenuity in the choice of algorithm can enable a great deal to be done in double precision arithmetic, so one should always think carefully before resorting to higher precision arithmetic, with its attendant increase in cost. Nevertheless, today's multiprecision computational landscape offers great scope for clever exploitation, presenting exciting opportunities to researchers in our community.

References

[1] Bornemann, F., Laurie, D., Wagon, S., & Waldvogel, J. (2004). *The SIAM 100-Digit Challenge: A Study in High-Accuracy Numerical Computing*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

[2] Carson, E., & Higham, N.J. (2017). Accelerating the Solution of Linear Systems by Iterative Refinement in Three Precisions. *MIMS EPrint 2017.24*, The University of Manchester.

[3] Langou, J., Langou, J., Luszczek, P., Kurzak, J., Buttari, A., & Dongarra, J. (2016). Exploiting the performance of 32 bit floating point arithmetic in obtaining 64 bit accuracy (revisiting iterative refinement for linear systems). In *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*. Tampa, FL.

[4] Roldao-Lopes, A., Shahzad, A., Constantinides, G.A., & Kerrigan, E.C. (2009). More Flops or More Precision? Accuracy Parameterizable Linear Equation Solvers for Model Predictive Control. In *17th IEEE Symposium on Field Programmable Custom Computing Machines* (pp. 209-216).

[5] Yamazaki, I., Tomov, S., & Dongarra, J. (2015). Mixed-Precision Cholesky QR Factorization and its Case Studies on Multicore CPU with Multiple GPUs. *SIAM J. Sci. Comp., 37*, C307-C330.

Nicholas Higham is the Richardson Professor of Applied Mathematics at the University of Manchester. He is the current president of SIAM.

Cubism

Continued from page 1

invariant in the image domain, meaning that they are copies generated by scaling and shifting a wavefront pattern known as the mother wavelet. Therefore, the pattern of the mother wavelet can effectively capture local features of images, with the mother wavelet's choice determining the class of images being modeled. Dictionary learning is a more flexible local image representation. Instead of prescribing designed local patterns like the mother wavelets in wavelet bases, a dictionary of representative patterns is learned from fixed-size patches extracted from an image or a collection of images (a training dataset). These adaptive patch patterns are then used to more efficiently decompose images consisting of patches similar to the training dataset.

A nonlocal image representation, on the other hand, focuses on repetition rather than decomposition of patterns in an image. Such a model is "nonlocal" because similar patches from one image are not necessarily localized in the image domain. One popular nonlocal image model comes from manifold learning, where image patches are assumed to vary smoothly in the patch space—forming a manifold—with a small degree of freedom (dimension of the manifold) compared to patch size (dimension of the ambient patch space). In practice, a graph whose nodes are sample patches approximates the patch manifold. Furthermore, one can define a diffusion process on the patches with respect to their similarity; the corresponding (graph) Laplacian induces an orthonormal spectral basis that encodes the connection between similar patches.

Considering local wavelet decomposition and nonlocal manifold learning (spectral decomposition) as specific examples, we demonstrate a novel way to combine local and nonlocal image representations by convolution. Given a discrete image $f \in \mathbb{R}^n$, its decomposition with respect to a J-level (overcomplete) wavelet basis generated by the mother wavelet ψ is

$$f = \sum_{i,j} a_{i,j}\psi(2^j(\cdot - i)), \quad i = 1,\cdots,n, j = 1,\cdots,J. \tag{1}$$

Because the wavelet transform is shift-invariant, we can rewrite the above decomposition as a sum of convolutions $\sum_j A_j * \psi(2^j\cdot)$, where $A_j \in \mathbb{R}^n$ is the set of wavelet coefficients $a_{\cdot j}$ associated with the scaled mother wavelet $\psi(2^j\cdot)$, with translations in the image domain. Alternatively, if we look at patches of size $2^J \times 2^J$ centering on each pixel in the image (with periodic boundary extension), we see that they are decomposed against the same set of basic patterns, i.e., mother wavelets in different scales $\psi(2^j\cdot)$. Therefore, two similar patches, p_s, p_t—centering at s and t respectively—have coefficients $A_j(s)$ and $A_j(t)$ that are close for $j = 1,\cdots,J$. In other words, coefficient vectors A_j indicate the similarity between patches. On the other hand, if we construct a graph using all patches p_i, then the spectral basis ϕ_k, $k = 1,\cdots,n$ generated from the graph Laplacian in manifold learning is an orthonormal basis of $\mathbb{R}^n$. Therefore, we can use the spectral basis to decompose the coefficient vectors $A_j = \sum_k c_{j,k}\phi_k$, which results in a reformulation of the original image decomposition (1) as a linear combination of convolution components generated from the wavelet basis and spectral basis

$$f = \sum_{j,k} c_{j,k}\phi_k * \psi(2^j\cdot) \tag{2}$$

In fact, Proposition 1 in [1] shows that given any orthonormal basis ψ_j in $\mathbb{R}^\ell$ and any orthonormal basis ϕ_k in $\mathbb{R}^n$, the bases generate a tight frame of $\mathbb{R}^n$ consisting of convolution components $v_{j,k} := \psi_j * \phi_k$ with the frame constant $\sqrt{\ell}$; $v_{j,k}$ are called convolution framelets.

Figure 1. *Les Demoiselles d'Avignon, a painting by Pablo Picasso, embodies some of the bold traits of Cubism. Public domain image.*

Combining a local and nonlocal basis results in convolution framelets with stronger representation power than either basis alone. To observe this, we consider a simulated image f containing two patterns, ψ_1 and ψ_2, whose supports divide the image domain into D_1 and D_2. In this case, the leading (nontrivial) spectral basis vector is $\phi_1 = 1_{D_1} - 1_{D_2}$ (up to a constant) and the image is thus a linear combination of four convolution framelets $f = 0.5\psi_1 * (\phi_1 + \phi_0) + 0.5\psi_2 * (\phi_0 - \phi_1)$, where $\phi_0 = 1$ is the trivial spectral basis vector (up to a constant). A pair of local and nonlocal bases can be viewed in the form of an autoencoder [1], a type of neural network whose output is the same as input, with dimensionality reduction on the input. We also find that applying regularization induced by convolution framelets improves the reconstruction result, when compared with regularization on the corresponding nonlocal bases alone.

In general, one can obtain a convolution component $v = \psi * \phi$ by distributing the pattern ψ in the image domain with respect to the layout ϕ; v inherits the regularity from both ψ and ϕ. Imagine an artist working step by step on a painting. Each time the artist paints part of the painting with a certain type of brush stroke to create a specific pattern, he/she adds a "convolution component" to the painting. The set of patterns to choose from depends on the painting's style and the artist's skill set, whereas the layout of the patterns is more closely related to the painting's content. As with art, there are many ways to represent an image, yet none is optimal. The convolution framelets introduced here present our inspiration from classical models to be further explored in the future.

References

[1] Yin, R., Gao, T., Lu, Y.M., & Daubechies, I. (2017). A tale of two bases: Local-nonlocal regularization on image patches with convolution framelets. *SIAM Journal on Imaging Sciences, 10*(2), 711-750.

Rujie Yin received her Ph.D. in applied mathematics from Duke University in May 2017. Her research focuses on multi-resolution representation and modeling of high-dimensional data, especially images.

Downwind, Faster Than the Wind

By Mark Levi

The feasibility of the title's suggestion depends on one's definition of sailing. A regular sailboat cannot exceed wind speed when going dead downwind, i.e., exactly in the wind's direction. But if propellers and gears are used instead of sails, then the seemingly impossible becomes possible.[1] In principle (and before going into any detail), it stands to reason that one can harvest energy from the relative motion of two media (air and water) and use this energy in an engine. The question is whether this can be done "in practice." Figure 1 offers a "constructive proof of concept." Two propellers are mounted on the boat as shown. Assuming that the boat is moving forward, the water propeller—connected to an electricity generator—is dragged through the water with speed v_{boat}, generating electric power

$$P_{\text{generated}} = F_{\text{drag}} v_{\text{boat}}; \tag{1}$$

we assume an ideal propeller and no losses. Here, F_{drag} refers to the drag on the propeller only — the drag on the hull is neglected. The air propeller, on the other hand, pulls

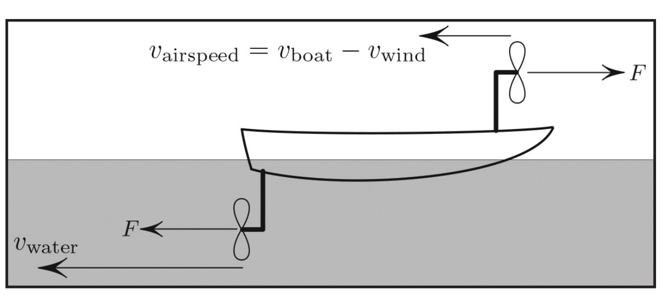


Figure 2. *Reference frame of the boat. Since $v_{\text{water}} > v_{\text{airspeed}}$, we have $P_{\text{generated}} > P_{\text{consumed}}$.*

the boat forward and is driven by an electric motor, requiring power

$$P_{\text{consumed}} = F_{\text{pull}} v_{\text{airspeed}} = F_{\text{pull}}(v_{\text{boat}} - v_{\text{wind}}). \tag{2}$$

MATHEMATICAL CURIOSITIES
By Mark Levi

Does the power generated by the "dragger" suffice to feed the puller so as to maintain *constant speed* $v_{\text{boat}} > v_{\text{wind}}$? The answer is yes, because $F_{\text{drag}} = F_{\text{pull}}$ for constant speed, so that (1) and (2) imply

$$P_{\text{generated}} > P_{\text{consumed}}. \tag{3}$$

Incidentally, the surplus is exactly what a stationary windmill would generate (assuming the same force).

Figure 2 gives an alternative view, from the boat's frame of reference; the key is that the oncoming water is faster than the oncoming air. And since the power generated/consumed depends on the propeller's speed relative to the medium, higher speed means greater power. The water propeller therefore generates more than the air propeller consumes.

Of course, the above idea is not limited to boats, and has been realized.[2]

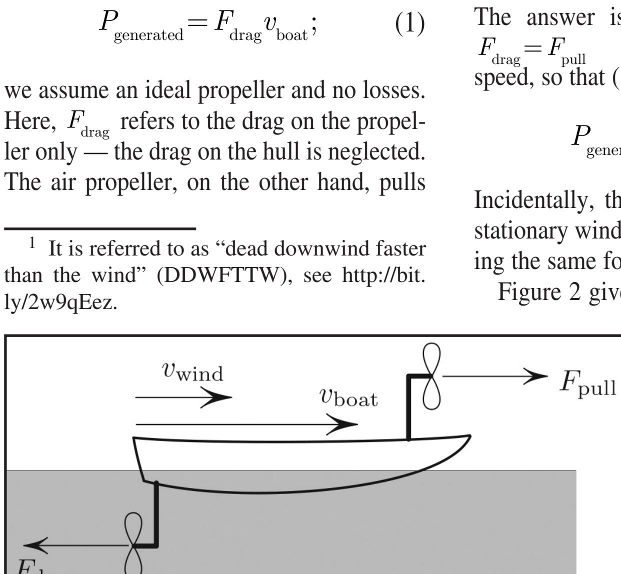


Figure 1. *Reference frame of the water. In steady motion, $F_{\text{pull}} = F_{\text{drag}} = F$. Here, F_{drag} is the drag on the propeller; the drag on the hull is ignored, as are other "imperfections."*

A Solution to Last Month's Puzzle[3]

Refer to the caption of Figure 3, which restates the puzzle. Imagine a firework exploding at point O in Figure 3, sending a myriad of shards in all directions, each with the same initial speed v. Ignoring the air resistance, the shards form an expanding circle[4] of radius vt at time t. And the center of this circle undergoes free fall, descending by $gt^2/2$ in time t. *The safety parabola is the envelope of this family of circles*, as illustrated in Figure 4. In other words, the safety ceiling serves double duty as one envelope of two different families of curves. Because of this double role, any point T on the envelope is the point of tangency with a parabolic trajectory, and also with a circle (see Figure 5). Since the circle's center undergoes free fall, $OC = gt^2/2$, where t is the time of free fall. But this is the same t as the parabolic flight time from O to T, because only one shard ever reaches T. This completes the solution.

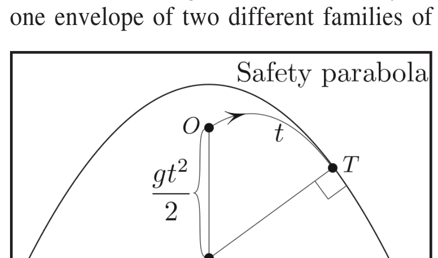


Figure 3. *Last month's puzzle. Show, without calculation, that the normal at the point of tangency T intersects the vertical at the distance $gt^2/2$ from the launch point O, where t is the time of flight from O to T.*

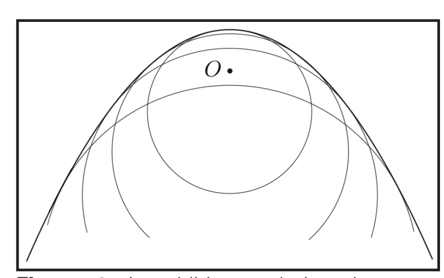


Figure 4. *In addition to being the envelope of a family of trajectories, the safety parabola is also the envelope of a one-parameter family of expanding circles with a free-falling center.*

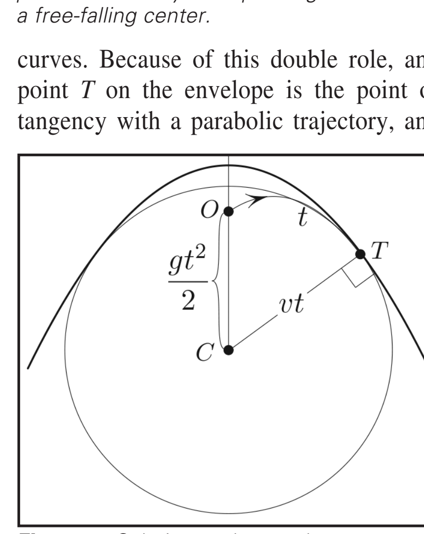


Figure 5. *Solution to the puzzle.*

The figures in this article were provided by the author.

Mark Levi (levi@math.psu.edu) is a professor of mathematics at the Pennsylvania State University.

[1] It is referred to as "dead downwind faster than the wind" (DDWFTTW), see http://bit.ly/2w9qEez.

[2] https://www.wired.com/2010/06/downwind-faster-than-the-wind/

[3] *SIAM News, 50*(7), September 2017. https://sinews.siam.org/Details-Page/parabola-of-safety-and-the-jacobian

[4] We are considering a two-dimensional cross-section of the three-dimensional picture.

From Global to Local: Getting More from Compressed Sensing

By Alexander Bastounis, Ben Adcock, and Anders C. Hansen

Over the last decade, compressed sensing and sparse recovery techniques have had a significant impact on applied mathematics and its uses in science and engineering. Compressed sensing applications have moved beyond experimentation and are beginning to be seen in new implementations. An area of particular note is imaging, where compressed sensing can be used in magnetic resonance imaging (MRI), electron tomography, and radio interferometry, among other applications. With this in mind, it is timely to revisit the mathematics of compressed sensing as it pertains to imaging. While the standard theory of compressed sensing is justifiably celebrated, it falls somewhat short of explaining phenomena that result from the application of these techniques to imaging. In this article, we describe recent work that seeks to bridge this gap. As we demonstrate, our approach yields significant practical benefits in imaging, allowing researchers to further push the limits of performance.

Standard Compressed Sensing

Compressed sensing [3, 4, 6] concerns the recovery of an object from an incomplete set of linear measurements. In a discrete setting, one can formulate this as the linear system

$$y = Ax,$$

where $y \in \mathbb{C}^m$ is the vector of measurements, $x \in \mathbb{C}^N$ is the object to recover, and $A \in \mathbb{C}^{m\times N}$ is the so-called measurement matrix. In practice, the number of measurements m is often substantially smaller than the dimension N, making recovery generally impossible. To overcome this, compressed sensing leverages two key properties: *sparsity* of the vector x and *incoherence* of the measurement vectors (rows of A). The first property asserts that x should have at most $s \le m$ significant components, with the remaining components relatively small, while the second states that the measurement vectors should be (in a formally-defined sense) spread out, rather than concentrated around a small number of entries.

A popular tool in compressed sensing theory is the *Restricted Isometry Property* (RIP). A matrix has the RIP or order s if there exists a $\delta \in (0,1)$, such that

$$(1-\delta)\|x\|^2_{\ell^2} \le \|Ax\|^2_{\ell^2} \le (1+\delta)\|x\|^2_{\ell^2},$$

for all s-sparse vectors x.

For instance, if recovery is performed by solving the convex *basis pursuit* problem

$$\min_{z\in\mathbb{C}^N} \|z\|_{\ell^1}, \text{ subject to } Az = y, \quad (1)$$

then the RIP (with sufficiently small δ) implies exact recovery of any s-sparse x and robustness with respect to perturbations in x (i.e., inexact sparsity) or y (i.e., noise).

Typically, the rows of A are drawn independently according to some random distribution. An elegant demonstration of compressed sensing mathematics considers Gaussian random vectors. These are incoherent, and a signature result asserts that A has the RIP with an optimal number of measurements $m \approx Cs\log(N/s)$.

The Flip Test

Imaging is an ideal fit for compressed sensing, and one of its original areas of application [5, 8, 11]. Though not typically sparse themselves, images can be represented sparsely in certain dictionaries, such as wavelets. Furthermore, acquisition devices found in many imaging applications are modelled with the Fourier transform, which tends to be fairly incoherent.

In their seminal work [3], Emmanuel Candès, Justin Romberg, and Terence Tao demonstrated the benefits of compressed sensing by recovering the classical Shepp-Logan phantom from incomplete Fourier measurements. This experiment is repeated in Figures 1a-1c. The theoretical basis for this result is twofold. First, the image $x \in \mathbb{C}^N$ has a sparse representation in a wavelet basis. Specifically, if $\Phi \in \mathbb{C}^{N\times N}$ is the matrix of the wavelet transform, then

$$x = \Phi c$$

for some s-sparse vector of coefficients $c \in \mathbb{C}^N$. Second, the matrix

$$A = PF\Phi \quad (2)$$

has the RIP, where $F \in \mathbb{C}^{N\times N}$ is the discrete Fourier matrix and $P \in \mathbb{C}^{m\times N}$ is the matrix of the *sampling map*, i.e., P selects rows of F according to frequencies shown in Figure 1b. As per the theory, x should recover to high accuracy as $\hat{x} = \Phi\hat{c}$, where $\hat{c}$ is a solution of (1).

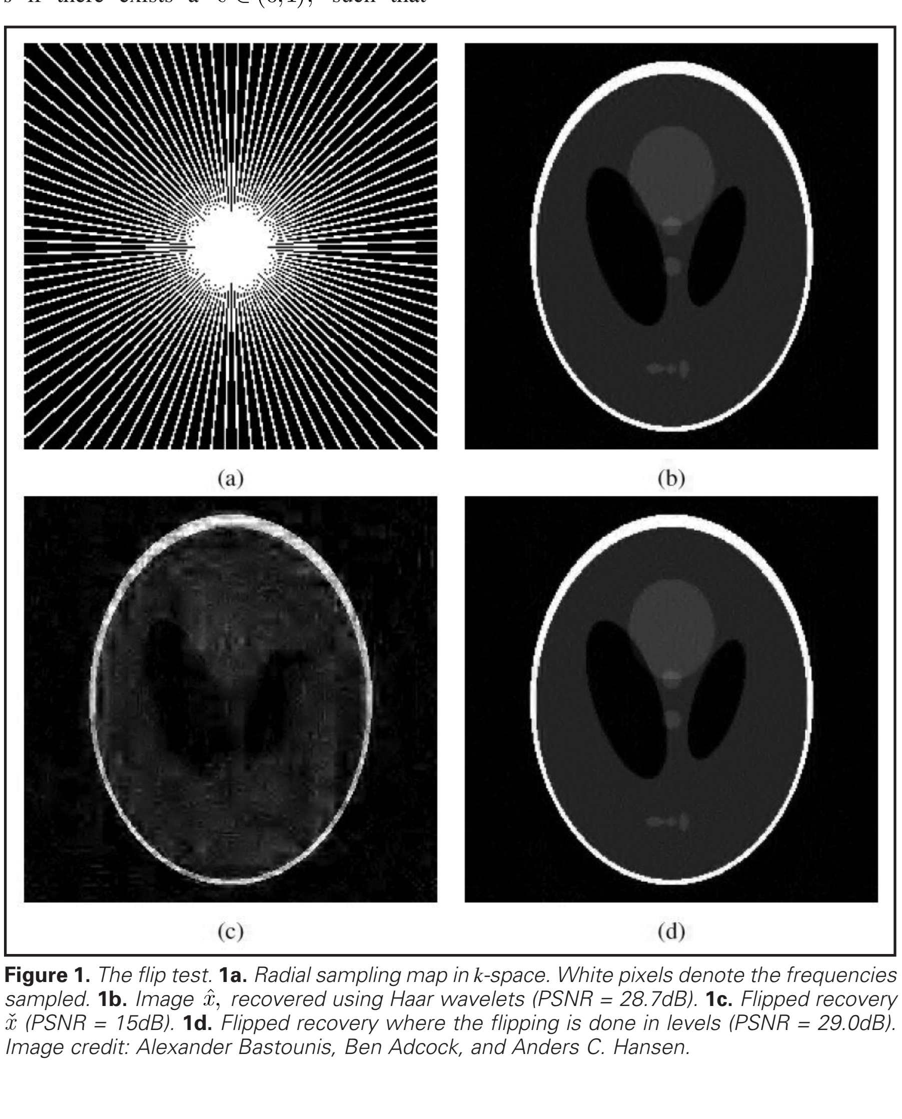


Figure 1. *The flip test.* **1a.** *Radial sampling map in k-space. White pixels denote the frequencies sampled.* **1b.** *Image $\hat{x}$, recovered using Haar wavelets (PSNR = 28.7dB).* **1c.** *Flipped recovery $\check{x}$ (PSNR = 15dB).* **1d.** *Flipped recovery where the flipping is done in levels (PSNR = 29.0dB). Image credit: Alexander Bastounis, Ben Adcock, and Anders C. Hansen.*

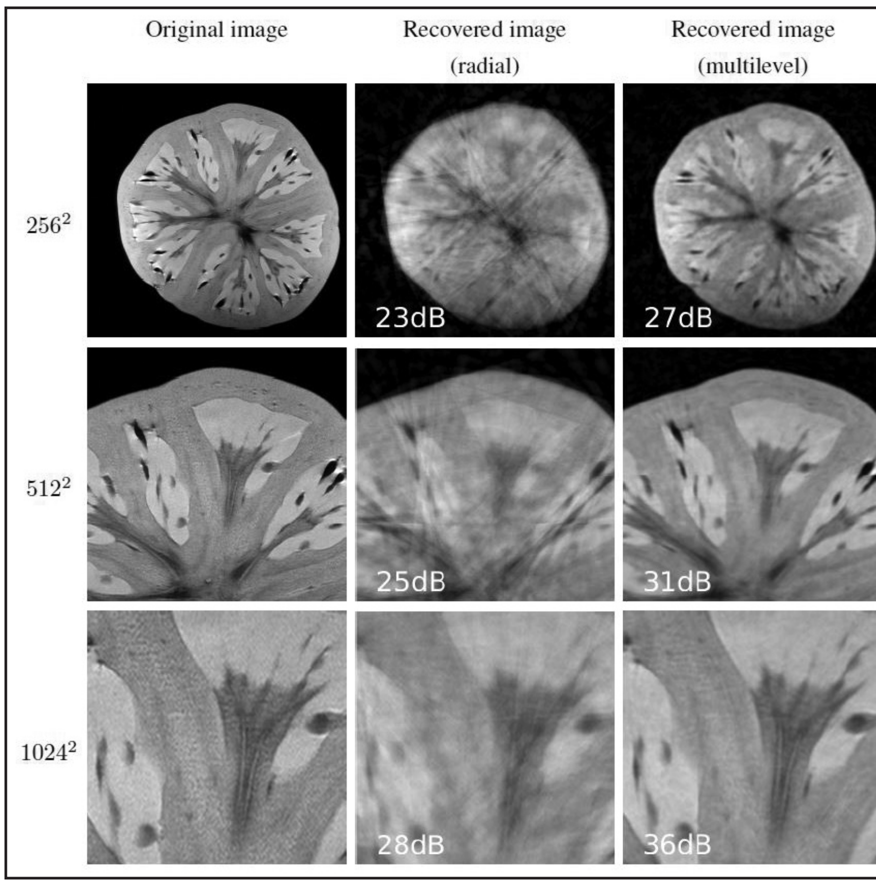


Figure 3. *Compressed sensing using 6.25% Fourier measurements at various resolutions. Original image courtesy of Andy Ellison, recovered images by Alexander Bastounis, Ben Adcock, and Anders C. Hansen.*

To examine the extent to which this theory explains the results observed in Figure 1, we perform the following experiment, known as the *flip test* [1]. Let $Q \in \mathbb{C}^{N\times N}$ be a permutation matrix and define the permuted coefficients $c_p = Qc$ and corresponding image $x_p = \Phi c_p$. Now let $\hat{c_p}$ be the coefficients recovered by solving (1) with $y = PFx_p$, and define $\check{c} = Q^{-1}\hat{c_p}$ and $\check{x} = \Phi\check{c}$. Since permutations do not affect sparsity, coefficients c_p are s-sparse and image x_p has an s-sparse representation in the wavelet basis. Hence, if matrix A has the RIP, one would expect both the *unflipped* reconstruction $\hat{x}$ and *flipped* reconstruction $\check{x}$ to yield similar recoveries of the original image x.

Figure 1d demonstrates that this is not the case. The flipped reconstruction—in this example, the permutation simply reverses the ordering of the coefficients—is significantly worse than the unflipped reconstruction. We therefore conclude that the RIP cannot hold. Additionally, since certain sparse vectors are recovered better than others, *distribution* of the wavelet coefficients is crucial to recovery.

Classical wavelet theory can intuitively explain these results. As illustrated in Figure 2a (on page 1), wavelet coefficients, when arranged according to dyadic scales, are sparser at finer scales than coarser scales. Moreover, wavelets at a given scale are essentially concentrated in square annular regions of k-space (see Figure 2c). The radial sampling pattern samples less densely in regions corresponding to the fine scales, where the image is more sparse, and more densely at coarse scales, where the image is less sparse. However, if the coefficients are permuted (see Figure 2b), too many coefficients exist at fine scales (compared to the number of high-frequency samples) to ensure good recovery.

A Levels-based Framework

To provide a more comprehensive compressed sensing framework, the approach in [1, 2] first replaces the global concepts of sparsity and incoherence with suitable local quantities. Specifically, let r be a number of levels and $\mathrm{M} = (M_1,\ldots,M_r)$, where $1 \le M_1 < \ldots < M_r = N$, a vector of *sparsity levels*. These may typically correspond to wavelet scales. Rather than a single sparsity index s, the new model considers a vector $\mathrm{s} = (s_1,\ldots,s_r)$ of local sparsities, with s_k as the sparsity at the kth level. We refer to vector $x \in \mathbb{C}^N$ with this sparsity pattern as (s,M)-*sparse in levels*.

Note that permutations performed in Figure 2c (on page 1) destroy sparsity in levels but not global sparsity. Conversely, Figure 2d demonstrates that permutations within scales do not unduly alter the reconstruction quality, thus demonstrating the appropriateness of the (s,M)-sparsity model.

A modified version of the RIP helps analyze recovery with this model [2]. Matrix A has the *RIP in levels* (RIPL) of order (s,M) if there exists a $\delta \in (0,1)$, such that

$$(1-\delta)\|x\|^2_{\ell^2} \le \|Ax\|^2_{\ell^2} \le (1+\delta)\|x\|^2_{\ell^2},$$

for all (s,M)-sparse vectors x.

Much like the standard RIP, if A has the RIPL (for small $\delta_{\mathrm{s,M}}$), then all (s,M)-sparse vectors can be robustly recovered by solving (1).

Returning to Fourier sampling with wavelet sparsity, this novel sparsity model calls for a new type of sampling, known as *multilevel random subsampling*. The idea is to break up the rows of the matrix U into levels [1], following the block-diagonal structure illustrated in Figure 2d (on page 1). Specifically, we introduce *sampling levels* $\mathrm{N} = (N_1,\ldots,N_r)$, where $1 \le N_1 < \ldots < N_r = N$, and a vector $\mathrm{m} = (m_1,\ldots,m_r)$ of local numbers of measurements. Within each sampling level, m_k samples are chosen uniformly at random. Using this approach, one can show that the matrix (2) satisfies the RIPL (in the one-dimensional setting) [7], provided

$$m_k \approx C\left(s_k + \sum_{\substack{l=1\\ j\neq k}}^{r} 2^{-|l-k|}s_l\right)\cdot \log^3(N)\cdot\log^2(s), \qquad k = 1,\ldots,r. \quad (3)$$

That is, the number of measurements m_k required to capture each wavelet scale should be roughly proportional to the corresponding sparsity s_k.

Applications and Benefits

By refining the sparsity model and sampling procedure, this framework not only explains the observations of the flip test but also significantly enhances compressed

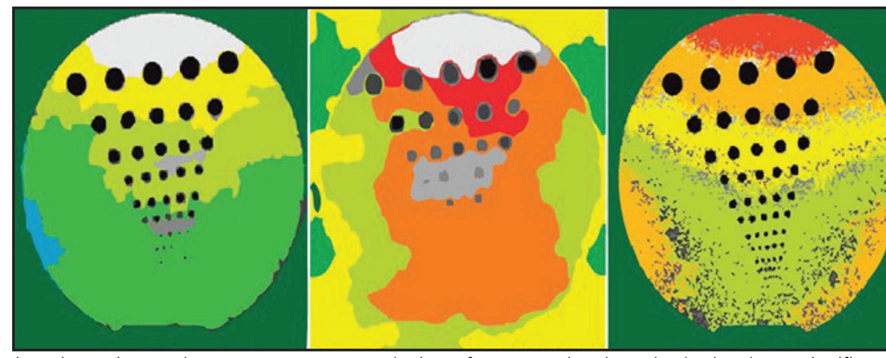
See **Compressed Sensing** *on page 7*

Convergence in Imaging Sciences

By Eric Miller

For those of us trained in the mathematical sciences, the notion of convergence has a very specific connotation of coming together without ever moving apart (you know the drill: for every ϵ, there exists a δ such that ….). Here I will focus on a more expansive idea of convergence as the basis for divergence — an explosion of new developments and opportunities, at least in the area of imaging sciences. In recent years, imaging sciences has experienced a rather marked increase in fundamentally new advances enabled by the convergence of technological capabilities and interests, some of which are far removed from the world of applied mathematics.

One source of these developments is the wealth of novel—and in many cases, challenging—sensor technologies. The role of applied mathematics in sensor data modeling and processing is certainly not new. The search for hydrocarbons in Earth's subsurface is perhaps the quintessential example of a highly successful collaboration, dating back to at least the 1970s, between those who built sensors (seismic, acoustic, electromagnetic, etc.) and those tasked with modeling and extracting information from the resulting data. However, the quantity and diversity of sensing technologies that have emerged over the past 10 to 15 years is unprecedented. This is perhaps most evident in the general field of optics. From the single-pixel camera developed by Richard Baraniuk's group at Rice University to the gigapixel camera created by David Brady and his team at Duke University, there is no shortage of examples that intimately tie a new sensing method's success with a suite of associated mathematical models and processing methods. Biomedical applications are driving many of these advances. Laura Waller (University of California, Berkeley), Vasilis Ntziachristos (Technical University of Munich), and Lihong Wang (California Institute of Technology) are developing sensing systems that represent some of the most compelling instances of new imaging modalities employing light; in many cases these are "mixed" with sound, giving rise to improvements in both computational imaging methods and the mathematical analysis accompanying the resulting inverse problems. Peter Kuchment's (Texas A&M University) work on the analysis of photoacoustic imaging problems is a good example.



Imaging sciences has seen a recent explosion of computational methods that have significantly advanced the field. This figure shows compressed sensing reconstruction of a magnetic resonance image. Compressed sensing allows images and signals to be reconstructed from small amounts of data. Here, a Split Bregman method is applied to a compressed sensing problem that arises in magnetic resonance imaging. Image credit: Tom Goldstein and Stanley Osher, adapted from [1].

As physicists and engineers create new sensing methods, convergence is also evident within the mathematics community proper. I would like to focus specifically on the area of inverse problems in which a physical model stands between the data one possesses and the information one desires. In some cases, researchers can develop closed form, analytical methods for turning sensor data into images; the best known among these is convolution back-projection (also called filtered back projection or Radon inversion), originally developed for parallel beam X-ray computed tomography and since generalized in a variety of mathematically-interesting and practically-useful directions, including fan beam, cone beam, and helical scan cases. Implementation of these methods generally has low computational overheard, i.e., they are "fast." However, they also tend to apply to very specific sensing geometries and assumptions about the underlying physics, a fact that hasn't deterred their recent, rather remarkable expansion. Using sophisticated ideas in microlocal analysis, mathematicians and their colleagues—including Todd Quinto (Tufts University), Margaret Cheney (Colorado State University), and Bill Lionheart (University of Manchester)—have developed interesting methods for solving imaging problems when the sensing geometry is less than ideal. They have demonstrated the utility of these ideas not only in the case of X-ray imaging, but more broadly to problems of wave propagation, including sonar, radar, and more recently Compton scatter imaging. I would be remiss to not acknowledge that these recent advancements build on an existing base of work dating back at least (to the best of my knowledge) to the efforts of folks like Gregory Beylkin, Douglas Miller, Michael Oristaglio, and others who in the 1980s pioneered many of these ideas in the context of geophysical sensing for hydrocarbon exploration.

A large body of work in the use of numerical/computational techniques for solving inverse problems also exists. The intent is to discretize the physical model and pose image formation as the answer to a variational problem in which a "good" solution balances fidelity to the data against information one possesses in addition to the data itself, often quantified mathematically in terms of some degree of smoothness of the image or its derivatives. Interpreting the variational problem through a probabilistic lens (a technique known for decades) has recently produced some rather compelling results in the area of uncertainty quantification (UQ), where the output is not a single image but an entire probabilistic model. This model offers insight into not only the most likely image but also the level of confidence in such an estimate, which is valuable information when deciding how best to collect new data. The work of Omar Ghattas's group at the University of Texas and Youssef Marzouk's group at the Massachusetts Institute of Technology provide great examples of this line of inquiry.

In contrast to the analytical methods, the computational ones do provide more flexibility for addressing nonideal problems in which sensors may be arbitrarily located, the underlying medium inhomogeneous, or the physics not well approximated in a "nice" manner. The price is computational: this approach typically demands the solution to a high-dimensional, non-convex optimization problem, where both gradient information and the evaluation of the

See **Convergence** *on page 9*

Compressed Sensing

Continued from page 5

sensing performance in various imaging settings. By following sparsity patterns of the wavelet coefficients, one can exploit (3) to develop sampling patterns that target the sparsity in levels structure and thereby enhance reconstruction quality.

We conclude by demonstrating these benefits in several practical settings (see [10] for further experiments). First, Figure 3 (on page 5) compares the recovery of a magnetic resonance image at various resolutions from Fourier measurements, taken according to radial and multilevel sampling patterns. Multilevel sampling is consistently superior to radial sampling because it better targets the image's sparsity structure. This benefit also increases with the resolution, since the multilevel sampling pattern aligns increasingly well with the wavelet coefficients' *asymptotic* sparsity.

From this latter observation we conclude the following: instead of using compressed sensing at lower resolutions to reduce acquisition time, one can best realize the full benefits by subsampling at higher resolutions and seeking to improve image quality. In other words, compressed sensing is most beneficial as a *resolution enhancer*. Figure 4 demonstrates this effect. For a fixed budget of measurements, subsampling from higher resolutions yields a vastly superior reconstruction when compared to full sampling at low frequencies. Both Siemens—a leading manufacturer of MRI scanners [13]—and [10] further verify this phenomenon in a practical MRI setting.

Finally, Figure 5 considers a class of problems informally known as *compressive imaging* [11]. In these problems—the applications of which include single-pixel [5] and lensless imaging, infrared imaging, and fluorescence microscopy [12]—one can choose the measurement matrix A, provided that its entries are binary. In this case, a randomly-subsampled Hadamard transform with scrambled columns is a standard choice for A. This is a computationally-efficient procedure whose performance mimics that of random Gaussian sampling; it is near-optimal for recovering sparse vectors. It may therefore seem surprising that the reconstruction quality can be improved. However, as Figure 5 shows, a multilevel subsampled Hadamard transform (without column scrambling) does precisely this. Even though wavelet coefficients are sparse, the procedure targets the image's fine details (captured by the fine-scale wavelet coefficients) to achieve a significant performance gain.

Acknowledgments: The authors thank Andy Ellison for the MR images used in Figures 3 and 4.[1]

[1] See https://insideinsides.blogspot.co.uk

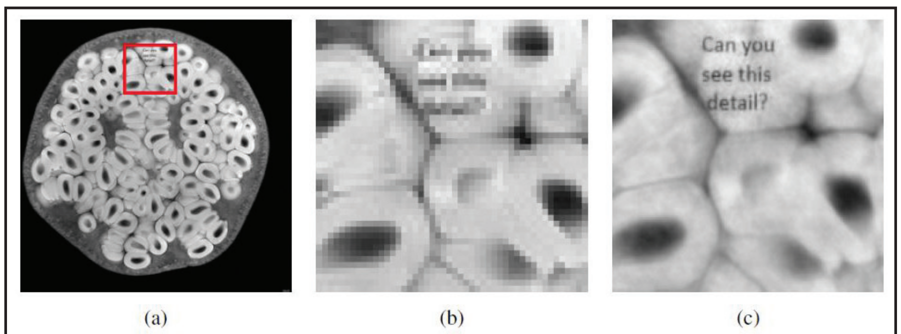


Figure 4. *Resolution enhancing in MRI.* **4a.** *Original image with small synthetic detail added.* **4b.** *Linear recovery from the lowest $m=256^2$ Fourier measurements.* **4c.** *Compressed sensing with multilevel subsampling using $m=256^2$ measurements. Original image courtesy of Andy Ellison, recovered images by Alexander Bastounis, Ben Adcock, and Anders C. Hansen.*

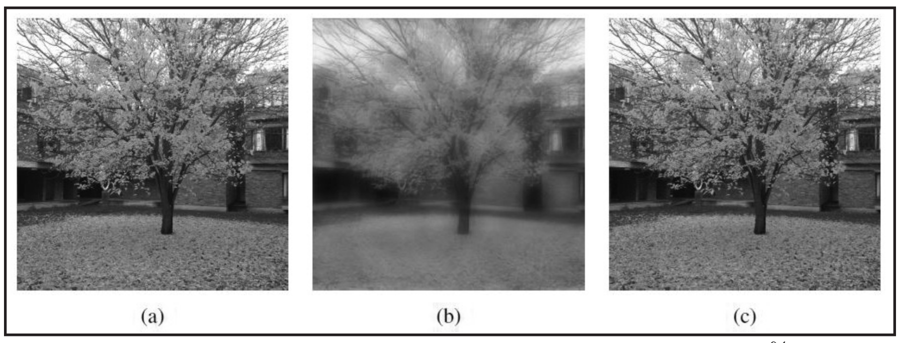


Figure 5. *Compressive imaging.* **5a.** *Original image.* **5b.** *Recovery from $m=16.5\%$ scrambled Hadamard measurements.* **5c.** *Recovery from $m=16.5\%$ multilevel subsampled Hadamard measurements. Image credit: Alexander Bastounis, Ben Adcock, and Anders C. Hansen.*

References

[1] Adcock, B., Hansen, A.C., Poon, C., & Roman, B. (2017). Breaking the coherence barrier: A new theory for compressed sensing. *Forum of Mathematics, Sigma*, 5.

[2] Bastounis, A., & Hansen, A.C. (2017). On the absence of uniform recovery in many real-world applications of compressed sensing and the restricted isometry property and nullspace property in levels. *SIAM J. Imaging Sci., 10*(1), 335-371.

[3] Candès, E.J., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory, 52*(2), 489-509.

[4] Donoho, D.L. (2006). Compressed sensing. *IEEE Trans. Inform. Theory, 52*(4), 1289-1306.

[5] Duarte, M.F., Davenport, M.A., Takhar, D., Laska, J., Kelly, K., & Baraniuk, R.G. (2008). Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag., 25*(2), 83-91.

[6] Foucart, S., & Rauhut, H. (2013). *A Mathematical Introduction to Compressive Sensing: Applied and Numerical Harmonic Analysis*. New York, NY: Birkhauser.

[7] Li, C., & Adcock, B. (2017). Compressed sensing with local structure: uniform recovery guarantees for the sparsity in levels class. *Appl. Comput. Harmon. Anal.* (to appear). *arXiv:1601.01988*.

[8] Lustig, M., Donoho, D.L., Santos, J.M., & Pauly, J.M. (2008). Compressed Sensing MRI. *IEEE Signal Process. Mag., 25*(2), 72-82.

[9] Poon, C. (2015). On the role of total variation in compressed sensing. *SIAM J. Imaging Sci., 8*(1), 682-720.

[10] Roman, B., Calderbank, R., Adcock, B., Nietlispach, D., Bostock, M., Calvo-Almazán, I., Graves, M., & Hansen, A.C. (2017). Undersampling improves fidelity of physical imaging and the benefits grow with resolution. *Preprint*.

[11] Romberg, J. (2008). Imaging via compressive sampling. *IEEE Signal Process. Mag., 25*(2), 14-20.

[12] Studer, V., Bobin, J., Chahid, M., Moussavi, H., Candès, E., & Dahan, M. (2011). Compressive fluorescence microscopy for biological and hyperspectral imaging. *Proc. Natl Acad. Sci., 109*(26), 1679-1687.

[13] Wang, Q., Zenge, M., Cetingul, H.E., Mueller, E., & Nadar, M.S. (2014). Novel sampling strategies for sparse MR image reconstruction. *Proc. Int. Soc. Mag. Res. in Med., 22*.

Alexander Bastounis is a final-year Ph.D. student in the Applied Functional and Harmonic Analysis group in the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge. Ben Adcock is an assistant professor at Simon Fraser University. He received the Leslie Fox Prize for Numerical Analysis in 2011, an Alfred P. Sloan Research Fellowship in 2015, and the CAIMS-PIMS Early Career Award in 2017. Anders C. Hansen is the head of the Applied Functional and Harmonic Analysis group at the University of Cambridge, where he is a reader (associate professor) in mathematics. He is also a Royal Society University Research Fellow.

Unhidden Figures

By Karthika Swamy Cohen

It's not easy to be a black woman in a field dominated by white men. But certain factors—impactful mentors, inspiring role models, and supportive academic environments—alleviate some of the challenges. That was the overriding theme in the stories of four African American women who spoke to a packed auditorium at the "Hidden Figures" panel, which took place at the 2017 SIAM Annual Meeting, held in Pittsburgh, Pa., this July.

"For many of us who are underrepresented, the question is not what encouraged you to study mathematics," Shelby Wilson, assistant professor at Morehouse College, said. "It's *who* encouraged you to study mathematics."

Plenty of whos inspired three of the panelists—current female mathematicians navigating a male-dominated field—including fellow panelist and hidden figure Christine Darden, a former human computer-turned-aerospace engineer at NASA's Langley Research Center. Darden's early findings in the 1960s and 70s resulted in a revolution of aerodynamics design to produce low-boom sonic effects. She was featured in *Hidden Figures: The American Dream and the Untold Story of the Black Women Mathematicians Who Helped Win the Space Race*, a 2016 book by Margot Lee Shetterly that inspired the film of the same name.

"I'm fortunate enough to have many of my whos in the room right now, many of them across the country," Wilson said. "And I'm eternally grateful that my grandmother and other hidden figures like Dr. Darden not so much directly affected me, but put my whos into place."

For Erica Graham, assistant professor at Bryn Mawr College, mentors didn't always share the same background or experiences, but offered support and encouragement nonetheless. "Although the collection of teachers I've had over the years was less diverse, I was never made to question my right to be where I was, so I never did," Graham said. "The first best decision I made was to choose a college where I had professors who saw my potential and did what they could to draw it out of me. And the most important aspect of completing my graduate education was the support network I developed through various mentors and fellow graduate students, characterized more by collaboration than competition with faculty who clearly wanted us to succeed."

This is not to say that Graham didn't face the challenges that come with being an African American woman in a field dominated by people from other racial and ethnic backgrounds. However, she tried to ignore the differences. "In an environment where stereotype threat is very real and imposter syndrome runs rampant, it was essential for me to shed as much unnecessary weight as I possibly could and pretend as though I was just any other graduate student," she said. "It wasn't always easy."

Talitha Washington, associate professor of mathematics at Howard University, attributes much of her success to three mentors who continue to inspire her. "As they always say, behind every successful woman there are a few good men," she began. Washington spoke about her college mentor and primary inspiration behind her decision to pursue a Ph.D. in mathematics. "I had zero aspirations of going to graduate school," she said. "I wanted to work in business because that is what I knew. Then along came Dr. Jeffrey Ehme, who took me on as a student researcher in my senior year and forced—yes, he forced me—to apply to graduate school. Had he not made me apply, I wouldn't be here today."

Life in graduate school at the University of Connecticut, however, was difficult for Washington. "People always asked me what country I was from," she recalled. "I was totally confused. I told them, 'I'm from Indiana, is that a country?'" But then she found her second supportive mentor, Joe McKenna. "I remember sitting in his office devastated by the environment, the workload, and the graduate life," Washington continued. "He told me, 'You outwork anyone here, you are good.' Those words encouraged me to work even harder and see it through."

In 2001, Washington became the first African American to graduate from the University of Connecticut with a Ph.D. in mathematics. Her third noteworthy mentor, Ronald Mickens of Clark Atlanta University, has helped guide much of her professional career. "He gave me a book on difference equations when I was an undergrad, but it accumulated a little bit of dust," Washington said. "Little did I know that I would dust it off a decade later, and we would actually begin doing research together."

Wilson also spoke highly of those who inspired her. "I've never lacked role models and mentors in mathematics," she said. "I've had mentors who are mathematical biologists like me, women like me, black like me, socioeconomically privileged like me, and many more who are not like me in all the ways that you can think of."

See **Unhidden Figures** *on page 9*

Unhidden Figures

Continued from page 8

While influential mentors come in many stripes, sometimes it helps when role models look like you and have undergone and relate to similar life experiences. For this reason, Graham talked fondly of the Enhancing Diversity in Graduate Education (EDGE)[1] program, intended to increase the number of women and minorities who complete graduate programs in the mathematical sciences. "EDGE was the first time in my life I met black women who had already done what I was getting ready to spend the next several years of my life doing," she said. "And for me, there is such a normality to their presence that I sometimes forget about the actual composition of the mathematical community. In the years since that first summer at EDGE, I've been lucky enough to have a network of friends, colleagues, mentors, and research collaborators, without whom I'd be less likely to be where I am today." EDGE's success demonstrates how being in an environment devoid of "otherness" perhaps helped these young women excel.

Washington spoke of the benefits of attending a black women's college where she did not have to singularly represent an entire race. "I went to Spelman College in Atlanta, Ga., which is a black women's college, and I was immediately blessed with not being the only black female in the class," she recalled. "I did not have to explain my race or how I looked. I could simply learn and absorb copious amounts of knowledge that was centered around my perspective as a black woman in a multicultural world."

Wilson noted that women's colleges inspire the same kind of unfettered productivity. "If you put women in an environment where their womanhood is not questioned—the term 'woman' is non-descript, where a woman is not an active description of you—you have the opportunity to flourish in mathematics," she said. "It allows you to be more confident and comfortable. I think these [women's] programs do away with some of the stigma and are helping women."

What made Darden unique among the panelists wasn't just that she is profiled in a best-selling book, but also that she came to mathematics at a time when there were no such support groups for women and hardly any female black role models in the field.

None of that stopped her, however. When her father insisted she become a teacher to ensure employment after college, Darden earned her teacher's certificate while taking several math classes so she could pursue her dream. "I still had this idea that I wanted to do something that was not teaching," she confessed. "I took my 30 hours of education in student teaching, but I took about 16-18 hours of math that I didn't have to take as electives. I went on and taught school for a couple of years, and while I was teaching I started going to Virginia State University, taking in-service classes in higher mathematics."

When Darden learned that women were getting passed over for promotions during her time at NASA, she stood up to her supervisors. "I worked as a human computer for five years until I found out that a lot of the men we were supporting—the engineers—were math majors too," she said. She asked her supervisor about transferring to an engineering area, but to no avail. A few months later, Darden decided to talk to somebody higher up. "I went to a director who was about three to four levels higher and said, 'I just want to know why men and women coming here with the same background are assigned to such different areas — you are putting the women in the computer pools where they don't write papers and don't get promoted. Men with the same degree are going into engineering, they are working on their own projects, they are writing papers, and they are getting promoted.' He said nobody ever asked that question before."

Christine Darden (formerly of NASA) addresses a packed room during the "Hidden Figures" panel at the 2017 SIAM Annual Meeting, which took place in Pittsburgh, Pa., this July. Other speakers included (from left to right) Erica Graham, Talitha Washington, and Shelby Wilson. SIAM photo.

Within two weeks of that conversation, Darden was promoted and transferred to an engineering section. "That's what I really felt was the beginning of my career at NASA," she said.

Washington referenced a pioneer in the African American community who still motivates her. "At Howard University, our departmental meetings are held in a room with a picture of Dr. Elbert Frank Cox, the first black person in the world to receive a Ph.D. in mathematics," she said. "I didn't learn about Cox until I was fully grown, post-Ph.D. Even though we grew up in the same neighborhood in Evansville, Ind., for me, he was hidden. He spent most of his career at Howard, as [have] I. I often stare at his photo and I'm encouraged just by looking at it, I'm encouraged to continue the pursuit of mathematics and justice."

Young mathematicians from underrepresented communities today owe much to people like Cox and Darden—and Dorothy Vaughan, Mary Jackson, and Katherine Johnson, all of whom were featured in *Hidden Figures*—who paved the way for future generations.

"As a mathematician, mother, and activist, I hope that we all remain unhidden so that our children can see that they too can become mathematicians," Washington said.

Karthika Swamy Cohen is the managing editor of SIAM News.

[1] https://www.edgeforwomen.org/

Convergence

Continued from page 6

cost function require the solution of tens, hundreds, or even thousands of discretized partial differential equations. Thus, regardless of whether one seeks a single solution to a variational problem or an entire UQ model, the corresponding mathematical challenges tend to center around problems in numerical linear algebra (including fast linear systems solves, preconditioning, and reduced order modeling) as well as optimization. Recent studies focus on theory and methods that use randomization as a tool for reducing system size and hence processing complexity.

Moving forward, I have to believe that there will be opportunities to bring together these two separate approaches to imaging—the analytical and the computational—because the strengths of one balance the shortcomings of the other. Samuli Siltanen (University of Helsinki) and his collaborators developed complex geometric optics methods for an array of inverse problems, most notably electrical impedance tomography, which may offer a clue to a possible union. Their work is based on a rather deep and analytically elegant mathematical formulation of the physics of the problem, which requires the solution to a numerical inverse problem at one crucial point. Perhaps these ideas will lead to progress in combining some of the aforementioned areas. Or maybe a totally different variety of insight will be necessary. Regardless of the details, one thing is certain: imaging sciences will continue to provide relevant, intellectually-stimulating problems that allow applied mathematicians and their collaborators to impact the field for years to come.

References

[1] Goldstein, T., & Osher, S. (2009). The Split Bregman Method for L1-Regularized Problems. *SIAM Journal on Imaging Sciences, 2*(2), 323-343.

Eric Miller is a professor and chair of the Department of Electrical and Computer Engineering at Tufts University. He is the chair of—and the SIAM News liaison for—the SIAM Activity Group on Imaging Sciences.

NSF PIC Math Grant Sponsors Data Analytics Workshop

By Thomas Wakefield

A quick online search for "data scientist" reveals a wealth of sites that convey the growth of this profession, rank it as one of the most desirable careers, and accentuate the field's value. As data science continues to flourish, it is increasingly important that mathematicians understand both the field itself, and the way in which college and university faculty train and prepare students for entry.

To that end, 70 mathematics and statistics faculty gathered at Brigham Young University (BYU) in late May for a four-day workshop that introduced data analytics, machine learning, statistics, and programming to faculty with little to no expertise in these areas. Michael Dorff of BYU and Suzanne Weekes of Worcester Polytechnic Institute (WPI) organized the workshop as part of Preparation for Industrial Careers in Mathematical Sciences (PIC Math). PIC Math is a program of SIAM and the Mathematical Association of America, with support provided by the National Science Foundation (NSF).

The PIC Math Workshop on Data Analytics exposed attendees to the techniques and software used in data analytics problems (particularly classification problems), which strengthened their understanding of data analytics and machine learning. Participants also learned how to identify data analysis projects and mentor undergraduate students on such projects.

Randy Paffenroth of WPI conducted tutorials on the use of Python and sample code to implement various supervised classification algorithms, particularly k-nearest neighbors, decision trees, support vector machines, and linear discriminant analysis. He emphasized the trade-off between bias and variance; stressed the importance of cross-validation in machine learning; and overviewed techniques to prepare data for machine learning, including construction of a validation set, principal component analysis, and bootstrapping. Participants worked in groups to implement a classification algorithm on a data set from the University of California, Irvine Machine Learning Repository.[1] They presented their results and discussed pitfalls and issues that arose in their implementation of the algorithms on the given data. "The main focus of the PIC Math Workshop was to give everyone the opportunity to get their hands dirty working with real data, no matter their background," Paffenroth said. "I thought that the team exercises were the most important part of the workshop, and I hope all of the attendees enjoyed them!"

Participants at the PIC Math Workshop on Data Analytics, held this May at Brigham Young University, listen to Randy Paffenroth introduce the Python programming language and its machine learning packages. Photo courtesy of Mikayla Sweet of the Mathematical Association of America.

Jonathan Nolis, director and lead of Insights and Analytics at Lenati, offered valuable advice for faculty interested in preparing and engaging students pursuing data analytics. "It's great that PIC Math exists and math professors are getting more involved in data science," he said. "These professors are receiving valuable training in the sorts of jobs available in industry and how to prepare students for them. As someone who frequently hires for analytics jobs, I value students who understand the connection between mathematics and industry."

The workshop was but one activity of the PIC Math grant, which increases students' awareness and pursuit of career options outside academia. The grant exposes students to real problems from business, industry, and government, and provides faculty with the support necessary to offer students these opportunities. More information about PIC Math can be found on the website.[2]

Acknowledgments: The NSF supports the PIC Math program with NSF grant DMS-1345499.

Thomas Wakefield is a professor of mathematics and statistics at Youngstown State University and a fellow of the Society of Actuaries. With support from the PIC Math program and industrial sponsors, Wakefield runs undergraduate mathematics research courses that allow students the opportunity to work on problems originating from industry.

[1] http://archive.ics.uci.edu/ml/

[2] http://www.maa.org/picmath

Professional Opportunities and Announcements

*Send copy for classified advertisements and announcements to **marketing@siam.org**; For rates, deadlines, and ad specifications visit **www.siam.org/advertising**.*

*Students (and others) in search of information about careers in the mathematical sciences can click on "Careers and Jobs" at the SIAM website (**www.siam.org**) or proceed directly to **www.siam.org/careers**.*

Dartmouth College

Department of Mathematics

The Department of Mathematics announces a tenure-track opening for the 2018-2019 academic year. There is a preference for a junior appointment, but we welcome applicants suitable for a high initial rank. The successful applicant will have a research profile with a concentration in applied or computational mathematics, and have demonstrated ability to work across disciplines such as biology, physics, or computer science, as well as those at the Geisel School of Medicine at Dartmouth, the Norris Cotton Cancer Center, or the Thayer School of Engineering. Applicants should apply online at **www.mathjobs.org**, **Position ID: APACM #10567**. Applications received by **December 15, 2017** will receive first consideration. For more information about this position, please visit our website at **https://www.math.dartmouth.edu/activities/recruiting/**.

Dartmouth College is an equal opportunity/affirmative action employer with a strong commitment to diversity and inclusion. We prohibit discrimination on the basis of race, color, religion, sex, age, national origin, sexual orientation, gender identity or expression, disability, veteran status, marital status, or any other legally protected status. Applications by members of all underrepresented groups are encouraged.

California Institute of Technology

Computing and Mathematical Sciences Department

The Computing and Mathematical Sciences (CMS) Department at the California Institute of

See **Professional Opportunities** *on page 11*

Professional Opportunities

Continued from page 10

Technology (Caltech) invites applications for a tenure-track faculty position. CMS is a unique environment where innovative, interdisciplinary, and foundational research is conducted in a collegial atmosphere. Candidates in all areas of applied and computational mathematics, computer science, and statistics are invited to apply. Areas of interest include (but are not limited to) scientific computing, optimization, statistics, probability, networked systems, control and dynamical systems, robotics, theory of computation, security, privacy, machine learning, and algorithmic economics. In addition, we welcome applications from candidates who have demonstrated strong connections between computer science, engineering, and applied mathematics, and to other fields such as the physical, biological, and social sciences.

A commitment to world-class research, as well as high-quality teaching and mentoring, is expected. The initial appointment at the assistant professor level is for four years, and is contingent upon the completion of a Ph.D. degree in applied mathematics, computer science, engineering, or a related field.

Applications will be reviewed beginning **November 15, 2017**, and applicants are encouraged to have all their application materials on file by this date. For a list of documents required and full instructions on how to apply online, please visit **http://www.cms.caltech.edu/search**. Questions about the application process may be directed to **search@cms.caltech.edu**.

We are an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation, gender identity, national origin, disability status, protected veteran status, or any other characteristic protected by law.

Mathematical Sciences Research Institute

Director of Advancement & External Relations

m/Oppenheim Associates is assisting the Mathematical Sciences Research Institute (MSRI) in the search for a new **Director for Advancement & External Relations**. The organization seeks a proven fundraising professional to shape significant endowment campaigns, drive annual contributed revenue, and engage new donors from various fields of endeavor who will support fresh research into mathematics. For more information, please review the complete position description at **http://www.moppenheim.com/wp-content/uploads/MSRI-Director-for-Advancement-position-description-Final.pdf**.

MSRI is the world's preeminent center for collaborative research in mathematics, and advances research into the key unsolved mathematical problems that underlie core mathematics and applications in the physical sciences, economics, engineering, computing, communications, statistical analysis, and the global financial system. The prime objective for this position is to develop strategies and campaigns that:

- **Develop a five-year campaign to add $27 million to the current MSRI endowment of $23 million**, and then a second campaign to add another $50 million to the endowment by 2030, for a total endowment of $100 million.
- **Deliver annual fund contributions of $4.5 million by 2020** (up from about $3.5 million in 2017).
- **Diversify and broaden donor support for MSRI** through use of conventional and social media marketing that communicates MSRI's relevance to donors of different sectors (technology, finance, economics, medicine, etc.).

Founded in 1982 and located in Berkeley, Calif., MSRI has a 2017 budget of $9.9 million, a staff of 21, and a Board of 33. The Advancement Team includes the executive director, an associate director, admin support, and subcontractors including but not limited to web development, videography, and design of communications materials.

For additional information or to apply, please contact Mark Oppenheim or Patrick Salazar at **info@moppenheim.com**.

More information can be found on the MSRI website (http://www.msri.org/web/cms) and on Numberphile (https://www.youtube.com/user/numberphile), an MSRI-supported YouTube channel with short and entertaining videos on mathematics.

INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES

The *Institute for Computational Engineering and Sciences (ICES)* at *The University of Texas at Austin* is searching for exceptional candidates with expertise in computational science and engineering to fill several Moncrief endowed faculty positions at the Associate Professor level and higher. These endowed positions will provide the resources and environment needed to tackle frontier problems in science and engineering via advanced modeling and simulation.

This initiative builds on the world-leading programs at ICES in Computational Science, Engineering, and Mathematics (CSEM), which feature 16 research centers and groups as well as a graduate degree program in CSEM. Candidates are expected to have an exceptional record in interdisciplinary research and evidence of work involving applied mathematics and computational techniques targeting meaningful problems in engineering and science. For more information and application instructions, please visit: www.ices.utexas.edu/moncrief-endowed-positions-app/.

This is a security sensitive position. The University of Texas at Austin is an Equal Employment Opportunity/Affirmative Action Employer.

THE UNIVERSITY OF TEXAS AT AUSTIN

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
NATURWISSENSCHAFTLICHE FAKULTÄT

The Faculty of Sciences invites applications for a

W3 Professorship for Applied Mathematics

at the Department of Mathematics to be filled by 1 October 2018. The successful candidate is expected to represent the field of applied mathematics adequately in teaching and research and have an outstanding international profile in this field. The research activities of the professorship will focus on optimization with partial differential equations and optimal control of dynamical systems. Proven expertise in interdisciplinary research as well as active involvement in Transregio / CRC 154 and other research activities at FAU are expected.

For further information and the application guidelines please see
https://www.fau.eu/university/careers-at-fau/professorships/

Please submit your complete application documents (CV, list of publications, list of lectures and courses taught, copies of certificates and degrees, list of third-party funding) online at https://berufungen.fau.de/ by **31 October 2017**, addressed to the Dean of the Faculty of Sciences. Please contact nat-dekanat@fau.de with any questions.

FAMILIE IN DER HOCHSCHULE
Mitglied des Best Practice-Clubs

www.fau.de

Williams College

The Williams College Department of Mathematics and Statistics invites applications for a **tenure-track position in Statistics**, beginning fall 2018, at the rank of assistant professor (a more senior appointment is possible under special circumstances). The candidate should have a Ph.D. in Statistics or a closely related field by the time of appointment. We are seeking candidates who show evidence and/or promise of excellence in teaching students from diverse backgrounds and a strong research program that can engage undergraduate students. The candidate will become the sixth tenure-track statistician in the department, joining a vibrant and innovative group of statisticians with an established statistics major. For more information on the Department of Mathematics and Statistics, visit *http://math.williams.edu/*.

At Williams, we are committed to building a diverse and inclusive community where members from all backgrounds can live, learn, and thrive. In your application materials, we ask you to address how your teaching, scholarship, mentorship and/or community service might support our commitment to diversity and inclusion. Candidates may apply via http://apply.interfolio.com/43065 by uploading a cover letter addressed to Professor Klingenberg, a curriculum vitae, a teaching statement, a description of your research plans, and three letters of recommendation on teaching and research.

Expectations: The teaching load is two courses per 12-week semester and a winter term course every other January. The candidate will be expected to teach introductory statistics, core courses for the statistics major, and electives in their area of expertise. The successful candidate will establish an independent research program that results in scholarly publications. Williams College provides broad support for start-up funds, funding for student research assistants, faculty professional development funds, and a shared computer cluster for parallel computation.

Review of applications will begin on or after **November 1st** and will continue until the position is filled. All offers of employment are contingent upon completion of a background check. Further information is available at *https://faculty.williams.edu/prospective-faculty/background-check-policy/*.

Williams College is a coeducational liberal arts institution located in the Berkshire Hills of western Massachusetts with easy access to the culturally rich cities of Albany, Boston, and New York City. The College is committed to building and supporting a diverse population of approximately 2,000 students, and to fostering an inclusive faculty, staff and curriculum. Williams has built its reputation on outstanding teaching and scholarship and on the academic excellence of its students. Please visit the Williams College website, ***http://www.williams.edu/***.

Tenured/Tenure-track Faculty Positions

Cornell University's School of Operations Research and Information Engineering (ORIE) seeks to fill multiple tenured/tenure-track faculty positions for its Ithaca campus. We will primarily consider applicants with research interests in the areas of integer programming and financial engineering, though we welcome strong applicants from all research areas represented within ORIE, especially those in resonance with the College of Engineering Strategic Areas: **www.engineering.cornell.edu/research/strategic**.

Requisite is a strong interest in the broad mission of the School, exceptional potential for leadership in research and education, an ability and willingness to teach at all levels of the program, and a Ph.D. in operations research, mathematics, statistics, or a related field by the start of the appointment. Salary will be appropriate to qualifications and engineering school norms.

Cornell ORIE is a diverse group of high-quality researchers and educators interested in probability, optimization, statistics, simulation, and a wide array of applications such as e-commerce, supply chains, scheduling, manufacturing, transportation systems, health care, financial engineering, service systems, and network science. We value mathematical and technical depth and innovation, and experience with applications and practice. Ideal candidates will have correspondingly broad training and interests.

Please apply online at **https://academicjobsonline.org/ajo/jobs/9654** with a cover letter, CV, statements of teaching and research interests, sample publications, at least three reference letters and, for junior applicants, a Doctoral transcript. We strongly encourage applicants attending the INFORMS annual meeting to submit all application materials by October 15, 2017. All applications completed by November 15, 2017 will receive full consideration, but we urge candidates to submit all required material as soon as possible. We will accept applications until we fill the positions.

ORIE and the College of Engineering at Cornell embrace diversity and seek candidates who can contribute to a welcoming climate for students of all races and genders. Cornell University seeks to meet the needs of dual career couples, has a Dual Career program, and is a member of the Upstate New York Higher Education Recruitment Consortium to assist with dual career searches.

Visit **www.unyherc.org/home** to see positions available in higher education in the upstate New York area.

Diversity and Inclusion are a part of Cornell's heritage. We are a recognized employer and educator valuing AA/EEO, Protected Veterans, and Individuals with Disabilities. We strongly encourage qualified women and minority candidates to apply.

Separating Shape and Intensity Variation in Images

By Line Kühnel, Stefan Sommer, Akshay Pai, and Lars Lau Raket

One of the main goals of image analysis is to describe variations between images. Understanding variations in a population of images allows researchers to classify new subjects as potential members of the population distribution. Such classification is especially essential in the field of computational anatomy [2], where locating similarities and differences in imaging data for sick and healthy populations is an important task. Recognizing these patterns aids in the detection of sickness or quantification of disease severity in newly-observed subjects. Our recent work presents a novel, flexible class of mixed-effects models that separates variation in images [3]. The framework combines image analysis with theory from functional data analysis, and uses statistical methods to simultaneously estimate the template image and variation effects.

The two largest modes of image variation are *intensity variation* and variation in *point correspondence*. Point correspondence, or *warp* variation, is the shape variability compared to the template image. Intensity variation is the spatially correlated variation left after compensating for the true warping effect. For example, the spatial intensity variation could describe either systematic error in an image or anatomical variation, such as tissue density or texture.

Intensity variation has previously been considered a nuisance that could be handled by pre-processing images. Hence, the analysis focused primarily on estimating shape variation/warp effects. This approach can be problematic because one does not account for uncertainty of the intensity modifications—made when pre-processing the images—in the subsequent analysis. This underestimates the intensity variation, producing an analysis with overconfident estimates regarding the precision of the estimated warp effects [9]. Several works have considered the intensity variation as an integral part of the model, using statistical methods to simultaneously model intensity and warp variation [1, 5]. Compared to other models, the proposed mixed-effects model distinguishes between systematic intensity variation and independent noise, making it useful for denoising images. This model simultaneously estimates intensity and warp variation by an alternating maximum-likelihood estimation and prediction; as a result, the model chooses the most likely separation of the random effects, based on patterns of variation in the data. It consequently prevents the problem of bias in parameter estimates of the random effects. Figure 1 illustrates the idea behind the model.

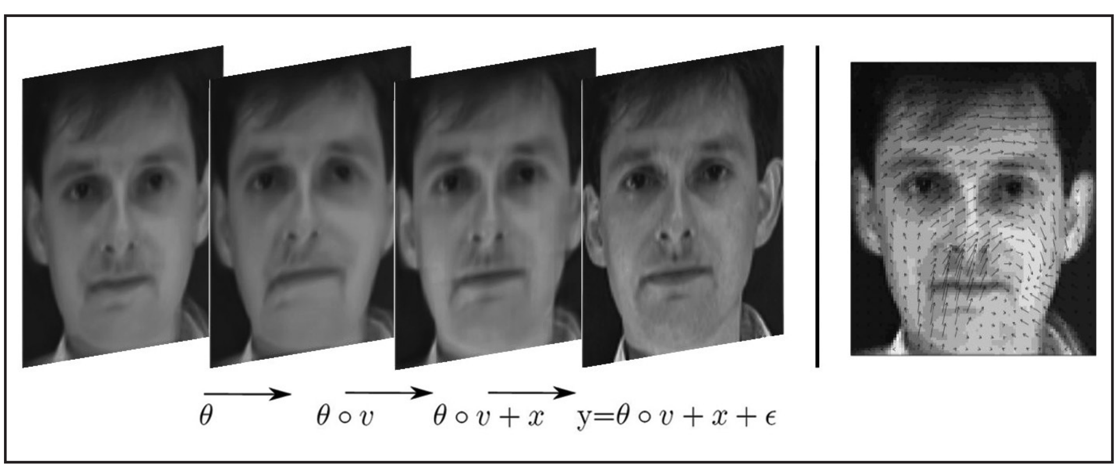


Figure 1. *Fixed and random effects.* **Left.** *The template (θ: leftmost) perturbed by random warp ($\theta \circ v$: 2nd from left) and warp+spatially correlated intensity ($\theta \circ v + x$: 3rd from left), together with independent noise ϵ (y: 4th from left).* **Right.** *The warp field v brings the observation into spatial correspondence, with θ overlayed on the template. Estimation of template and model hyperparameters is conducted simultaneously with prediction of random effects, allowing for separation of the different factors in the nonlinear model. Image courtesy of [3].*

Images included in the model are considered spatial functional data from $\mathbb{R}^2$ to $\mathbb{R}$. Each observation y_i is a vector of function values on a regular lattice, with $m = m_1 m_2$ grid points (s_j, t_k), i.e., $y = (y_i(s_j, t_k))_{j,k}$ for $j = 1, \ldots m_1$, $k = 1, \ldots m_2$. The proposed model is defined by

$$y_i(s_j, t_k) = \theta(v_i(s_j, t_k)) + x_i(s_j, t_k) + \varepsilon_{ijk}, \tag{1}$$

where the template image $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ is modeled as a fixed effect. The spatially correlated intensity variation x_i is assumed to arise from a Gaussian field $(x_i(s_j, t_k))_{j,k} \sim \mathcal{N}(0, S)$, while the last term is Gaussian white noise $\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2)$. The model differentiates spatially correlated noise and independent white noise, as both can be present in the observed images. In this model, warping functions $v_i\colon \mathbb{R}^2 \to \mathbb{R}^2$ are assumed to satisfy

$$v_i(s,t) = v(s,t,w_i) = \begin{pmatrix} s \\ t \end{pmatrix} + \varepsilon_{w_i}(s,t) \tag{2}$$

for $\varepsilon_{w_i}\colon \mathbb{R}^2 \to \mathbb{R}^2$, denoting a coordinate-wise bilinear spline interpolation of a displacement vector, $w_i \in \mathbb{R}^{m_w^1 \times m_w^2 \times 2}$, on a lattice spanned by $s_w \in \mathbb{R}^{m_w^1}$, $t_w \in \mathbb{R}^{m_w^2}$. The displacement vectors w_i are modeled as random effects following a normal distribution $w_i \sim \mathcal{N}(0, C)$.

Parameter estimation in (1) is based on a maximum-likelihood approach. The model contains nonlinear transformations of the random warping functions $v_i(s_j, t_k)$; hence, a closed-form expression for the likelihood function is unavailable. To overcome this, we iteratively linearize the likelihood function around the predicted warp parameters w_i, which at each step enables the use of methodology from linear mixed-effects models. Linearization of nonlinear mixed-effects models is effective in standard situations and present in many software packages [4, 7, 8]. However, one cannot use these existing implementations with the large sizes of image data and combination of linear and nonlinear random effects in (1). We model the covariance matrix for the spatially correlated intensity effect with a sparse inverse to make the computations feasible. This modeling choice is equivalent to assuming conditional independence between pixels that are far from each other, given all other pixels — often a reasonable assumption.

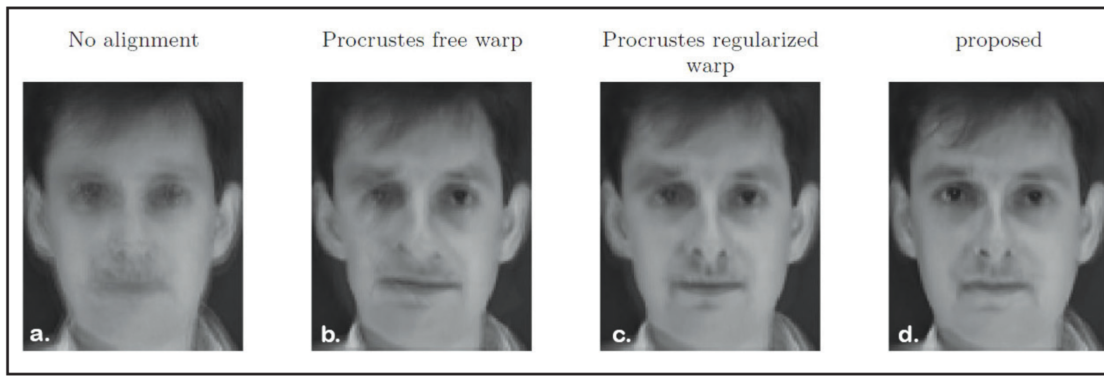


Figure 2. *Estimates for the fixed effect θ using different models. The models used to calculate the estimates are as follows:* **2a.** *Model assuming no warping effect and Gaussian white noise for the intensity model.* **2b.** *2a. with a free warping function based on 16 displacement vectors.* **2c.** *2a. with a penalized estimation of warping functions.* **2d.** *The full model (1). Image courtesy of [3].*

To illustrate the model's applications, we have analyzed data from 10 facial images of the same person [10] and 50 two-dimensional sagittal magnetic resonance imaging (MRI) slices of brains from the Alzheimer's Disease Neuroimaging Initiative database [6].

Figure 2 compares an estimated template image of the proposed model to frequently-used models for the face data. Sharpness and representativeness of the estimates increase when going from left to right. Upon determining estimates of the template image θ, the covariance matrix of the intensity function S, and the covariance matrix for the warping effect C, we can use the parameters to split observations in different variations. Figure 3 depicts an example of a brain observation, split into the different variation effects.

In conclusion, we have presented a class of models that avoids the classical problems of biased estimation of variations caused by sequential preprocessing in image analysis. We achieved this by simultaneously modeling the major modes of random variation in object shape and recorded intensities. This in turn allowed maximum-likelihood-based estimation of parameters, which would have been otherwise manually tuned. The maximum-likelihood-based approach leads to parameter estimates that induce a most likely separation of shape and intensity variation.

Acknowledgments: This work was supported by the Centre for Stochastic Geometry and Advanced Bioimaging (CSGB), funded by a grant from the Villum Foundation.

References

[1] Allassonniere, S., Durrleman, S., & Kuhn, E. (2015). Bayesian mixed effect atlas estimation with a diffeomorphic deformation model. *SIAM Journal on Imaging Sciences, 8*(3), 1367-1395.

[2] Joshi, S., Davis, B., Jomier, M., & Gerig, G. (2004). Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage, 23*, 151-160.

[3] Kühnel, L., Sommer, S., Pai, A., & Raket, L.L. (2017). Most likely separation of intensity and warping effects in image registration. *SIAM Journal on Imaging Sciences, 10*(2), 578-601.

[4] Lindstrom, M.J., & Bates, D.M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics, 46*, 673-687.

[5] Ma, J., Miller, M.I., Trouvé, A., & Younes, L. (2008). Bayesian template estimation in computational anatomy. *NeuroImage, 42*, 252-261.

[6] Pai, A., Sommer, S., Sorensen, L., Darkner, S., Sporring, J., & Nielsen, M. (2015). Kernel bundle diffeomorphic image registration using stationary velocity fields and Wendland basis functions. *IEEE Transactions on Medical Imaging, 35*(6), 1369-1380.

[7] Pinheiro, J., & Bates, D. (2006). *Mixed-effects models in S and S-PLUS: Statistics and Computing*. New York, NY: Springer Science & Business Media.

[8] Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2007). Linear and nonlinear mixed effects models. *R package version, 3*, 57.

[9] Raket, L.L., Sommer, S., & Markussen, B. (2014). A nonlinear mixed-effects model for simultaneous smoothing and registration of functional data. *Pattern Recognition Letters, 38*, 1-7.

[10] Samaria, F.S. (2002). *The Database of Faces*. AT&T. Retrieved from http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

Line Kühnel, Stefan Sommer, and Akshay Pai are affiliated with the Department of Computer Science at the University of Copenhagen in Denmark. Lars Lau Raket is affiliated with the Department of Mathematical Sciences at the University of Copenhagen.

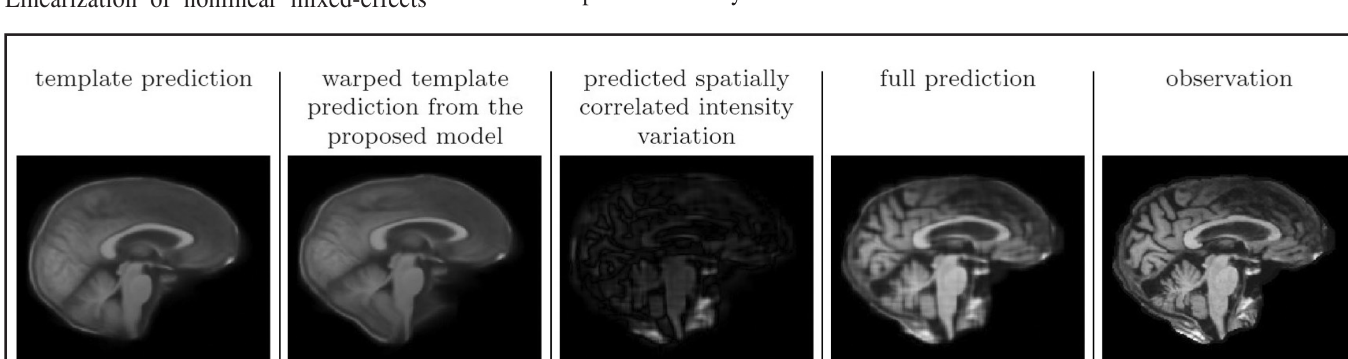


Figure 3. *Model predictions of a mid-sagittal brain slice (shown on the far right). From left to right: The estimated template for the proposed model, the warped template from the proposed model, the absolute value of the predicted spatially correlated intensity variation from the proposed model, and the full prediction. Image courtesy of [3].*