Guest Editorial: Special Section on CVPR 2013

William T. Freeman, *Fellow, IEEE*, Richard Szeliski, *Fellow, IEEE*, and Gregory D. Hager, *Fellow, IEEE*

THIS special section contains selected papers from the premier annual computer vision conference, IEEE Computer Vision and Pattern Recognition (CVPR), June, 2013, jointly sponsored by the IEEE and the Computer Vision Foundation. In 2013, there were 1,816 submissions to CVPR, from which 472 were selected for oral or poster presentation. From the accepted CVPR 2013 papers, we, program chairs from that conference, selected a set that were especially appreciated by the reviewers for their technical contributions, and for their expected impacts to the field going forward. We invited those authors to submit a longer manuscript to this special section, which resulted in the six articles appearing here. We found these papers to be delightful, and we hope that you will, too. We thank the reviewers of these manuscripts, who made many helpful suggestions for changes to the papers.

The first paper, by C. Lawrence Zitnick, Ramakrishna Vedantam, and Devi Parikh, "Adopting Abstract Images for Semantic Scene Understanding" studies high-level issues of semantics, salience, and memorability with clip-art images, rather than photographs, thus avoiding the difficult tasks of object detection and recognition from pixel data. This world of cartoon characters lets the researchers explore various image interpretation questions, such as what variety of scenes convey a given semantic description, or how the occurrence of any object contributes to an image's meaning.

In the next paper, "Photometric Ambient Occlusion for Intrinsic Image Decomposition," by Daniel Hauagge, Scott Wehrwein, Kavita Bala, Noah Snavely, the authors study a static image observed under varying lighting conditions. By measuring the ratio of two different intensity averages, the authors can infer the surface albedo, the illumination image, and a quantity called the ambient occlusion—how much light reaches the surface without being blocked by other parts of the scene. This is a powerful approach to disentangle shading, reflectance, illumination for a static scene.

The third paper, "Map-Based Probabilistic Visual Self-Localization," by Marcus A. Brubaker, Andreas Geiger, and Raquel Urtasun, addresses self-localization, given visual data and a map. The authors assume that GPS data is not available or is corrupted (as it can be in urban environments). Their algorithm uses video data from the vehicle in combination with an online, crowd-sourced street map. Within less than a minute of driving, from an initial region of localization containing kilometers of roads, the algorithm can localize the car to within a few meters. The corresponding CVPR paper received recognition as the Best Paper, runner-up.

"Leveraging the Wisdom of the Crowd for Fine-Grained Recognition," by Jia Deng, Jonathan Krause, Michael Stark, and Li Fei-Fei, uses human labelers to learn where the important features are for the recognition of fine-grained object categories. The authors construct an online game to elicit from humans features useful for fine-grained object recognition. These learned features significantly improve performance in fine-grained object recognition over previous state of the art.

The next paper, by Uwe Schmidt, Jeremy Jancsary, Sebastian Nowozin, Stefan Roth, and Carsten Rother, titled, "Cascades of Regression Tree Fields for Image Restoration," is an extended version of the CVPR paper that received the Best Student paper award. While discriminative approaches for image enhancement can be powerful, they may need to be trained for every possible degradation condition, a major drawback. For example, a discriminatively trained image deblurring algorithm would need to be pre-trained for every possible motion blur. Schmidt et al. show that, for Gaussian models of image corruption, such instance-specific parameters can be separated from the discriminative training. For deblurring, this allows coping with arbitrary blur kernels at test time without needing to retrain the model. They use this approach for image denoising and non-blind deblurring, achieving state-of-the-art image restoration results.

The final paper is "Intrinsic Scene Properties from a Single RGB-D Image" by Jonathan T. Barron and Jitendra Malik. The authors model the scene as shapes under a mixture of illuminations. From a single RGB image and a corresponding depth image, their algorithm produces an improved depth map, plus a set of surface normals, a reflectance image, a shading image, and a spatially-varying model of illumination, providing a significant step toward full low-level interpretation from a single image.

We hope you enjoy reading these papers as much as we have.

William T. Freeman is the Thomas and Gerd Perkins professor of Electrical Engineering and Computer Science (EECS) at MIT, and a member of the Computer Science and Artificial Intelligence Laboratory (CSAIL) there. He was an associate department head of EECS from 2011 to 2014. He is on leave from MIT, leading a computer vision research group in Google in Cambridge, MA. His current research interests include Bayesian models of visual perception, computational photography, and motion magnification. He received outstanding paper awards at computer vision or machine learning conferences in 1997, 2006, 2009, and 2012, and test-of-time awards for papers from 1990 and 1995. Previous research topics include steerable filters and pyramids, orientation histograms, the generic viewpoint assumption, color constancy, computer vision for computer games, and belief propagation in networks with loops. He is active in the program or organizing committees of computer vision, graphics, and machine learning conferences. He was the program co-chair for ICCV 2005, and for CVPR 2013, and is an editor of *Foundations and Trends in Computer Graphics and Vision*, and is an IEEE Fellow.

- *W.T. Freeman is with MIT and Google. E-mail: billf@mit.edu.*
- *R. Szeliski is with Facebook. E-mail: szeliski@fb.com.*
- *G.D. Hager is with Johns Hopkins University. E-mail: hager@cs.jhu.edu.*

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2016.2529898

Richard Szeliski received the PhD degree in computer science from Carnegie Mellon University, Pittsburgh, in 1988 and joined Facebook in 2016. He is the director in the Computational Photography group at Facebook. He is also an affiliate professor at the University of Washington. He has done pioneering research in the fields of Bayesian methods for computer vision, image-based modeling, image-based rendering, and computational photography, which lie at the intersection of computer vision and computer graphics. His research on photo tourism, photosynth, and hyperlapse are exciting examples of the promise of large-scale image and video-based rendering. Prior to Facebook, he was at Microsoft Research for twenty years, the Cambridge Research Lab of Digital Equipment Corporation for six years, and several other industrial research labs. He has published over 150 research papers in computer vision, computer graphics, neural nets, and numerical analysis, as well as the books *Computer Vision: Algorithms and Applications and Bayesian Modeling of Uncertainty in Low-Level Vision*. He was a program committee chair for CVPR 2013 and ICCV 2003, served as an associate editor of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* and on the Editorial Board of the *International Journal of Computer Vision*, and as a founding editor of *Foundations and Trends in Computer Graphics and Vision*. He is a member of the NAE, and a fellow of the ACM and IEEE.

Gregory D. Hager received the BA degree in mathematics and computer science summa cum laude at Luther College in 1983, and the MS and PhD degrees in 1986 and 1988, respectively, from the University of Pennsylvania. He is the Mandell Bellmore professor of computer science at Johns Hopkins University and the founding director in the Malone Institute for Engineering in Healthcare. His research interests include computer-integrated medicine, collaborative and vision-based robotics, time-series analysis of image data, and medical applications of image analysis and robotics. He has served on the editorial boards of *IEEE TRO*, *IEEE PAMI*, and *IJCV*, and served as a program chair of CVPR 2013 and general chair of ICCV 2015. He was a Fulbright fellow at the University of Karlsruhe, and was on the faculty of Yale University prior to joining Johns Hopkins. He served as the deputy director in the National Science Foundation (NSF) Engineering Research Center for Surgical Systems and Technology, and as the chair of Computer Science from 2010 to 2015. He is also the chair in the Computing Community Consortium, a board member of the Computing Research Association, and is a member of the governing board of the International Federation of Robotics Research. In 2014, he received a Hans Fischer fellowship in the Institute of Advanced Study of the Technical University of Munich, where he also holds an appointment in Computer Science. He is a fellow of the IEEE for his contributions to vision-based Robotics and a fellow of the MICCAI Society for his contributions to imaging and his work on the image guidance and computer-aided analysis of surgical technical skill.