

Guest Editorial: Best of CVPR 2015

Kristen Grauman, Eric Learned-Miller, Antonio Torralba, and Andrew Zisserman

CVPR 2015 was held in Boston, Massachusetts, at a time of unprecedented growth and excitement in computer vision. With more than 2,000 submissions and 2,800 attendees, the conference exceeded previous records for both. In addition to strong research contributions, the conference saw new levels of industry participation, ushering in an era in which computer vision is being widely deployed in applications as diverse as image search, virtual reality, and autonomous driving.

This special issue of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* highlights some of the best research from CVPR 2015, presenting work from many of the highest rated conference papers, with additional details, extensions, and new results. The papers were vetted through the standard *TPAMI* review process, with the CVPR 2015 Program Chairs acting as guest editors.

The papers in this special issue can be grouped into themes that give a snapshot of this unique time in computer vision. Not surprisingly, many of the best papers represent new techniques and ideas in convolutional neural networks.

One theme explores new convolutional architectures for image segmentation, rather than classification. This is represented by the two papers: “Object Instance Segmentation and Fine-grained Localization using Hypercolumns”, by Bharath Hariharan, Pablo Arbelaez, Ross Girshick, and Jitendra Malik; and “Fully Convolutional Networks for Semantic Segmentation” by Evan Shelhamer, Jonathan Long, and Trevor Darrell.

A second significant new theme is deep learning for vision and language, for example generating captions for images. Three of the top papers in this area are included in this special issue: “Show and Tell: Lessons learned from the 2015 MSCOCO Image Captioning Challenge” by Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan; “Deep Visual-Semantic Alignments for Generating Image Descriptions” by Andrej Karpathy and Fei-Fei Li; and “Long-Term Recurrent Convolutional Networks for Visual Recognition and Description” by Jeff Donahue, Lisa Anne Hendricks, Marcus Rohrbach, Subhashini Venugopalan, Sergio Guadarrama, Kate Saenko, and Trevor Darrell.

A number of excellent papers are also included that explore connections between 3D geometry or 3D representations and their 2D counterparts. These include: “Learning to Generate Chairs, Tables and Cars with Convolutional Networks” by Alexey Dosovitskiy, Jost Tobias Springenberg, Maxim Tatarchenko, and Thomas Brox; “Building Proteins in a Day: Efficient 3D Molecular Reconstruction” by Marcus

Brubaker, Ali Punjani, and David Fleet; and “Learning Category-Specific Deformable 3D Models for Object Reconstruction” by Shubham Tulsiani, Abhishek Kar, Joao Carreira, and Jitendra Malik.

Two papers introduce new types of imaging technique for inferring material properties of objects. These are: “Visual Vibrometry: Estimating Material Properties from Small Motion in Video”, by Abe Davis, Katherine Bouman, Justin Chen, Michael Rubinstein, Oral Buyukozturk, Fredo Durand and William Freeman; and “Recovering Inner Slices of Layered Translucent Objects by Multi-Frequency Illumination” by Kenichiro Tanaka, Yasuhiro Mukaigawa, Hiruyuki Kubo, Yasuyuki Matsushita, and Yasushi Yagi.

The paper by Tat-Jun Chin, Pulak Purkait, Anders Eriksson, and David Suter titled “Efficient Globally Optimal Consensus Maximisation with Tree Search” introduces an efficient algorithm for global maximization of consensus that is an alternative to traditional randomized approaches such as RANSAC.

Finally, a paper is included that has a novel approach to the important area of human action recognition by using the temporal ordering of video frames, “Rank Pooling for Action Recognition” by Basura Fernando, Efstratios Gavves, Jose Oramas M., Amir Ghodrati and Tinne Tuytelaars.

Together the papers in this special issue provide a window onto the cutting edge research in computer vision, and they illustrate the breadth of ongoing challenges being tackled in the field.

Kristen Grauman
Eric Learned-Miller
Antonio Torralba
Andrew Zisserman
Guest Editors



Kristen Grauman received the PhD degree from the Computer Science and Artificial Intelligence Laboratory, MIT. She is an associate professor in the Department of Computer Science at the University of Texas, Austin. Her research in computer vision and machine learning focuses on visual search and recognition. She is an Alfred P. Sloan research fellow and a recipient of NSF CAREER and ONR Young Investigator awards, the 2013 PAMI Young Researcher Award, the 2013 Computers and Thought Award from the International Joint Conference on Artificial Intelligence, and a 2014 Presidential Early Career Award for Scientists and Engineers (PECASE). She and her collaborators were recognized with the Best Student Paper award at CVPR 2008, the Marr Prize at ICCV 2011, and the Best Application Paper award at ACCV 2016. She serves on the Editorial Board for the *International Journal of Computer Vision (IJCV)* and as an associate editor in chief for the *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. She served as a program chair for CVPR 2015 in Boston.

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.2017.2663859



Erik Learned-Miller received the BA in psychology from Yale University, in 1988, the master's of science degree, in 1997, and the PhD degree, in 2002 from the Massachusetts Institute of Technology, both in electrical engineering and computer science. In 1989, he co-founded CORITechs, Inc., where he co-developed the second FDA cleared system for image-guided neurosurgery. He worked for Nomos Corporation, Pittsburgh, Pennsylvania, for two years as the manager of neurosurgical product engineering. In 2006, he received an NSF CAREER award for his work in computer vision and machine learning. He was a co-Program Chair for CVPR 2015 in Boston.



Antonio Torralba received the degree in telecommunications engineering from Telecom BCN, Spain, in 1994 and the PhD degree in signal, image, and speech processing from the Institut National Polytechnique de Grenoble, France, in 2000. From 2000 to 2005, he spent postdoctoral training at the Brain and Cognitive Science Department and the Computer Science and Artificial Intelligence Laboratory, MIT. He is now a professor of electrical engineering and computer science with the Massachusetts Institute of Technology (MIT). He is an associate editor of the *International Journal in Computer Vision*, and served as program chair for the Computer Vision and Pattern Recognition Conference in 2015. He received the 2008 National Science Foundation (NSF) Career award, the best student paper award at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) in 2009, and the 2010 J. K. Aggarwal Prize from the International Association for Pattern Recognition (IAPR).

Andrew Zisserman is the professor of computer vision engineering in the Department of Engineering Science, University of Oxford.

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.