Guest Editors' Introduction: Special Section on Higher Order Graphical Models in Computer Vision

Karteek Alahari, Dhruv Batra, Srikumar Ramalingam, Nikos Paragios, *Fellow, IEEE*, and Richard Zemel

Many tasks in computer vision, including low-level ones such as image segmentation and stereo estimation, as well as high-level ones such as object recognition and scene understanding, have been modelled as discrete labelling problems. To solve these problems discrete optimization has emerged as an indispensable tool over the last two decades. It is now a routine process to write an explicit energy function, incorporating prior models into it, and then depending on its properties, perform exact or approximate inference. Initial success with this approach derived from formulating a labelling problem in terms of an energy function comprising of unary and pairwise clique potentials. This assumption severely restricts the representational power of these models as they are unable to capture the rich statistics of natural images. More recently, a second wave of success can be attributed to the incorporation of higher order terms that have the ability to encode significantly more sophisticated priors and structural dependencies between variables.

This special section explores the main challenges in this framework—modeling novel priors, learning, inference—and presents innovative solutions in the form of six papers. We, the five guest editors, oversaw the selection process. Each of the submitted manuscripts were reviewed by three experts, who provided excellent feedback. Based on their suggestions, we selected a subset of manuscripts to undergo a major revision. The revised manuscripts were then re-reviewed by the original reviewers before the final decision was made according to their evaluations. We are grateful to all the reviewers for their detailed assessment of manuscripts, which has helped us immensely in compiling a high-quality special issue. We would also like to thank the former editor in chief (EIC) Ramin Zabih and the current EIC David Forsyth for making this special section possible. We are also thankful to the editorial staff for managing the submission process and providing us with assistance as and when it was needed.

The six papers in the special section cover the aspects of modeling novel priors, inference algorithms and parameter learning methods in the context of higher order graphical models.

For information on obtaining reprints of this article, please send e-mail to: reprints.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2015.2434651

The first paper in this section, "Generalized Flows for Optimal Inference in Higher Order MRF-MAP" by C. Arora, S. Banerjee, P. K. Kalra and S. N. Maheshwari, addresses the problem of inference for higher order energy functions, i.e., computing the maximum a posteriori (MAP) solution. While previous methods typically use techniques that reduce higher order functions to pairwise ones before computing the MAP solution, this paper proposes a new method that uses a special graph construction to handle higher order functions in the max-flow/min-cut problem. Experimental results demonstrate that this method is faster than popular algorithms such as Dual Decomposition for higher order functions.

The next paper, "Learning Weighted Lower Linear Envelope Potentials in Binary Markov Random Fields" by S. Gould, addresses an important problem of parameter learning for a special class of higher order functions. In particular, the lower linear envelope potentials are used to model label consistency over a large sets of pixels. The paper shows that it is possible to perform exact energy minimization on these models for binary Markov random fields (MRFs) and the parameters of these lower linear envelope potentials can be learned using a max-margin learning framework.

The paper "Submodular Relaxation for Inference in Markov Random Fields" by A. Osokin and D. Vetrov focuses on developing an approximate inference algorithm for higher order MRFs. The authors propose a submodular relaxation approach based on a Lagrangian relaxation of the original problem. Previous decomposition methods split the original problem into easier solvable problems and use Lagrangian relaxation to make them reach an agreement on their solutions. In contrast, this paper constructs a submodular energy that is minimized within the Lagrangian relaxation.

The paper "Context-Aware Activity Modeling using Hierarchical Conditional Random Fields" by Y. Zhu, N. Nanyak, and A. Roy-Chowdhury presents a hierarchical Conditional Random Field for joint activity localization and recognition in videos. The approach learns with weakly labeled training data, and is formulated as a max-margin problem. In addition to generating activity labels for individual activities, the proposed model simultaneously predicts an optimum structural label for the related activities in the scene.

The paper "Parameter Estimation and Energy Minimization for Region-based Semantic Segmentation" by M. Kumar,

H. Turki, D. Preston, and D. Koller addresses the fundamental problem of parameter estimation in energy-based models from weakly and partially annotated data, in the context of semantic segmentation. The proposed solution is formulated as a Latent Structured SVM, and a segmentation model is learned from (i) images segmented using generic foreground or background classes; (ii) images with bounding boxes specified for objects; and (iii) images labeled to indicate the presence of a class.

This section ends with the paper, "A Hypergraph-based Reduction for Higher-order Binary Markov Random Fields" by A. Fix, A. Gruber, E. Boros, and R. Zabih, which presents an order reduction technique for high-order binary Markov Random Fields. While previous works reduced each factor in the problem independently, the proposed technique uses the underlying hyper-graph structure to reduce all factors simultaneously, leading to better performance both theoretically and experimentally.

All these papers show the diversity of methods and algorithms for developing higher order graphical models in computer vision. We hope that these novel papers will further research in this area, and in particular, in relation to other exciting developments in the computer vision and machine learning areas.

Karteek Alahari
Dhruv Batra
Srikumar Ramalingam
Nikos Paragios
Richard Zemel
Guest Editors

Karteek Alahari received the BTech (with Honours) and MS degrees in computer science from IIIT Hyderabad, in 2004 and 2005, respectively, and the PhD degree from Oxford Brookes University, where he was in the Torr Vision Group, in 2010. He is a researcher in the LEAR team at Inria Grenoble-Rhône-Alpes, and is a recipient of the Google Faculty Research Award 2015. Before joining the LEAR team in 2013, he was a postdoctoral fellow in the WILLOW team at Inria Paris-Rocquencourt and École Normale Supérieure. He has been an associate member of the Visual Geometry Group, University of Oxford since 2006.

Dhruv Batra received the MS and PhD degrees from Carnegie Mellon University in 2007 and 2010, respectively, advised by Tsuhan Chen. He is an assistant professor in the Bradley Department of Electrical and Computer Engineering, Virginia Tech, where he leads the VT Machine Learning & Perception group. He is a member of the Virginia Center for Autonomous Systems (VaCAS) and the VT Discovery Analytic Center (DAC). Prior to joining VT, he was a research assistant professor in Toyota Technological Institute at Chicago (TTIC), a philanthropically endowed academic computer science institute located in the campus of University of Chicago. In past, he has held visiting positions in the Machine Learning Department, CMU, and at CSAIL MIT. He received the Carnegie Mellon Dean's Fellowship (2007), Google Faculty Research Award (2013), Virginia Tech Teacher of the Week (2013), Army Research Office (ARO) Young Investigator Program (YIP) Award (2014), the National Science Foundation (NSF) CAREER award (2014), and Virginia Tech CoE Outstanding New Assistant Professor award (2015). His research is supported by the NSF, ARO, ONR, Amazon, Google, Microsoft, and NVIDIA.

Srikumar Ramalingam received the PhD degree in 2006 from INRIA Rhone Alpes. He received the Marie Curie Fellowship during the PhD degree and his thesis received the INPG best thesis prize and the AFRIF thesis prize (honorable mention) from the French Association for Pattern Recognition. He is a senior principal research scientist at the Mitsubishi Electric Research Lab (MERL). He spent two years in Torr Vision group as a postdoctoral fellow in Oxford Brookes University. He received R&D 100 Award for MELFA 3D Vision system in 2014. His research interests include topics in computer vision, machine learning, robotics and discrete optimization.

Nikos Paragios received the BSc, MSc, and PhD degrees in 1994, 1996, and 2000, respectively. He is a professor of applied mathematics at Ecole Centrale de Paris (since 2005), scientific leader at Inria (since 2008), and a senior fellow of the Institut Universitaire de France (since 2014). He has published more than 200 papers in the most prestigious journals and conferences of his research fields. He currently serves as the editor in chief of the *Computer Vision and Image Understanding Journal* while serving/has served as a member of the editorial boards for numerous journals of his scientific fields (*IEEE T-PAMI*, *IJCV*, *MedIA*, *IVCJ*, *SIAM Journal on Imaging Sciences*, ...). He has received several awards and honors and his scientific interests are in the areas of computer vision, medical imaging, and machine learning. He is a fellow of the IEEE.

Richard Zemel received the BSc degree in history & science from Harvard University in 1984, and the PhD degree in computer science from the University of Toronto in 1993. He is a professor of computer science at the University of Toronto, where he has been a faculty member since 2000. Prior to that he was an assistant professor of computer science and psychology at the University of Arizona, and was a postdoctoral fellow at the Salk Institute and at Carnegie Mellon University. He has received several awards and honors, including a Young Investigator Award from the Office of Naval Research, and six Dean's Excellence Awards at the University of Toronto. He is a fellow of the Canadian Institute for Advanced Research, and a member of the NIPS Advisory Board. His research interests include topics in machine learning, vision, and neural coding. His recent research focuses on structured output models, image-text analysis, and fairness.