

## Editorial: Special Issue on Probabilistic Models for Image Understanding

Bill Triggs · Christopher K.I. Williams

Published online: 11 March 2010  
© Springer Science+Business Media, LLC 2010

Probabilistic models provide a compelling framework for describing image and video content at many levels, from small image patches to large-scale scene and motion structure. There has been a great deal of progress in this area in recent years: model structures and inference techniques have become increasingly sophisticated; reliance on simplistic assumptions such as wholesale independence and Gaussianity has been reduced by widespread use of learning from real examples; the labelling requirements for learning have been progressively weakened to allow the use of massive sets of unlabelled or very weakly labelled images; and different kinds of models have increasingly been combined to cover multiple levels of image structure.

We felt that the time was ripe for a special issue in this area, covering the full range of probabilistic image models. Topics solicited included: low level models (image patches, random fields); object recognition/detection; structural models/image parsing; structured models of human motion; probabilistic frameworks for image representation; efficient algorithms for learning such models; and frameworks and datasets for evaluating such models. We were particularly interested in approaches that exploit rich structure (deep, graphical, hierarchical, compositional, *etc.*) or that minimize the amount of labelled data required to learn new classes, *e.g.* by exploiting latent structure or reusing components or priors.

B. Triggs (✉)  
Laboratoire Jean Kuntzman, Grenoble, France  
e-mail: [Bill.Triggs@imag.fr](mailto:Bill.Triggs@imag.fr)

C.K.I. Williams  
University of Edinburgh, Edinburgh, UK  
e-mail: [ckw@inf.ed.ac.uk](mailto:ckw@inf.ed.ac.uk)

Submissions to the special issue closed in July 2008 and we hoped to complete the review process by autumn 2009. We received 29 submissions, seven of which were immediately rejected as out of scope. Each of the remaining 22 were evaluated by at least three reviewers. Three of these manuscripts (including one on which a guest editor was an author) were handled by other IJCV editors to avoid possible conflicts of interest. We have selected a set of eight papers for the initial part of the special issue. Six submissions are still under review and some of these may appear in a subsequent part of the issue.

The accepted papers cover a diversity of themes ranging over image classification, object detection, articulated motion, image segmentation, image parsing, and evaluation papers. Li Li-Jia and Fei-Fei Li describe an on-line Bayesian learning method for image classification using large image collections on the web, which was tested on Caltech-101 and other datasets. Ashish Kapoor, Kristen Grauman, Raquel Urtasun and Trevor Darrell evaluate a discriminative Gaussian Processes model incorporating multiple kernel function learning for local feature based image classification on the Caltech-4 and Caltech-101 datasets. Camille Vidal and Bruno Jedynak present a low-level probabilistic model for template-based object detection, evaluating it on an MRI brain scan dataset. David A. Ross, Daniel Tarlow and Richard S. Zemel present a generative “stick figure” model for learning 2D or 3D articulated structure and motion from tracked feature points, testing it on an assortment of video sequences. Diane Larlus, Jakob Verbeek and Frédéric Jurie propose a method for category-level image segmentation based on local topic models and random fields that models images as compositions of regions representing single object instances, testing it mainly on the MSRC and PASCAL VOC 2007 datasets. Jake Porway, Qiongchen

Wang and Song Chun Zhu describe a multilayer contextual grammar-based model for whole-image parsing, testing it on 120 aerial images of urban regions from the Lotus Hill dataset.

Finally, two papers focus on evaluation: Tinne Tuytelaars, Christoph H. Lampert, Matthew B. Blaschko and Wray Buntine compare a range of different approaches for

unsupervised learning of image classes on the Caltech-256 and MSRC-2 datasets; and Mark Everingham, Luc Van Gool, Christopher K.I. Williams, John Winn, and Andrew Zisserman give a detailed report on the results of the 2007 PASCAL Visual Object Classes Challenge with respect to the classification and detection problems.