Guest Editorial: Tracking, Detection and Segmentation

R. Bowden · J. Collomosse · K. Mikolajczyk

Published online: 4 September 2014
© Springer Science+Business Media New York 2014

This special issue contains papers selected by representatives of the British Machine Vision Association (BMVA) describing recent advances made against major outstanding challenges in computer vision: detection; tracking and pose estimation; activity recognition; video segmentation; dense correspondence and video synthesis.

Establishing a dense correspondence between images remains an open problem; recent solutions based on randomised search have yielded promising advances in accuracy and efficiency. Our first paper, “PMBP: PatchMatch Belief Propagation for Correspondence Field Estimation” presents a unified framework in which the *Patch Match* algorithm is first described as a variant of belief propagation, and then extended using features of the latter to yield results that are both more accurate than the original *Patch Match* formulation and orders of magnitude faster than previous belief propagation approaches.

“Mixture of Trees Probabilistic Graphical Model for Video Segmentation” by Badrinarayanan, Budvytis and Cipolla, continues the theme of video analysis; proposing a new spatio-temporal representation, the Mixture of Trees for temporally coherent semantic video segmentation. This

approach allows propagation of semantic labels throughout a video sequence to provide joint tracking and segmentation.

Our next paper by Sapienza, Cuzzolin and Torr, “Learning Discriminative SpaceTime Action Parts from Weakly Labelled Videos”, also presents a novel spatio-temporal model for video analysis; a weakly supervised approach to action recognition which employs a deformable spatial bag of features that can deform in both space and time.

In “Using Bounded Diameter Minimum Spanning Trees to Build Dense Active Appearance Models”, Anderson et al. present an approach to producing dense active appearance models for video realistic synthesis. At the heart of the approach is the use of a spanning tree that solves the joint alignment problem and methods for densifying an AAM applying ideas from digital terrain modelling.

The next paper by Tang, Andriluka and Schiele presents an approach to tracking by detection of multiple people in crowded scenes. “Detection and tracking of Occluded people” works on the assumption that in many scenarios, occlusion typically occurs from other people and such occlusions produce characteristic appearance patterns which can be used to improve detection performance.

Continuing the theme of tracking people, our final contribution “Automatic and Efficient Human Pose Estimation for Sign Language Videos” by Charles, Pfister, Everingham and Zisserman presents an approach to arm and hand tracking in videos of singers taken from TV broadcasts. The approach combines co-segmentation and random forest regression and demonstrates robust long term tracking with challenging cluttered backgrounds.

We hope that you find this collection of excellent research interesting and we would like to thank the BMVA for their support in helping to compile this special issue.

R. Bowden · J. Collomosse · K. Mikolajczyk (✉)
CVSSP, University of Surrey, Guildford, Surrey, UK
e-mail: K.Mikolajczyk@surrey.ac.uk

R. Bowden
e-mail: R.Bowden@surrey.ac.uk

J. Collomosse
e-mail: J.Collomosse@surrey.ac.uk