Int J Comput Vis (2017) 122:409–410
DOI 10.1007/s11263-017-1008-8

Recognition, Tracking, and Optimisation

Xianghua Xie[1] · **Mark Jones**[1] · **Gary Tam**[1]

Published online: 23 March 2017

This special issue of the International Journal of Computer Vision contains eight selected contributions that showcase some of the most actively researched areas in Computer Vision, ranging from object recognition and identification, motion analysis and tracking, and optimisation. It also includes examples of real world applications where computer vision offers reliable quantitative solutions. Advances in fundamental methods, such as learning algorithm and feature representation, are essential in performing these tasks, as the papers in this special issue show.

Yu et al. present a novel algorithm for free-hand sketch recognition in "Sketch-a-Net: a Deep Neural Network that Beats Humans". A novel deep neural network architecture suitable for sketch recognition is proposed along with data augmentation strategies. The augmentation technique incorporates the temporal order of strokes and both local deformation of stroke splines at local level and large deformations of entire object at global level. A joint Bayesian fusion method is used to perform the ensemble. The proposed method is shown to outperform the earlier systems and is computationally more efficient.

In their paper "Deep Perceptual Mapping for Cross-Modal Face Recognition", Sarfraz and Stiefelhagen address the issue of thermal-to-visible face recognition. This paper provides a deep learning approach to multi-modal face recognition by matching visible RGB images to infra-red images. The authors propose a method for learning a non-linear mapping of local image descriptors across modalities using deep neural network, which is trained with features calculated from overlapping patches from the visible domain and corresponding features from the thermal domain. The method is very efficient at the testing stage and is suitable for real-time applications. Extensive quantitative evaluation is reported on multiple public datasets, demonstrating significant performance improvements compared to some most recent competing methods.

"Latent Structure Preserving Hashing", authored by Liu et al., introduces a novel method for the hashing of high-dimensional feature representations based on nonnegative matrix factorisation. The image representations are factored into low dimensions, regularised by the KL divergence to enforce a similarity between the distribution of the high dimensional data and low dimensional data. The proposed latent structure preserving hashing is also extended to multi-layer network to learn hierarchical representations. Three publicly accessible image retrieval datasets are used to validate the techniques and superior performances are achieved compared to several state-of-the-art methods.

Huang et al. propose a novel curvature penalised minimal path model in their work "Global Minimum for a Finsler Elastica Minimal Path Approach". This work incorporates the curvature penalisation to the classical geodesic energy by interpreting the Euler elastic bending energy using a novel Finsler elastic metric. The related Eikonal PDEs are solved using the fast marching method. The authors demonstrate the effectiveness of the proposed method on a number of computer vision problems, including interactive image segmentation, perceptual grouping and tubular structure extraction.

In "Learning Optimal Parameters for Multi-target Tracking with Contextual Interactions", Wang and Fowlkes propose a tracking by detection method for multi-target tracking. The authors extend the min-cost flow multi-target tracking by quadratic trajectory interactions in order to capture high order

✉ Xianghua Xie
x.xie@swansea.ac.uk
http://csvision.swan.ac.uk

[1] Department of Computer Science, Swansea University, Swansea SA2 8PP, UK

contextual interactions. The tracking is formulated as a discrete optimisation problem and a principled structured SVM framework is used to allow an end-to-end parameter learning. Two novel greedy extensions to dynamic programming are introduced to handle pairwise interactions with much improved efficiency. The authors also propose a decomposable tracking loss function that takes into account the length and localisation of transition links. Experimental results on two datasets comprehensively demonstrate its effectiveness and efficiency in multi-target tracking.

For detecting motion-based anomaly, Bensch et al. propose a novel approach to elastic registration of 3D and time trajectory patterns in their work "Spatiotemporal Deformable Prototypes for Motion Anomaly Detection". Hierarchical clustering is used to produce dominant motion patterns with various levels of abstraction and to group similar motions to a bundle, which in turn generates super-trajectories that represent groups of low level trajectories that overlap in time. This produces a learnt prototype model for motion pattern detection. A modified hashing approach is proposed to efficiently produce transformation hypotheses that are optimised by spatiotemporal elastic registration. The authors demonstrate the effectiveness of the method on both a new anomaly motion detection dataset and a biological dataset.

Wen et al. integrate the class-agnostic detection proposal method into a kernel based correlation filter tracker in their work "Applying Detection Proposals to Visual Tracking for Scale and Aspect Ratio Adaptability". The authors propose a background suppression optimisation technique in order to minimise the chance of tracking unwanted regions. The proposed kernelled correlation filter with detection proposed tracking is evaluated on a publicly available visual tracking dataset and sequences with significant scale variation and aspect ratio changes.

Finally, Hughes et al. present their work on visual identification of individual great white sharks from dorsal fin imagery in "Automated Visual Fin Identification of Individual Great White Sharks". The authors present a comprehensive system to tackle a specialised application with a challenging dataset. An open contour stroke model is proposed for fin detection and a multi-scale, combinatorial method is used to encode object boundaries. A scoring mechanism based on the local naive Bayesian nearest neighbour classification algorithm is used to perform identification. The method is evaluated on a dataset containing 2456 images of great white sharks. This work provides important insight into the distribution of individuality over the shark fin.

We believe these eight papers in this special issue form a diverse representation of current active research on deep learning, optimisation, tracking, recognition and identification. This issue also showcases a number of interesting applications. We hope the readers will find these papers informative and inspirational. We are grateful to the authors for their contributions and to the reviewers for their constructive comments and suggestions.