Editorial

Guest Editorial: Local Image Descriptors in Computer Vision

Due to their unique blend of theoretical interest and practical appeal, local image descriptors have attracted much attention from the computer vision community during the last twenty years or so. As a matter of fact, local image descriptors play a key role in a multitude of visual tasks dealing with point correspondences among several images, such as the detection, classification, registration and 3D reconstruction of objects and scenes. Moreover, the applications of image matching through local descriptors are innumerable, and include, among others: image and video retrieval in multimedia databases, augmented reality for the film, television and gaming industries, cultural heritage preservation and valorisation, biomedical imaging, assistive computer vision and robotics, autonomous driving, and advanced human-machine interaction.

In the last few years, the simultaneous onset of big data and deep learning, along with the development of faster and powerful hardware, provided a further boost to local image descriptor research. Despite the tremendous progress in both hand-crafted and data-driven descriptors, new challenges and application scenarios have also arisen, demanding even more effective solutions to old and new problems. These encompass, to name only a few: improving time and storage efficiency, providing descriptors with statistical and spatial context awareness, and achieving robustness in the presence of strong viewpoint and lighting changes, especially for general 3D scenes. Other modern research directions point towards designing unbiased evaluation metrics, effective benchmark protocols, and realistic datasets.

This Special Issue includes seven original research papers that cover diverse and significant aspects of local image descriptor research. In particular, the order in which papers appear reflects the main phase they address, in an ideal computational pipeline starting with the localisation of salient points in an image and ending with the incorporation of spatial and temporal features in descriptor construction.

In 'Fusion of visual salience maps for object acquisition', Greenberg *et al.* discuss how to merge different visual saliency maps in order to detect stable and robust image keypoints. This critical and early step, aimed at identifying the characteristic and distinctive image regions, is fundamental for the extraction of good local image descriptors.

In 'Locally lateral manifolds of normalised Gabor features for face recognition', Zaman shows how classical feature descriptors like Gabor filters, properly combined with well-known non-linear dimensionality reduction methods like locally linear embedding, can provide interesting solutions. In the specific case, the resulting approach is shown to improve the accuracy in face recognition.

The problem of speeding-up descriptor computation is addressed in 'SRP-AKAZE: an improved accelerated KAZE algorithm based on the sparse random projection' by Li *et al.* Data dimensionality reduction through sparse random projection is applied to the popular scale invariant feature transform (SIFT), which can be considered the forefather of the current hand-crafted local image descriptors. The resulting descriptor maintains SIFT discriminability yet improving on its computational efficiency.

An important issue is the tradeoff between robustness and computational efficiency in descriptor design. In 'RootsGLOH2: embedding RootSIFT 'square rooting' in sGLOH2', Bellavia and Colombo demonstrate how hand-crafted local descriptors can still achieve accuracy levels comparable with the state-of-the-art deep descriptors by focusing on the improvement of critical issues such as the similarity distance and the patch orientation, although at the expense of increasing the overall computational complexity.

Another example of the need for devising good approaches to handle descriptor orientation is described in 'Local descriptor for retinal fundus image registration' by Ramli *et al*. In this paper, retinal image registration is improved thanks to better correspondences achieved by a new rotationally invariant local descriptor, obtained by considering statistical information on rings around the key points.

Finally, local image descriptors can also be exploited on more than two dimensions. In 'SGHs for 3D local surface description' by Ao *et al.*, a new three-dimensional local descriptor for characterising surfaces is designed and successfully applied. The proposed descriptor is based on the computation of histograms obtained on multiple partitions and interpolations of the local surface area and can be used for 3D object recognition and registration.

The added dimension may not be limited to space but can also include time. In 'Directional dense-trajectory-based patterns for dynamic texture recognition' by Nguyen *et al.*, a new spatio-temporal local descriptor for the representation and the classification of dynamic textures is presented, extending the concept of local vector pattern inside a novel framework, denoted as 'directional dense trajectory patterns.'

We warmly wish all readers to enjoy the papers in this Special Issue and hope that they serve as a source of inspiration for original future work in local image descriptor research.

Guest Editor Biographies

Carlo Colombo obtained a PhD in Robotics from the Sant'Anna School of University Studies and Doctoral Research, Pisa, Italy, in 1996. He is associate professor of Computational Vision at the Department of Information Engineering, University of Florence, Italy, where he leads the local Computational Vision Group. His main research interests are in geometric computer vision and its applications to autonomous robotics, biomedicine and aids to disabled people, advanced human-machine interaction, multimedia systems and image forensics. He has published more than a hundred articles in international journals, books, and conference proceedings. He has been associate/area editor for the journals *Robotics and Autonomous Systems* (2001–2011), *Journal of Multimedia* (2006–2009), and *Computer Vision and Image Understanding* (2009–2012). He was guest editor of the Special Issue 'Large-Scale Computer Vision: Geometry, Inference, and Learning' of the *International Journal of Computer Vision* (2014), and since 2018 he has served as associate editor for *IET Computer Vision*. He was general chair of the 12th European Conference on Computer Vision ECCV 2012 (Florence), and currently serves as tutorial chair for ECCV 2020 (Glasgow). Carlo Colombo is, or has been in the past, program committee member for several international conferences, and a frequent reviewer for leading computer vision journals. He

has also participated in several research and technology transfer projects within regional, national, European and international programs.

Fabio Bellavia received his PhD degree in Mathematics and Computer Science from the University of Palermo, Italy, in 2011. His research interests include computer vision and image processing, and in particular: local image detectors and descriptors, image matching, 3D reconstruction, image mosaicing and color correction. From 2012 to 2019 he joined the Computational Vision Group of the University of Florence, Italy, where he worked on computer vision for autonomous robotics, image forensics and cultural heritage. Since 2019 he is a research associate at the Department of Math and Computer Science, University of Palermo, working on new computer vision methods for cultural heritage. He has co-authored several articles in international journals and conferences, serving either as reviewer or program committee member for many of them. He was among the local organizers of 'Which is Which? Evaluation of local descriptors for image matching in real-world scenarios', a contest held at CAIP2019. Fabio Bellavia is currently member of CVPL (associated to IAPR) and GRIN.