Trends and Controversies

Hugo Proença
University of Beira Interior,
IT: Instituto de Telecomunicações

Mark Nixon
University of Southampton

Michele Nappi
Università di Salerno

Performing covert biometric recognition in surveillance environments has been regarded as a “grand” challenge, considering the adversity of the conditions where recognition should be carried out (e.g., poor resolution, bad lighting, off-pose and partially occluded data). This special issue compiles a group of approaches to this problem.

Progress in biometrics research has been concentrated on improving the robustness of recognition against poor quality data, consistent with less constrained data acquisition environments and protocols. Among the most obvious ambitions of this research topic is the development of automata able to work effectively in conditions that are currently confined to visual surveillance, so called “recognition-in-the-wild.” In such conditions data is acquired covertly, from large distances, and has poor discriminability due to limited resolution, blur, and other degradation factors.

One interesting possibility to acquire data in visual surveillance scenarios is the use of PTZ (pan-tilt-zoom) devices. According to this concept, the QUIS-CAMPI surveillance system was recently introduced, enabling the automated acquisition of face imagery of subjects at-a-distance and on-the-move (up to 50 meters away). This dataset was the basis of the “ICB-RW: International Challenge on Biometric Recognition-in-the-Wild” competition, of which the primary goal was fostering the development of biometric recognition algorithms capable of working in surveillance scenarios.

The ICB-RW competition took place from September to December, 2015. There were a total of 19 registrations in the competition, most of these from academic/research institutions, also with a small number coming from private companies. A learning set from the QUIS-CAMPI database was initially released for all participants and, by the end of the contest, a disjoint subset was used in performance evaluation. Based on the obtained results, seven methods were selected, and their authors were invited to contribute to this department.

Ekenel et al. align the probe and gallery face images with respect to eye centers, considering only frontal images as gallery elements. A convolutional neural network (CNN) is used for face representation purposes, with 1-nearest neighbor rule based on signal correlation being used for matching.

Grm and Struc generated an augmented version of the learning set by oversampling the training images via bounding box noise and horizontal flipping. The pre-trained Visual Geometry Group (VGG) face deep convolutional network was used as a feature extractor and a soft max classifier to discriminate between genuine and imposter pairwise comparisons.

Shi et al. used a feature set extracted from a deep convolutional network model trained on the CASIA-Webface database, and a similarity measure based on cosine distance. Ten models were learned independently from different facial parts, and subsequently fused. Also, multi-pose gallery data was synthesized to ease the matching phase.

Gutfeter and Pacut provided an information fusion approach that relies on the responses given by a set of convolutional neural networks that perform face recognition, each one specialized in handling samples from a specific 3D angle.

Brogan and Scheirer started by frontalizing both the gallery and probe data. Next, feature extraction was carried out based on a SLMSimple Neural Network with four bins created to represent different versions of the gallery samples. Finally, probe descriptors are matched with one of the four bins according to yaw angle of the head, and the resulting pairs of feature vectors feed a support vector machine that performs biometric recognition.

Gonzalez-Sosa et al. (Universidad Autónoma de Madrid, Spain) extracted Local Binary Patterns from nine facial regions of frontalized versions of the images. Next, illumination is compensated, and a fused distance score is determined by only considering the five best individual facial regions of each sample.

Finally, Riccio, Nappi, and de Maio started by locating a set of facial key points using an Active Shape Model. This step provided the information to remap (align) the face regions into 64 x 100 images of constant dimension. Next, local light adjustment techniques are used to compensate for the dynamic lighting conditions, with matching being carried out according to an optimized localized version of the spatial correlation index.

We hope that this collection of seven papers provides an overview of the current research in this extremely ambitious sub-field of biometric recognition research. We wish to thank all the people that enabled the publication of this special issue. First of all, we wish to thank Dr. Daniel Zeng, the editor-in-chief emeritus of this magazine, for accepting this idea with enthusiasm and for his support and motivation. Also, we would like to acknowledge the work carried out by João C. Neves, both in the management of the ICB-RW contest and in the performance evaluation of the submitted algorithms.

Acknowledgments

This work is supported by ‘’FCT – Fundação para a Ciência e Tecnologia” (Portugal), through the project “UID/EEA/50008/2013”.