Advances in Visual Computing: 6th International Symposium, ISVC 2010, Las Vegas, NV, (USA) November 29-December 1, 2010. Proceedings. Lecture Notes in Computer Science, Volumen 6453. Springer, 2010. 461-468

DOI: http://dx.doi.org/10.1007/978-3-642-17289-2_44

Acquisition Scenario Analysis for Face Recognition at a Distance

P. Tome[a], J. Fierrez[a], M.C. Fairhurst[b] and J. Ortega-Garcia[a].

[a]Biometric Recognition Group - ATVS
Escuela Politecnica Superior - Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco 28049 Madrid, Spain
{pedro.tome, julian.fierrez, javier.ortega}@uam.es

[b]Department of Electronics, University of Kent, Canterbury, Kent CT2 7NT, UK
{M.C.Fairhurst}@kent.ac.uk

Abstract. An experimental analysis of three acquisition scenarios for face recognition at a distance is reported, namely: close, medium, and far distance between camera and query face, the three of them considering templates enrolled in controlled conditions. These three representative scenarios are studied using data from the NIST Multiple Biometric Grand Challenge, as the first step in order to understand the main variability factors that affect face recognition at a distance based on realistic yet workable and widely available data. The scenario analysis is conducted quantitatively in two ways. First, we analyze the information content in segmented faces in the different scenarios. Second, we analyze the performance across scenarios of three matchers, one commercial, and two other standard approaches using popular features (PCA and DCT) and matchers (SVM and GMM). The results show to what extent the acquisition setup impacts on the verification performance of face recognition at a distance. [1]

Key words: Biometrics, face recognition, at a distance, on the move.

1 Introduction

Face and iris are two of the most relevant biometrics used nowadays in many user recognition applications [1,2]. A new research line growing in popularity is focused on using these biometrics in less constrained scenarios in a non-intrusive way, including acquisition "On the Move" and "At a Distance" [3]. Imagine a scenario where the people do not have to stop in front of a sensor to acquire a picture of the face: simply, they walk through an identification bow. This kind of scenarios are still in their infancy, and much research and development is

[1] P. T. is supported by a FPU Fellowship from Univ. Autonoma de Madrid. Part of this work was conducted during a research stay of P.T. at Univ. of Kent funded by the European Action COST2101. This work has been partially supported by projects Bio-Challenge (TEC2009-11186), Contexts (S2009/TIC-1485), TeraSense (CSD2008-00068) and "Cátedra UAM-Telefónica".

needed in order to achieve the levels of precision and performance that certain applications require.

The new field of biometrics at a distance is enabled mainly thanks to: 1) recent advances in sensing technology [2], and 2) new algorithms and methods to deal with varying factors (e.g., illumination, movement, pose, distance to the camera), which in this case are less controlled than the ideal situations commonly considered in biometrics research.

As a result of the interest in these biometric applications at a distance, there is now a growing number of research works studying how to compensate for the main degradations found in uncontrolled scenarios [4]. Nevertheless, there is almost no experimental knowledge about the main variability factors found in specific scenarios, which may help in devising robust methods for biometrics at a distance tailored to specific applications of practical importance. The contribution of the present paper is toward this end, by analyzing quantitatively three scenarios of face recognition at a distance, namely: close, medium and far distance between subject and camera. This analysis is conducted quantitatively at two levels for the considered scenarios: 1) main data statistics such as information content, and 2) performance of recognition systems: one commercial, and two other based on popular features (PCA and DCT) and matchers (SVM and GMM).

The scenarios under study are extracted from the NIST Multiple Biometric Grand Challenge [5], which is focused on biometric recognition at a distance using iris and face. In particular, we use a subset of this benchmark dataset consisting of images of a total of 112 subjects acquired at different distances and varying conditions regarding illumination, pose/angle of head, and facial expression.

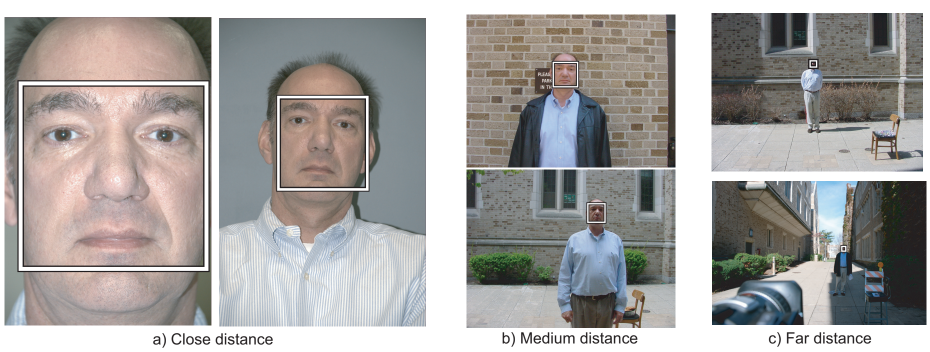

Fig. 1. Example images of the three scenarios: a) close distance, b) medium distance, and c) far distance.

The paper is structured as follows. Sect. 2 describes the dataset and scenarios under study. Sect. 3 analyzes the main data statistics of the scenarios. Sect. 4

studies the performance of the three considered recognition systems on the different scenarios. Sect. 5 finally discusses the experimental findings and outlines future research.

2 Scenario Definition

The three scenarios considered are: 1) "close" distance, in which the shoulders may be present; 2) "medium" distance, including the upper body; and 3) "far" distance, including the full body. Using this three general definitions we marked manually all the 3482 face images from the 147 subjects present in the dataset NIST MBGC v2.0 Face Stills [5]. Some examples images are depicted in Fig. 1. A portion of the dataset was discarded (360 images from 89 subjects), because the face was occluded or the illumination completely degraded the face. Furthermore, although this information is not used in the present paper, all the images were marked as indoor or outdoor.

Finally, in order to enable verification experiments considering enrollment at close distance and testing at close, medium, and far distance scenarios, we kept only the subjects with at least 2 images in close and at least 1 image in both of the two other scenarios. The data selection process is summarized in Table 1, were we can see that the three considered scenarios result in 112 subjects and 2964 face images.

Num. users	Close distance	Medium distance	Far distance	Discarded images	Total
147	1539	870	713	360	**3482**
	At least **2** *images per user*	*At least* **1** *images per user*			
112	1468	836	660		**2964**

Table 1. Number of images of each scenario constructed from NIST MBGC v2.0 Face Visible Stills.

3 Scenario Analysis: Data Statistics

3.1 Face Segmentation and Quality

We first segmented and localized the faces (square areas) in the three acquisition scenarios using the VeriLook SDK discussed in Sect. 4.1. Segmentation results are shown in Table 2, where the segmentation errors increase significantly across scenarios, from only 1.43% in close distance to 82.57% in far distance. Segmentation errors here mean that the VeriLook software could not find a face in the image due to the small size of faces and increment of variability factors. For all the faces detected by VeriLook, we conducted a visual check, where we observed 3 and 10 segmentation errors for medium and far distance, respectively.

	Close distance	Medium distance	Far distance	Discarded	Total
Num. Images	1468	836	660	360	**3324**
Errors	21	151	545		**848**
Errors(%)	1.43%	18.06%	**82.57%**		

Table 2. Segmentation results based on errors produced by face Extractor of VeriLook SDK.

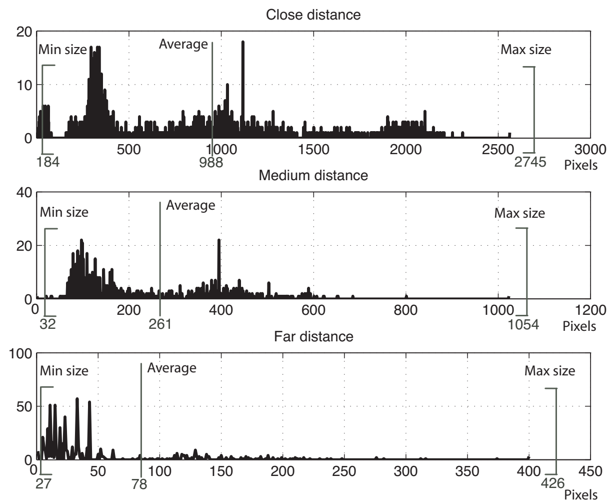

Fig. 2. Histograms of face sizes for each scenario (side of the square area in pixels).

All the segmentation errors were then manually corrected by manually marking the eyes. The face area was then estimated based on the marked distance between eyes.

The resulting sizes of the segmented faces are shown in Fig. 2, where we observe to what extent the face size decreases with the acquisition distance. In particular, the average face size in pixels for each scenario is: 988×988 for close, 261×261 for medium, and 78×78 for far distance.

Another data statistic we computed for the three scenarios is the average face quality index provided by VeriLook (0 = lowest, 100 = highest): 73.93 for close, 68.77 for medium, and 66.50 for far distance (see Fig. 3, computed only for the faces correctly segmented by VeriLook). As stated by VeriLook providers, this quality index considers factors such as lightning, pose, and expression.

3.2 Information Content

The entropy of the face images in the different acquisition scenarios represents a quantitative assessment of the information content in the gray levels of the images. In principle, an image acquired in controlled conditions (illumination, clean background, neutral pose, ...) would have less entropy than other image acquired at a distance in uncontrolled conditions. In Fig. 4 (top), this effect

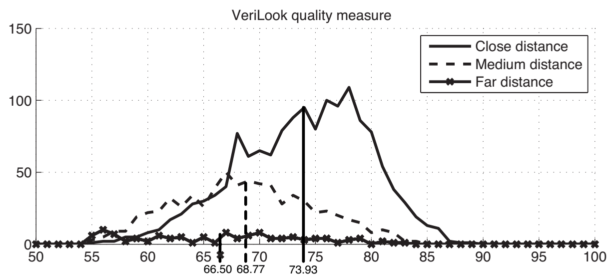

Fig. 3. Histogram of face quality measures produced by VeriLook SDK.

is patent: the farther the distance the higher the entropy. When considering only the information within the segmented faces, as shown in Fig. 4 (down), the opposite occurs: the farther the distance the lower the entropy. These two measures (increase in entropy of the full image, and decrease in entropy of the segmented faces), can therefore be seen, respectively, as a quantitative measure of the scenario complexity increase due to background effects, and the reduction in information within the region of interest due to acquisition scenario change.

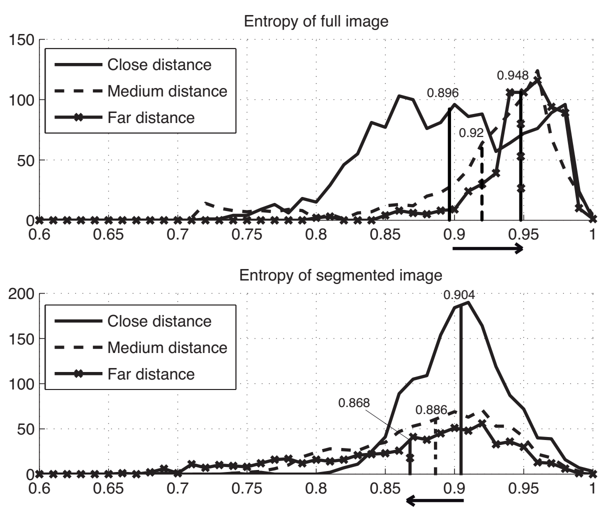

Fig. 4. Histograms of entropy for full images (top) and segmented faces (down) for the three scenarios with their corresponding average value.

4 Scenario Analysis: Verification Performance Evaluation

4.1 Face Verification Systems

- **VeriLook SDK**. This is the commercial face recognition system provided by Neurotechnology [2].
- **PCA-SVM system**. This verification system uses Principal Component Analysis (PCA). The evaluated system uses normalized and cropped face images of size 64 × 80 (width × height), to train a PCA vector space where 96% of the variance is retained. This leads to a system where the original image space of 5120 dimensions is reduced to 249 dimensions. Similarity scores are computed in this PCA vector space using a SVM classifier with linear kernel.
- **DCT-GMM system**. This verification system also uses face images of size 64 × 80 divided into 8 × 8 blocks with horizontal and vertical overlap of 4 pixels. This process results in 285 blocks per segmented face. From each block a feature vector is obtained by applying the Discrete Cosine Transform (DCT); from which only the first 15 coefficients ($N = 15$) are retained. The blocks are used to derive a world GMM Ω_w and a client GMM Ω_c [6]. From previous experiments we obtained that using ($M = 1024$) mixture components per GMM gave the best results. The DCT feature vector from each block is matched to both Ω_w and Ω_c to produce a log-likelihood score.

4.2 Experimental Protocol

Three main experiments are defined for the verification performance assessment across scenarios:

- *Close2close*. This will give us an idea about the performance of the systems in ideal conditions (both enrollment and testing using close distance images). About half of the close distance subcorpus (754 images) is used for development (training the PCA subspace, SVM, etc.), and the rest (714 images) is used for testing the performance.
- *Close2medium*, and *close2far* protocol. These two other protocols use as training set the whole close distance dataset (1468 face images). For testing the performance of the systems, we use the two other datasets: 836 medium distance images for *close2medium*, and 660 far distance images for *close2far*.

4.3 Results

In Fig. 5 we show the verification performance for the three considered scenarios: *close2close*, *close2medium*, and *close2far*. We first observe that VeriLook is the best of the three systems in *close2close* with an EER around 7%. At the same time, this commercial system is the most degraded in uncontrolled conditions, with an EER close to 40% in *close2far*, much worse than the other two

[2] http://www.neurotechnology.com/

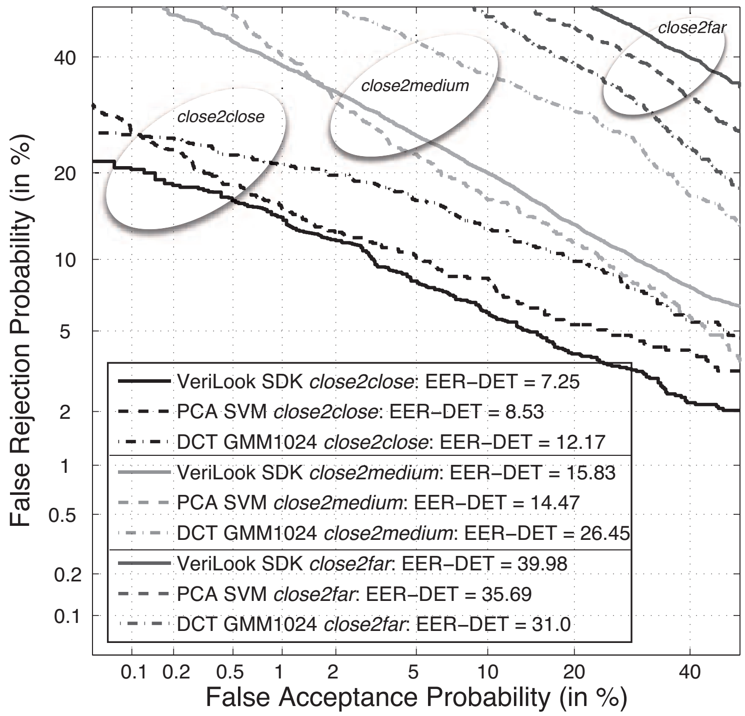

Fig. 5. Verification performance results for the three scenarios and three systems considered.

much simpler systems. This result corroborates the importance of analyzing and properly dealing with variability factors arising in biometrics at a distance.

We also observe in Fig. 5 that the GMM-based system works better in far distance conditions than the other systems, although being the less accurate in *close2close* and *close2medium*. This result demonstrates the greater generalization power of this simple recognition approach, and its robustness against uncontrolled acquisition conditions.

Based on this observation, we finally conducted a last experiment simplifying the DCT-GMM complexity in order to enhance its generalization power, seeking for a maximum of performance in the challenging *close2far* scenario. The verification performance results are given in Table 3 as EER for decreasing DCT-GMM complexity ($N =$ DCT coefficients, $M =$ Gaussian components per GMM). The results indicate in this case that decreasing the recognition complexity (i.e., improving the generalization power) of this simple recognition method does not help in improving its robustness against uncontrolled conditions. In other words, the DCT-GMM recognition complexity initially considered ($N = 15, M = 1024$), is the most adequate for the *close2far* scenario studied here.

EER	M Gaussians								
N Coeff.	*close2close*			*close2medium*			*close2far*		
DCT	1024	128	8	1024	128	8	1024	128	8
15	12.17	14.62	20.06	26.45	29.06	36.19	31.01	32.52	38.74
10	13.22	15.97	19.62	26.09	28.72	34.90	29.80	32.83	38.58
5	17.66	19.80	22.15	31.72	34.60	35.43	33.46	37.07	39.37

Table 3. Verification performance of the DCT-GMM system for different configurations.

5 Discussion and Future Work

An experimental approach towards understanding the variability factors in face recognition at a distance has been reported. In particular, we have conducted a data-driven analysis of three realistic acquisition scenarios at different distances (close, medium, and far), as a first step towards devising adequate recognition methods capable to work in less constrained scenarios. This data-driven analysis has been made for a subset of the benchmark dataset NIST MBGC v2.0 Face Stills.

Our analysis has been focused on: 1) data statistics (segmented face sizes, quality and entropy measures), and 2) verification performance of three systems. The results showed that the considered systems degrade significantly in the far distance scenario, being more robust to uncontrolled conditions the most simple approach.

Noteworthy, the scenarios considered in the present paper differ not only in the distance factor, but also in illumination and pose (being the illumination variability much higher in far distance than in close distance). Based on the data statistics obtained and the performance evaluation results, a study of the effects of such individual factors is source for future research.

References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Comput. Surv. **35** (2003) 399–458
2. Matey, J., Naroditsky, O., Hanna, K., Kolczynski, R., LoIacono, D., Mangru, S., Tinker, M., Zappia, T., Zhao, W.: Iris on the move: Acquisition of images for iris recognition in less constrained environments. Proc. of the IEEE **94** (2006) 1936–1947
3. Li, S.Z., Schouten, B., Tistarelli, M.: Biometrics at a Distance: Issues, Challenges, and Prospects. In: Handbook of Remote Biometrics for Surveillance and Security. Springer (2009) 3–21
4. Robust2008: Robust biometrics: Understanding science and technology, (http://biometrics.cylab.cmu.edu/ROBUST2008)
5. MBGC: Multiple biometric grand challenge. (NIST - National Institute of Standard and Technology)
6. Galbally, J., McCool, C., Fierrez, J., Marcel, S., Ortega-Garcia, J.: On the vulnerability of face verification systems to hill-climbing attacks. Pattern Recognition **43** (2010) 1027–1038