Face Sample Quality

Kui Jia[1] and Shaogang Gong[2]

[1] Shenzhen Institute of Advanced Integration Technology, CAS / CUHK, Shenzhen, China
kui.jia@sub.siat.ac.cn
[2] Queen Mary, University of London, Mile End Road, London, UK
sgg@dcs.qmul.ac.uk

Synonyms

Face sample standardization; Face sample utility

Definition

Face is a human biometric attribute that can be used to establish the identity of a person. A face-based biometric system operates by capturing probe face samples and comparing them against gallery face templates. The intrinsic characteristics of captured face samples determine their effectiveness for face authentication. Face sample quality is a measurement of these intrinsic characteristics. Face sample quality has significant impact on the performance of a face-based biometric system. Recognizing face samples of poor quality is still a challenging problem. A number of factors can contribute towards degradation in face sample quality. They include, but not limited to, illumination variation, pose variation, facial expression change, face occlusion, low resolution and high sensing noise.

Main Body Text

Introduction

A typical face-based biometric system operates by capturing face data (images or videos), and comparing the obtained face data against face templates of different individuals in a gallery set. Whilst face templates in the gallery set are normally captured under constrained imaging conditions (e.g. from frontal view, at a short distance from the camera, and under consistent illumination), it is unrealistic to assume controlled acquisition of probe face data. Face data captured under uncontrollable environment usually contain many kinds of defects caused by poor illumination, improper face positioning and imperfect camera sensors [2]. For instance, when face data is captured in a natural outdoor environment, inconsistent illumination is typically casted on human faces resulting in uneven, extremely strong or weak lightings. Face rotation can also cause significant appearance variations, and at the extreme, face can be self occluded (Fig. 1). When distances between human faces and cameras increase, captured face data will be at low resolution, in low contrast, and likely to contain high imaging noise. In some instances people may wear sunglasses, have varying facial expression, and be with heavy makeup. All of these factors contribute towards potential degradation in the quality of captured face samples, resulting in disparities to those of face templates stored in the gallery set.

Face sample quality has significant impact on the performance of face-based biometric systems. Assessing the quality of face samples before applying them in any biometric system may help improve the authentication accuracy. For example, an intruder may wear sunglasses intending to disguise himself, quality assessment of intruder's face samples can give an alert to such a situation. Quantitative measures on the quality of face samples can also be integrated into biometric systems to increase or decrease relevant thresholds. In a people enrollment stage, such quantitative measures of quality also help procure gallery face templates of good quality. Many approaches assess face sample quality using general image properties including contrast, sharpness and illumination intensity [3]. However, these properties cannot properly measure face sample degradation caused by inconsistent illumination, face rotation, or large face-camera distance. There are a few recent works assessing face sample quality by considering such kinds of degradation. For example in [2], facial symmetry based methods are used to measure facial asymmetries caused by non-frontal lighting and improper facial pose.

When only poor quality face data can be acquired at the authentication stage, face recognition becomes significantly more challenging due to: (i) *Illumination variation* to which the performance of most existing face recognition algorithms and systems is highly sensitive. It has been shown both experimentally [4] and theoretically [5] that face image differences resulting from illumination variation are more significant than either inherent face differences between different individuals, or those from varying face poses [6]. State-of-the-art approaches addressing this problem include heuristic methods, reflectance-model methods, and 3D-model-based methods [1]. Although performance improvement is achieved, none of these methods are truly illumination invariant. (ii) *Pose variation* which causes face recognition accuracy to decrease significantly, especially when large pose variations between gallery and probe faces are present. The difficulties would further increase if only an unknown single pose is available for each probe face. In such a situation, an extra independent training set, different from the galley set and containing multiple face images of different individuals under varying poses, will be helpful. Three-dimensional face model or statistical relational learning between different poses can be employed to generate virtual face poses. By generating virtual poses, one can either normalize probe faces of varying poses to a predefined pose, e.g. frontal, or expand the gallery to cover large pose variations. (iii) *Low resolution* face data will be acquired when face-camera distances increase, which is rather typical in surveillance imagery. The performance of existing face recognition systems decreases significantly when the resolution of captured face data is reduced below a certain level. This is because the missing high-resolution details in facial appearances and image features make facial analysis and recognition ineffective, either by human operators or by automated systems. It is therefore useful to generate high-resolution face images from low-resolution ones. This technique is known as face hallucination [7] or face **super-resolution**.

Assessment of Face Sample Quality

The performance of face authentication depends heavily on face sample quality. Thus the significance of face sample quality assessment and standardization grows as more practical face-based biometric systems are required. Quality assessment of probe face samples can either reject or accept a probe in order to improve later face verification or identification accuracy. Quantitative assessment of face sample quality can also be used to assign weights in a biometric fusion scheme.

ISO/IEC WD 29794-1 [8] considers that biometric sample quality can be defined by character (inherent features), fidelity (accuracy of features), or utility (predicted biometrics performance). Many efforts have been made on biometric sample quality assessment for fingerprint, iris or face data. Most of those on face data are based on general image properties including contrast, sharpness and illumination intensity [3]. However, the face sample degradation that severely affects face authentication accuracy is from uncontrollable imaging conditions that cause variations in illumination and head pose, and/or very low resolution in facial appearance. There are a few attempts made on assessing face sample quality caused by these kinds of degradation.

In [9], two different strategies for face sample quality assessment are considered: one is for illumination variation and pose change, another is for facial expression change. In the first strategy, specific measures are defined to correlate with levels of different types of face sample degradation. A polynomial function is then utilised based on each measure for predicting the performance of a Eigenface technique on a given face sample. Quality goodness is assessed by selecting a suitable threshold. Since the measurement of facial expression intensity is difficult, in the second strategy, a given face sample is classified into good or poor quality based on its coarse similarity to neutral facial expression. Then the training procedure for each class is achieved by dividing the training set into two subsets, based on whether the samples are recognizable by the Eigenface technique. Then these two subsets are described by Gaussian mixture models (GMMs). In [2], facial symmetry based quality scores are used to assess facial asymmetries caused by non-frontal lighting and improper facial pose. In particular, local binary pattern (LBP) histogram features are applied to measure the lighting and pose asymmetries. Moreover, the inter-eye distance is also used to estimate the quality score for whether a face is at a proper distance from the camera.

Recognizing Face Samples of Poor Quality

In general, face recognition under varying illumination is difficult. Although existing efforts to address this challenge have not led to a fully satisfactory solution for illumination invariant face recognition, some performance improvements have been achieved. They can be broadly categorized into: heuristic methods, reflectance-model methods, and 3D-model-based methods [1]. A typical heuristic method applies subspace learning, e.g. principal component analysis (PCA), using training face samples. By discarding a few most significant, e.g. the first three, principal components, variations due to lighting can be reduced. Reflectance-model methods employ a Lambertian reflectance model with a varying albedo field, under the assumption of no attached and cast shadows. The main disadvantage of this approach is the lack of generalization from known objects to unknown objects [10]. For 3D face model based approaches, more stringent assumptions are often made and it is also computationally less reliable. For example in [11], it is assumed that the 3D face geometry lies in a linear space spanned by the 3D geometry of training faces and it uses a constant albedo field. Moreover, 3D model-based methods require complex fitting algorithms and high-resolution face images.

There are also attempts to address the problem of face recognition across varying facial poses. In real-world applications, one may have multiple face samples of varying poses in training and gallery sets (since they can be acquired offline), whilst each captured probe face can only be at an unknown single pose. Three-dimensional model-based methods [12] or statistical learning-based methods can be used to generate virtual face poses [13], by which either probe faces can be normalized to a predefined pose, e.g. frontal view, or gallery faces can be expanded to cover large pose variations. For example in [12], a 3D morphable model is used. The specific 3D face is recovered by simultaneously optimizing the shape, texture and mapping parameters through an analysis-by-synthesis strategy. The disadvantage of 3D model-based methods is slow speed for real-world applications. Learning-based methods try to learn the relations between different facial poses and how to estimate a virtual pose in 2D domain, e.g. the view-based active appearance model (AAM) [14]. This method depends heavily on the accuracy of face alignment, which unfortunately introduces another open problem in practice.

When the resolution of captured face data falls below a certain level, existing face recognition systems will be significantly affected. Face super-resolution techniques have been proposed to address this challenge. Reconstruction-based approaches require multiple, accurately aligned low-resolution face samples to obtain a high-resolution face image. Their magnification factors of image resolution are however limited [7]. Alternatively, learning-based face super-resolution approaches model high-resolution training faces and learn face-specific prior knowledge from them. They use the learned model prior to constrain the super-resolution process. A super-resolution factor as high as 4×4 can be achieved [7]. The face super-resolution process can also be integrated with face recognition. For example in [16], face image super-resolution is transferred from pixel domain to a lower dimensional eigenface space. Then the obtained high-resolution face features can be directly used in face recognition. Simultaneous face super-resolution and recognition in tensor space have also been considered [15]. Given one low-resolution face input of single modality, the proposed method can integrate and realize the tasks of face super-resolution and recognition across different facial modalities including varying facial expression, pose or illumination.

Summary

Many face-based biometric systems have been deployed in applications ranging from national border control to building door access, which normally solve the sample quality problem at the initial face acquisition stage. Given ongoing progress on standardization of face sample quality and technical advancement in authenticating face samples of poor quality, the availability of more reliable and convenient face authentication systems is only a matter of time.

Related Entries

Biometric Sample Quality, Face Illumination Analysis, Face Pose Analysis, Face Recognition

References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face Recognition: A Literature Survey. ACM Computing Surveys **35** (4) (2003) 399–458
2. Gao, X.F., Li, S.Z., Liu, R., Zhang, P.R.: Standardization of Face Image Sample Quality. In: Proceedings of Second International Conference on Biometrics (ICB), Seoul, Korea (2007) 242–251
3. Brauckmann, M., Werner, M.: Technical Report. In: Proceedings of NIST Biometric Quality Workshop (2006)
4. Adini, Y., Moses, Y., Ullman, S.: Face Recognition: The Problem of Compensating for Changes in Illumination Direction. IEEE Transactions on Pattern Analysis and Machine Intelligence **19** (7) (1997) 721–732
5. Zhao, W., Chellappa, R.: Robust Face Recognition Using Symmetric Shape-from-Shading. Technical Report, Center for Automation Research, University of Maryland (1999)
6. Tarr, M.J., Bulthoff, H.H.: Image-Based Object Recognition in Man, Monkey and Machine. Cognition **67** (1998) 1–20

7. Baker, S., Kanade, T.: Limits on Super-Resolution and How to Break Them. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (9) (2002) 1167–1183
8. ISO/IEC JTC 1/SC 37 N 1477: Biometric Sample Quality Standard - Part 1: Framework (2006)
9. Abdel-Mottaleb, M., Mahoor, M.H.: Application Notes - Algorithms for Assessing the Quality of Facial Images. IEEE Computational Intelligence Magazine **2** (2) (2007) 10–17
10. Baker, S., Kanade, T.: Appearance Characterization of Linear Lambertian Objects, Generalized Photometric Stereo, and Illumination-Invariant Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **29** (2) (2007) 230–245
11. Atick, J., Griffin, P., Redlich, A.: Statistical Approach to Shape from Shading: Reconstrunction of 3-Dimensional Face Surfaces from Single 2-Dimentional Images. Neural Computation **8** (2) (1996) 1321–1340
12. Blanz, V., Vetter, T.: Face Recognition based on Fitting a 3-D Morphable Model. IEEE Transactions on Pattern Analysis and Machine Intelligence **25** (9) (2003) 1063–1074
13. Li, Y., Gong, S., Liddell, H.: Constructing Facial Identity Surfaces for Recognition. International Journal of Computer Vision **53** (1) (2003) 71–92
Cootes, T.F., Walker, K., Taylor, C.J.: View-based Active Appearance Models. In: Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition (FG), Grenoble, France (2000) 227–232
14. Cootes, T.F., Walker, K., Taylor, C.J.: View-based Active Appearance Models. In: Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition (FG), Grenoble, France (2000) 227–232
15. Jia, K., Gong, S.G.: Multi-Modal Tensor Face for Simultaneous Super-Resolution and Recognition. In: Proceedings of Tenth International Conference on Computer Vision (ICCV), Beijing, China **2** (2005) 1683–1690
16. Gunturk, B.K., Batur, A.U., Altunbasak, Y., Hayes, M.H., Mersereau, R.M.: Eigenface-Domain Super-Resolution for Face Recognition. IEEE Transactions on Image Processing **12** (5) (2003) 597–606

Definitional Entries

Super-Resolution

The techniques that form an enhanced-resolution image by fusing together multiple low-resolution and/or learning from high-resolution training images are known as super-resolution. Super-resolution can be performed in either frequency or spatial domain.

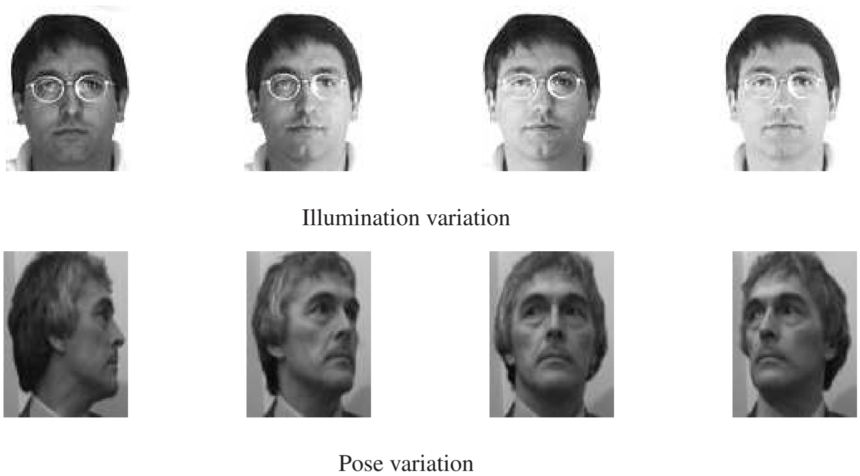

Fig. 1. Face samples of illumination and pose variations from AR and UMIST databases.