International Journal of Innovative Research in Electronics and Communications (IJIREC)
Volume 2, Issue 1, January 2015, PP 18-23
ISSN 2349-4042 (Print) & ISSN 2349-4050 (Online)
www.arcjournals.org

Face Recognition Using Active Shape Model

Sake Kiran Kumar[1], S. Raghavendra Swami[2], S. Ravi Kumar[3]

[1]PG Scholar, Electronics and Communication Engineering, PVKK Institute of Technology AP, India
[2]Asst Professor, Electronics and Communication Engineering, PVKK Institute of Technology, AP, India
[3]Asst Professor, Electronics and Communication Engineering, PVKK Institute of Technology, AP, India
sakekirankumar6@gmail.com

Abstract: *Face images captured by surveillance cameras usually have poor resolution in addition to uncontrolled poses and illumination conditions which adversely affect performance of face matching algorithms. In this paper, we develop a novel approach for matching surveillance quality facial images to high resolution images in frontal pose which are often available during enrollment. The proposed approach uses Multidimensional Scaling to simultaneously transform the features from the poor quality probe images and the high quality gallery images in such a manner that the distances between them approximate the distances had the probe images been captured in the same conditions as the gallery images. Thorough evaluation on the Multi-PIE dataset and comparisons with state-of-the-art super-resolution and classifier based approaches are performed to illustrate the usefulness of the proposed approach. Experiments on real surveillance images further signify the applicability of the framework.*

Keywords: *Video based image capturing, Face detection, pose and illumination changes, security authentication, Tagging*

1. Introduction

The face recognition in unconstrained scenarios is still a major challenge. Face recognition has been an active research area and many types of algorithms and techniques have been proposed. Some of the most popular and well-established algorithms for face recognition are principal component analysis (PCA), LDA, ICA, and SVM to assess the feasibility of real world face recognition in uncontrolled setting. In typical surveillance scenarios, people are usually walking free, and they are impossible to always keep their faces frontal or looking to the cameras. Face images captured by surveillance systems are non-ideal, because they are often affected by many factors: pose illumination, expression, occlusion, distance, weather and so on. 3D methods are always based on a 3D face model, which may be a single model, or a deformable model in certain parametric forms. In typical face recognition applications, the enrolled face images (gallery) are usually captured under controlled environment. [1] 3D methods can be divided into four categories depending on how to use of the 3D model:

1. Pose Normalization: Face images in the probe are normalized to frontal view based on the 3D model, and then match the normalized probe to the gallery .

2. Pose Synthesis: Use the 3D model to generate some virtual face images with various poses for the face images in the gallery, and then match the probe to the virtual face images.

3. Recognition by Fitting: Fit all face images in the gallery and probe by the 3D model. The texture and shape parameters are used for face recognition.

4. Filter Transformation: Transform the filters according to the pose and shape of face image, and then use the pose adapted filters for feature extraction.

When the same face will appears differently due to the variation in lighting and poses changes, the illumination problem will occur and there exists head rotation. The uneven lightning brings variations in illumination which affects the classification greatly since the facial features that are being used for classification gets effected due to this variation. The Local Binary Pattern (LBP) is an invariant feature extraction type [5] algorithm. It was first proposed for the use of texture description and it has been used to normalize illumination in face detection and recognition contexts. In LBP the center pixel cannot be compared with itself. So LBP cannot capture the local structure of the image. These images should be reduced in the size, to perform this operation, the convex cone. This convex cone is also known as illumination cone. These images are formed from the set of images with the same posture. This illumination cone can be well approximated by a low-dimensional linear subspace. Under variable lighting, the set of various images are characterized by a illumination cones parameterized by the pose.

2. Face Detection, Sample Quality, and Normalization

This section starts with face and landmark detection. It then describes the module responsible to derive distortion indices SP and SI, which estimate the distortions of some face image in pose and illumination. Normalization ("correction") methods suitable for such changes close the section.

A. Face Detection

The first step in a face-recognition-based application is the detection of one or more faces within an image. In our FACE, the face and its characteristic points ("landmarks") are located through the approach presented in [33], namely, the extended Active Shape Model (STASM) algorithm. The latter is used to locate features in frontal views of upright faces. It first submits the image to a global face detector (Viola-Jones or Rowley [38]), which extracts all regions of interest (ROI) from an image that includes at least one face. Images that lack faces are discarded. ROI are individually fed to STASM algorithm, which searches for relevant landmarks by minimizing a global distance between candidate image points and their homologues using a general model (shape model), which is precomputed ("learned") over a wide set of training images. The algorithm locates 68 interest points. The precision of the location procedure depends on the amount of face distortion. In the section presenting experimental results, we will discuss how the precision in the detection of relevant points influences the accuracy of recognition. These points are exploited for normalization, i.e., to render the face image using a canonical pose and illumination representation. Let us note that STASM may be not a solution fits all. However, it is among the few ones that are accessible as open source and can be handled without too much effort. This is a critical point, since solving such problems is out of the scope for this paper. The quality indices proposed and implemented, such as the one based on symmetry, which work even with an inaccurate point location, aim at handling such situations by asking for a new capture/selection of a different sample ("identity management") or for user validation in the specific case of tagging (when a new sample may not be available or cannot be recaptured). To the extent that STASM helps its use seems justified. In addition, we tested two further implementations of ASM, namely, ASM Library [48] and Open ASM [58]. The obtained results were in favor of using STASM.We also trained it on sample images from the image sets used. It is further worth noticing that, if one obtains good results with an imprecise location technique, the whole system, which is our challenge, can only get better with a better location module

B. Measuring Pose and Illumination Distortions

A way to measure the quality of a FACE probe is to consider the amount of "effort" that would be needed to correct for pose and illumination distortions in the image, with larger corrections yielding lower quality ("distortion") indices. The normalization procedure aims at recovering a frontal pose of the face presented in the input image, starting from the points located using the STASM approach in [33], as sketched previously. The distribution of such points on the face is a good starting point to evaluate the degree of distortion that needs to be corrected by the pose normalization process. The index for pose quality is given by the linear combination of three parts, which are, respectively, inversely proportional to roll, yaw, and pitch, i.e., the considered distortion components. In general, these three components vary jointly. The pose variations are such that a good distortion index, however, can be derived using independent weights for each component.We notice here that an accurate measure of the rotation angles to quantify such distortions is beyond our goals.We are rather interested in an estimation of the influence of pose distortion on recognition. Towards that end, the intuitive approximations used are good enough for our purpose and easy to derive. The same parameters also drive subsequent processing, e.g., face normalization.

C) Pose and Illumination Normalization

Pose normalization is often employed to improve classification accuracy. As an example, Blanz and Vetter [7] method, which is among the most efficient algorithms available for such purposes, computes a 3-D face model starting from one or more 2-D images and modifying a 3-D generic model (morph able model). A similar idea is applied in Polar Rose (see [53]). In principle, it allows us to simultaneously correct for both pose and illumination changes. This comes, however, at a significant computational cost, particularly when processing a high number of faces (regardless of their distribution within images). FACE exploits a less complex yet equally effective approach to pose normalization. Although starting from the set of 68 relevant points that are located by STASM, it uses only 13 of them, as shown in Fig. 4 (the red/darker ones). The center of the eyes is used to correct head rolling [Fig. 4(a)]. The distances between the external corners of the left and right eyes and the tip of the nose, represented respectively by *dl* and *dr* [Fig. 4(b)], allow us to locate the better exposed half of the face. If this is the right half ($dr \geq dl$), the image is left unchanged; otherwise, it is reflected with respect to the vertical axis (horizontal flip). This allows us to always consider the right half of the face in the following processing steps. The points in Fig. 4(c) provide the line that marks the boundary between the right and left half regions of the face. Actually, the first point upward (the light one) is not provided by STASM but is derived as the median point of the link between its immediate neighbors (the first two points upward on nose sides). The points making up the central line that ideally divides the face and the face ROI borders delimit the right and left face regions. The following step only considers the right (best exposed) half of the face. A stretching operation is applied to all of the rows in order to make them of constant length. The processed half face is then divided into horizontal and vertical bands, which are delimited by the lines passing through some of the interest points as shown in Fig. 4(d). The choice of the points to use is such that it allows the delimitation of bands that are most significant to drive the normalization process, namely, those containing the eyebrow, the eye, the nose, and the mouth. The lines are resized, to make the chosen interest points fall in predetermined positions.

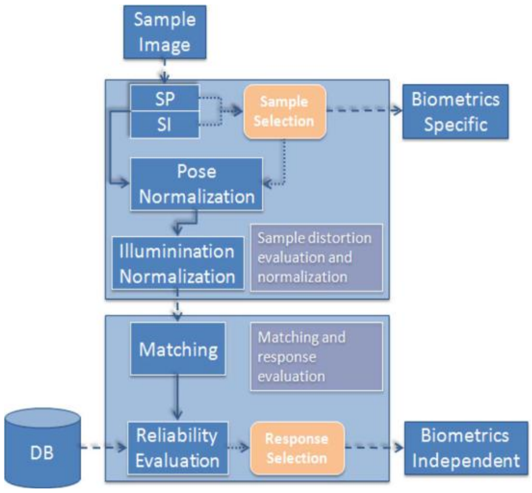

Fig.1. *FACE architecture*

3. IMAGE PREPROCESSING

Noise Reduction

A noise reduction filter is used in the binary image, which is used for eliminating single black pixel on white background. Eight neighbors of chosen pixels are examined if the number of black pixels are greater than white pixels then it is considered as black or else white. If there is any dirt on camera, scanner lens, imperfection in the scanner lighting etc, this shows the noise in the scanned face image.

The noise in the image can be removed with the help of filtering function, and works like a common function that replaces each pixel by its function.

Image Segmentation

The edges from shading have a large effect on the quality of an initial restored image. The edges may come from facial features include eyes, eyebrows, nose and mouth. Therefore, it is necessary to segment these facial feature regions and shading regions in order to reserve only the edges from shading for further processing. As the reconstructed images using PCA have a smaller number of noise points, these images, rather than the initial restored images [2], are used in segmentation for extracting edges. First, we can segment a reconstructed image to obtain the regions of the eyes, eyebrows, nose and mouth using an optimal threshold segmentation method. The presence of noise points should be removed. The uncovered small regions must also to be merged.

To perform these to operations, the erosion and dilation operators are used. In this the raw image is captured under an unknown light source to obtain the dark regions. The dark regions further divided into the darker region by means of the same segmentation method.

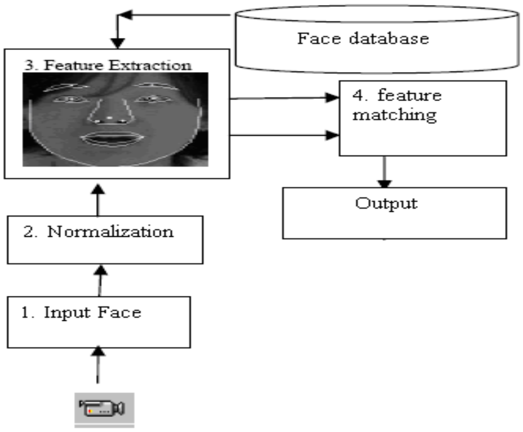

Fig.2. *Basic face recognition system*

Gaussian Filter

The restored image has a truer appearance if the size of the Gaussian filter is larger. However, the edges from shading and the noise in shading will also be magnified more and will appear in the restored image when a large filter size is used. Therefore, the uneven illumination variation is related to the gray-level intensities in shading and the edge strength from the shading in a raw image. These two factors are therefore used as the criteria in the determination of the maximum filter size in order to reduce the uneven illumination variation.

4. FACE MATCHING AND RESPONSE EVALUATION

Given quality indices SP and SI, for pose and illumination distortions, respectively, and normalization process that corrects estimated pose and illumination changes, we proceed to the second module of FACE, dealing with identification and decision making. We describe the matching algorithm and the reliability (“confidence”) for the matching scores. The latter relies on the gallery composition, namely, the relations holding among all of the measured distances from the probe image.

A. Matching Algorithm

Becker and Ortiz [4] benchmark studies, carried out using typical images from Facebook, report that many of the well-known techniques for face classification [1] are still too sensitive to image distortions. This prevents their use in commercial applications, which are typically run using uncontrolled settings. We propose here to perform image matching by a localized version of the spatial correlation index. The motivation for such an approach comes from the obvious observation that pose and illumination changes affect different parts of the face in different fashions. The changes experienced are thus local in nature and require individualized processing.

B. SRR

Once matching scores and rankings become available, we proceed with assessing the reliability of the results obtained. This is fed to the decision-making stage in order to filter out (“reject”) unreliable authentications and to accept (“recognize”) only those deemed sufficiently reliable. As discussed earlier, classical performance measures, such as RR, and figures of merit (FOM) like ROC, DET, or Tippet plots, whose aim is to express the discriminating power of a classifier, are suitable to compare the overall ability of different systems to perform correct recognition. Indices such as precision and recall, similar to the aforementioned FOMs, require access to ground truth and can only be evaluated *a posteriori*. All of the FOMs mentioned require access to the whole query set Q and the derivation of a full fledge response similarity matrix $S(G,Q)$. Response reliability (“confidence”) refers, however, to individual probes, it has no access to the whole query set Q, and its derivation needs to be immediate and confined to a single probe. This is because even a (generally) very well performing biometric engine may be occasionally deceived by a particularly “hard” probe. Therefore, matching score response reliability varies with each sample.

C. Some Summary Considerations

The biometric system, which integrates both quality and reliability indices, is not only robust but also flexible and fully interoperable even in a multimodal setting, since it provides a reliability measure for the classification accuracy and identity management. When it is possible to request a new sample, this is done; otherwise, such as it is the case for tagging, the system follows a different protocol, where answers are ordered according to their reliability and user feedback is sought. In this way, answer correction, which is typical of many applications of this kind, requires a lower workload for the user and increases overall system effectiveness. This is so because, as reliability increases, the probability of a wrong answer decreases. Notice that this holds for both quality and reliability indices. Moreover,

one can actually think of situations where tagging can access more than one collection for disambiguation purposes or expansion. The context relevance also affects re identification. Face exemplars stream online, and quality and reliability determine the following: 1) their label including confidence and 2) updates for their identity gallery set (not singletons). As a further consideration, database creation should be an ongoing process where samples can be dropped, updated, and or added, using suitable template updating strategies . Eventually, each identity ID should consist of a list of samples and their regions/parts to enable flexible matching. Finally, please note very recent results in [6] that detail evidence to the effect that good quality images do not necessarily match better than pairs of relatively lower quality. There is never going to be “perfect” data, and the challenge is that of interoperability subject to data variability.

5. CONCLUSION

This paper has described FACE, a new framework for face analysis including classification. FACE improves accuracy performance compared to state-of-the-art methods, for uncontrolled settings when the image acquisition conditions are not optimal. This is typical of applications such as photo tagging over social networks like Facebook or cataloguing of celebrities’ images in a magazine editorial office. FACE has access to multiple gallery instances for each subject and does not require expensive training to learn the face space, using instead straightforward correlation of local regions after proper pose and illumination normalization. FACE also has access to pose (SP) and illumination (SI) image quality indices, respectively, which can be used to *a priori* discard images whose quality is not sufficient to guarantee an accurate recognition response. Confidence in the system response is further assessed using SRR I and SRR II, two reliability indices based on the analysis of system responses in relation to the composition of the gallery. Experimental results show that FACE outperforms competing methods, with a significant increment in accuracy versus the next ranked methods. The improvement depends on the complexity of the data set at hand but is always worth of consideration. A number of research issues are still open. One of them regards scaling, i.e., the efficient use of our reliability measures with galleries consisting of millions of images, where such procedures can be very expensive. One possible solution can involve, for example, the precomputation of some pairwise metrics for the gallery, which might also allow using SRR measures for verification too. Another alternative is to perform preliminary clustering/binning of the gallery templates in order to reduce the search space without reducing accuracy below an acceptable level.

REFERENCES

[1]. A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recog. Lett.*, vol. 28, no. 14, pp. 1885– 1906, Oct. 2007.

[2]. T. Ahonen, A. Hadid, and M. Pietikainen, "Face recognition with local binary patterns," in *Proc. 8th ECCV*, N. Sebe, M. S. Lew, and T. S. Huang, Eds., 2004, pp. 469–481.

[3]. T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[4]. B. Becker and E. Ortiz, "Evaluation of face recognition techniques for application to Facebook," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recog.*, 2008, pp. 1–6.

[5]. P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar, "Localizing parts of faces using a consensus of exemplars," in *Proc. IEEE Conf. CVPR*, 2011, pp. 545–552.

[6]. J. R. Beveridge, P. J. Phillips, G. H. Givens, B. A. Draper, M. N. Teli, and D. S. Bolme, "When high-quality face images match poorly," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recog. Workshops*, 2011, pp. 572–578.

[7]. V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morph able model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1063– 1074, Sep. 2003.

[8]. D. S. Bolme, J. R. Beveridge, M. Teixeira, and B. A. Draper, "The CSU face identification evaluation system: Its purpose, features, and structure," in *Proc. ICVS*, 2003, pp. 304–311.

[9]. X. Chai, S. Shan, and W. Gao, "Pose normalization for robust face recognition based on statistical affine transformation," in *Proc. Inf., Commun. Signal Process./4th Pacific Rim Conf. Multimedia*, 2003, vol. 3, pp. 1413–1417.

[10]. M. De Marsico, M. Nappi, and D. Riccio, "FARO: Face recognition against occlusions and expression variations," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 40, no. 1, pp. 121–132, Jan. 2010.