13

Robust Face Alignment for Illumination and Pose Invariant Face Recognition

Fatih Kahraman[1], Binnur Kurt[2], Muhittin Gökmen[2]
Istanbul Technical University, [1]Informatics Institute, [2]Computer Engineering Department
Turkey

1. Introduction

Face recognition systems have matured from the systems working only in highly controlled indoor environments to the systems capable of identifying individuals in indoor or outdoor environments under severe conditions though some problems still remain, constraining their success to a limited degree. Illumination and pose variations are mostly responsible for dramatic changes on the face appearance. They produce such complex effects on the appearance of the acquired face that the face image pertains little to the actual identity. So any improvement in face appearance will enhance the recognition performance. Face recognition systems are usually required to handle highly varying illumination and pose conditions and more advanced techniques are needed to eliminate the undesired effects of variations from any sources. Research on face recognition has focused on solving issues arising from illumination and pose variations in one or more shots.

Lighting direction changes alter the relative gray scale distribution of face image and those changes due to lighting are larger than the one due to different personal identities. Consequently, illumination normalization is required to reach acceptable recognition rates. Varying illumination is a difficult problem and has received much attention in recent years. Several recent studies are centered around this issue: symmetric shape from shading (Zhao & Chellappa, 2000), for the illumination cones method (Georghiades & Belhumeur, 2001) theoretically explained the property of face image variations due to light direction changes. In this algorithm, both self shadow and cast-shadow were considered and its experimental results outperformed most existing methods. The main drawbacks of the illumination cone model are the computational cost and the strict requirement of seven input images per person.

Other directions of the photometric stereo in face recognition include introducing a more general illumination model, (Ramamoorthi, 2002) proposed a spherical harmonic representation for face images under various lighting conditions. (Basri & Jacobs, 2001) represent lighting using a spherical harmonic basis wherein a low-dimensional linear subspace is shown to be quite effective for recognition. The harmonic images can easily be computed analytically given surface normals and the albedos. (Zhou et al., 2007) extended a photometric stereo approach to unknown light sources. (Lee et al., 2005) empirically found a set of universal illumination directions, images under which can be directly used as a basis for the 9 dimensional illumination subspaces.

(Shashua & Riklin-Raviv, 2001) employ a very simple and practical image ratio method to map the face images into different lighting conditions. This method is suitable for modelling the variation in facial appearance caused by diffuse reflection and the proposed method is simply the ratio of albedo between a face image and linear combination of basis images, for each pixel. (Wang et al. , 2004) developed a reflectance estimation methods by using the idea of the ratio of the original image and its smooth version. In the same research direction (Zhang et al., 2007) and (An et al., 2008) proposed new methods to extract an illumination invariant representation of a face images from a raw facial images. Even though the proposed photometric normalization based representations increase the recognition performance, it is not suitable to say that these representations provide complete invariance against illumination. There are many recent works on illumination invariant face recognition. An extensive review of illumination invariant face recognition approaches is given by (Zou et al., 2007) and (Zhao & Chellappa, 2006).

There are several recent image-based studies on illumination invariant face recognition. Image-based methods are known to be robust to illumination variations. Main drawback of the image-based methods is that they always assume the face image is already aligned. Usually it is not an easy assumption to be satisfied especially when the input image is poorly illuminated. Appearance-based methods require training images of individuals taken under different illumination conditions. A method proposed by (Sim & Kanade, 2001) overcomes this restriction by using a statistical shape-from-shading model. Using this method they generate images for each of the individuals under different lighting conditions to serve as database images in a recognizer.

Face alignment is a crucial step to extracting good facial features and obtaining high performance in face recognition, expression analysis and face animation applications. Several face alignment methods were proposed by Active Shape Models (ASM) (Cootes et al., 1995) and Active Appearance Models (AAM) (Cootes et al., 2001; Stegmann et al., 2003), by Cootes *et al* which are two successful models for object localization. ASM utilizes local appearance models to find the candidate shape and global model to constrain the searched shape. AAM combines the constraints on both shape and texture variations in its characterization of facial appearance. In searching for a solution, it assumes linear relationships between appearance variation and texture variation and between texture variation and position variation. In this study, we have used AAM to solve the pose-invariant face alignment problem.

AAM is known to be very sensitive to illumination, particularly if the lighting conditions during testing are significantly different from the lighting conditions during training. Several variations of AAM appear in the literature to improve the original algorithm, namely Direct Appearance Models (Hou et al., 2001) and view-based AAM (Cootes et al., 2002). Cootes et al constructed three AAMs which are called as View-based AAMs. These models are linear model of frontal, profile and half profile views of faces. They also show how to estimate the pose from the model parameters. The approach in this study differs from their method in the way that only one AAM is constructed rather than three models. The motivation here is to reduce the three separate searching procedures to just one fitting procedure based on one linear statistical model. The model has better generalization performance in capturing pose variations than the one using three separate linear models. In order to construct the one linear model, a training dataset comprised of 8 different poses of 3 individuals captured under similar illumination conditions is used. Despite the success of

these methods, problems still remain to be solved. Moreover, under the presence of partial occlusion, the PCA-based texture model of AAM causes the reconstruction error to be globally spread over the image, thus degrading alignment. In this paper, we propose an approach based on histogram-fitting to overcome the problem explained above. A detailed explanation of the proposed approach is given in Section 2.

Yet another issue related to face recognition is to recognize different poses of the same person. Pose-invariant face recognition requires pose alignment where images are captured either by multiple cameras or by a single camera at different time instances. There are several works related to pose normalization. (Blanz & Vetter, 2003) use a statistical 3D morphable model (3DMM) to tackle with pose and illumination variations. Since their method requires textured 3D scans of heads, it is computationally expensive. The vertices in a 3DMM shape are much denser than an AAM shape. 3DMM achieved promising results for illumination invariant face recognition. However, fitting a dense model requires much higher computational effort, which is not suitable for real-time face recognition systems.

In Section 3, we will study the proposed AAM based approach capable of producing different poses of unseen person and explain how a non-frontal face is projected to a frontal face in detail. In this paper, we have focused on the problems induced by varying illumination and poses in face recognition. Our primary goal is to eliminate the negative effect of challenging conditions, especially illumination and pose, on the face recognition system performance through illumination and pose-invariant face alignment based on Active Appearance Model. The rest of the paper is structured as follows: Section 2 introduces Active Appearance Model (AAM) and Section 3 introduces illumination normalization inserted into the searching procedure of AAM. Section 4 is for the proposed pose invariant combined active appearance model. The experimental results and the conclusion are presented in Section 5 and 6, respectively.

2. Active Appearance Model

Active Appearance Models are generative models capable of synthesizing images of a given object class. By estimating a compact and specific basis from a training set, model parameters can be adjusted to fit unseen images and hence perform image interpretation. The modeled object properties are usually shape and pixel intensities (here denoted texture). AAM aims to find the optimal model parameters to represent the target image that belongs to the same object class by using an iterative scheme.

Training objects are defined by marking up each image with points of correspondence. Relying upon the landmarks, a triangulated mesh is produced for the reference position and orientation of the object. Before modeling variations, all shape vectors are normalized to a common reference shape frame by using Procrustes Analysis (Goodall, 1991). After obtaining the reference shape vector, all of the training images are warped to the reference shape by using a piecewise affine warping (Glasbey & Mardia, 1998), which is defined between corresponding triangles to obtain normalized texture vectors.

Using prior knowledge of the optimization space, AAMs can rapidly be fitted to unseen images with a reasonable initialization given. AAM uses principal component analysis (PCA) to model the variations of the shapes and textures of the images. Usage of PCA representation allows AAM to model and represent a certain image with a small set of parameters.

AAM works according to the following principle: An image is marked with n landmark points. The content of the marked object is analyzed based on a Principal Component Analysis (PCA) of both texture and shape. The shape is defined by a triangular mesh and the vertex locations of the mesh. Mathematically the shape model is represented as follows:

$$x_0 = \left((x_1, x_2, x_3, \ldots, x_n), (y_1, y_2, y_3, \ldots, y_n)\right) \in R^{2n} \tag{1}$$

Shape is reduced to a more compact form through PCA such that,

$$x = \bar{x} + \Phi_s b_s. \tag{2}$$

In this form, x is the synthesized shape in the normalized frame, Φ_s is a matrix that contains the *t*-eigenvectors corresponding to the largest eigenvalues and b_s is a *t*-dimensional vector of shape coefficients. By varying the parameters in b_s, the synthesized shape can be varied.

In the texture case one needs a consistent method for collecting the texture information (intensities) between the landmarks, i.e. an image warping function needs to be established. This can be done in several ways. Here, we used a piece-wise affine warp (Glasbey & Mardia, 1998) based on the Delaunay triangulation (Shewchuk, 1996).

All training images are warped to the reference shape and are sampled into a vector to obtain the texture vectors represented as *g*. Prior to the PCA modeling of the texture, we need to normalize all texture vectors. So, a photometric normalization of the texture vectors of the training set is done to avoid the side effects of global linear changes in pixel intensities. The aim of this normalization is to obtain texture vectors with zero mean and unit variance. Texture model can now be obtained by applying PCA to the normalized textures,

$$g = \bar{g} + \Phi_g b_g \tag{3}$$

where g is the synthesized texture, $\bar{g}$ is the mean texture and b_g is a *k*-dimension vector of texture parameters. In the linear model of texture, Φ_g is a set of orthogonal modes of variation.

To remove the correlation between shape and texture model parameters, a third PCA is applied on the combined model parameters, giving a further model,

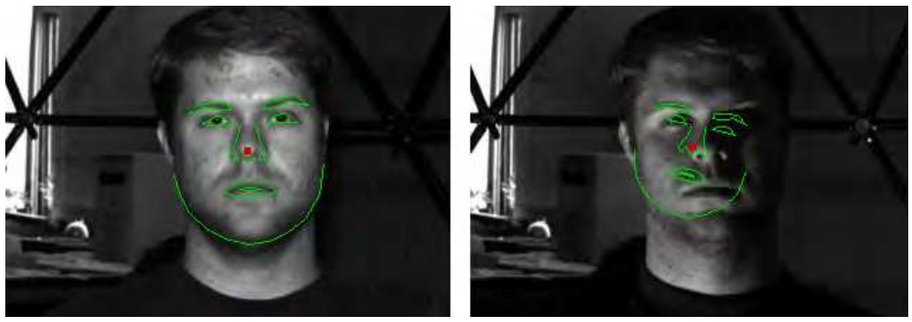
$$b = Q\, c \tag{4}$$

where Q is the eigenvectors and c is a vector of appearance parameters controlling both the shape and the texture of the model. Note we do not use a mean vector in this model since the shape, texture and appearance model parameters need to have zero mean. Due to the linear nature of the model, the shape and texture vectors can be expressed in terms of the appearance parameters c,

$$x = \bar{x} + \Phi_s W_s^{-1} Q_s c \tag{5}$$

$$g = \overline{g} + \Phi_g Q_g c, \tag{6}$$

where $b=[W_s b_s b_g]^T$ and $Q=[Q_s\ Q_g]^T$. In this form, W_s is a diagonal matrix of weights for each shape parameter, allowing for the difference in units between the shape and the grey models. Generally W_s is the square root of the ratio of the total intensity variation to the total shape variation. An example image can be synthesized for a given c appearance vector by generating the shape normalized image from the vector g and warping it using the control points described by x vector. Appearance parameters vector, c, controls both the shape and the grey-levels of the model. Q_s and Q_g are the eigenvectors of the shape and texture models respectively. An image can be represented by a vector p which is written in terms of x, g and c as $p=[x\ g\ c]^T$. It is possible to synthesize a new image by changing the parameter p.



(a) (b)

Fig. 1. Face alignment using standard AAM under good and extreme illumination.
(a) Normal illumination, (b) Extreme illumination.

The underlying problem with the classical AAM is demonstrated in Fig.1. In Fig.1 (a) a correct AAM search result is shown where the input image contains a frontal face which is also illuminated frontally. Since the model is constructed from a database containing frontally illuminated faces, the standard AAM searching procedure cannot converge to a meaningful solution for an extremely illuminated frontal face given in Fig.1 (b). We propose an illumination normalization method explained in Section 3 and insert it into the standard AAM searching procedure applied to the faces captured under different illumination conditions. The inserted normalization module guides the AAM to converge to a meaningful solution and also enhances the accuracy of the solution.

3. Illumination Normalization

We discuss here two light normalization methods and analyze their behavior in AAM searching. The first proposed method is ratio-image face illumination normalization method (Liu et al., 2005). Ratio-image is defined as the quotient between an image of a given face whose lighting condition is to be normalized and an image of the reference face. These two images are blurred using a Gaussian filter, and the reference image is then updated by an iterative strategy in order to improve the quality of the restored face. Using this illumination restoration method, a face image with arbitrary illumination can be restored to a face having frontal illumination.

The second normalization method discussed in this study is based on image histogram techniques. The global histogram equalization methods used in image processing for normalization only transfers the holistic image from one gray scale distribution to another. This processing ignores the face-specific information and cannot normalize these gray level distribution variations. To deal with this problem, researchers have made several improvements in recent years. The underlying problem is that well-lit faces do not have a uniform histogram distribution and this process gives rise to an unnatural face illumination. As suggested in (Jebara, 1996), it is possible to normalize a poorly illuminated image via histogram fitting to a similar, well illuminated image.

In this study, a new histogram fitting algorithm is designed for face illumination normalization taking the structure of the face into account. The algorithm is explained over poorly illuminated frontal face image where one side of the face is dark and the other side is bright. The main idea here is to fit the histogram of the input face image to the histogram of the mean face. The face is first divided into two parts (left/right) and then the histogram of each window is independently fitted to the histogram of mean face. For these two histograms, namely the histogram of the left window denoted as $H_{LEFT}(i)$ and the histogram of the right window denoted as $H_{RIGHT}(i)$, two mapping functions are computed: $f_{H_{LEFT}\to G}$ and $f_{H_{RIGHT}\to G}$ corresponding to the left and right windows, respectively. Here $G(i)$ is the histogram of the reference image which is also called as *mean face* in AAM. An artifact introduced by this mapping is the sudden discontinuity in illumination as we switch from the left side of the face to the right side. The problem is solved by averaging the effects of the two mapping functions with a linear weighting that slowly favors one for the other as we move from the left side to the right side of the face. This is implemented with the mapping function $f_{H_{TOTAL}\to G}$ defined as bellow:

$$f_{H_{TOTAL}\to G}(i) = leftness \times f_{H_{LEFT}\to G}(i) + \left(1 - leftness\right) \times f_{H_{RIGHT}\to G}(i) \tag{7}$$

Illumination normalization result is shown in Fig. 2 obtained by using the histogram fitting method explained above. As it can be seen from the figure the normalization method can produce more suitable images to be used in AAM search mechanism. The classical AAM search fails in all images given in the first row of Fig. 2. We will show in the next section that AAM search procedure can now converge to the correct shape for the restored image both in point-to-point error and point-to-curve error senses.

Fig. 3 presents several results obtained for Set 4 (left) and Set 3 (right) faces of different individuals having extremely dark and bright regions. A significant amount of improvement in quality can be easily verified from the experimental results. The dark parts now become somehow noisy whereas there are still some very bright areas.

Fig. 2. Illumination normalization using histogram fitting: On the top the input images are given, and on the bottom the normalized images are shown.

Fig. 3. Illumination normalization results for extreme cases: On the top the input images are given, and on the bottom the normalized images are shown.

4. Pose Normalization

Pose normalization is required before recognition in order to reach acceptable recognition rates. There are several studies related to pose normalization. (Blanz & Vetter, 2003) use a statistical 3D morphable model to tackle with pose and illumination variations. Since their method requires textured 3D scans of heads, it is computationally expensive. Cootes et al constructed three AAMs which are called as View-based AAMs (Cootes et al., 2002). We developed AAM based pose normalization method which uses only one AAM. There are two important contributions to the previous studies. By using the proposed method:

- One can synthetically generate appearances for different poses when a single frontal face image is available.
- One can generate frontal appearance of the face if only non-frontal face image is available.

Next section explains the proposed pose normalization and generation method in detail.

4.1 Pose Generation from 2D Images

The same variation in pose imposes similar effect on the face appearance for all individuals. Fig.4.a demonstrates how face texture and shape are affected by pose. Deformation mostly occurs on the shape whereas the texture is almost constant. Since the number of landmarks in AAM is constant, the wireframe triangles are translated or scaled as pose changes. Therefore, as we change pose, only wireframe triangles undergo affine transformation but the gray level distribution within these translated and rotated triangles remains the same. One can easily generate frontal face appearance if AAM is correctly fitted to any given non-frontal face of the same individual provided that there is no self-occlusion on face. Self-occlusion usually is not a problem for angles less than ±45.

For 2D pose generation, we first compute how each landmark point translates and scales with respect to the corresponding frontal counterpart landmark point for 8 different poses, and obtain a ratio vector for each pose. We use the ratio vector to create the same pose variation over the shape of another individual. In Fig.4.b, two examples are given where the landmark points of unseen individuals are synthetically generated using the ratio vector obtained from that different person.

Appearances are also obtained through warping in AAM framework, using synthetically generated landmarks given in Fig.4.b. These are shown in Fig.5. First column in Fig.5 shows the frontal faces and the second column shows appearances for various poses. It is important to note that the generated faces contain no information about the individual used in building the ratio matrix.

4.2 Training AAM for Pose Normalization

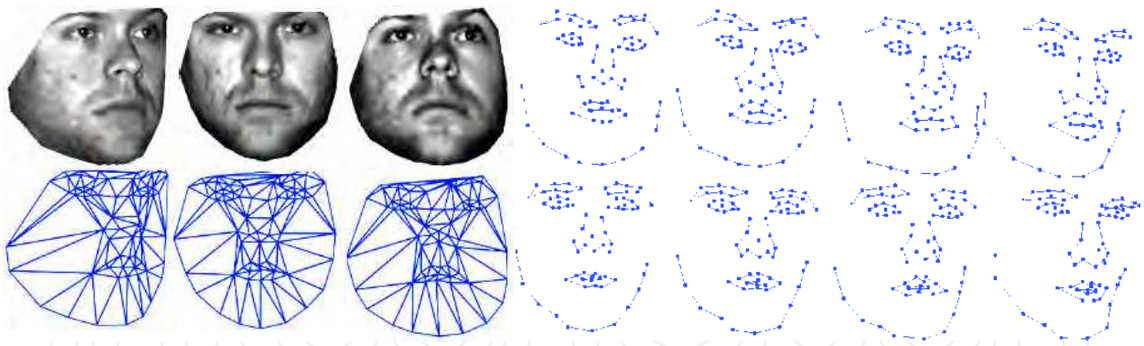
An AAM model trained by using only frontal faces can only fit into frontal faces well and fail to fit into non-frontal faces. Our purpose here is to enrich the training database by inserting synthetically generated faces at different poses so that AAM model trained by frontal faces can now converge to images at any pose.

We manually labeled 73 landmarks on 4920 images. Let us denote the landmark points on i^{th} frontal image as $S_i^0 = \left(\left(x_{i,1}, y_{i,1}\right), \left(x_{i,2}, y_{i,2}\right), \ldots, \left(x_{i,K}, y_{i,K}\right)\right) \in R^{2K}$ where $i = 1, 2, \ldots, N$. N is *4920* and K=*73* in our database. The shape-ratio vector explained in the previous subsection (3.1) is defined between the p-posed shape and the frontal shape as

$$r_p(S^p, S^0) = \left(\left(\frac{x_{p,1}}{x_{0,1}}, \frac{y_{p,1}}{y_{0,1}}\right), \ldots, \left(\frac{x_{p,K}}{x_{0,K}}, \frac{y_{p,K}}{y_{0,K}}\right)\right) \tag{8}$$

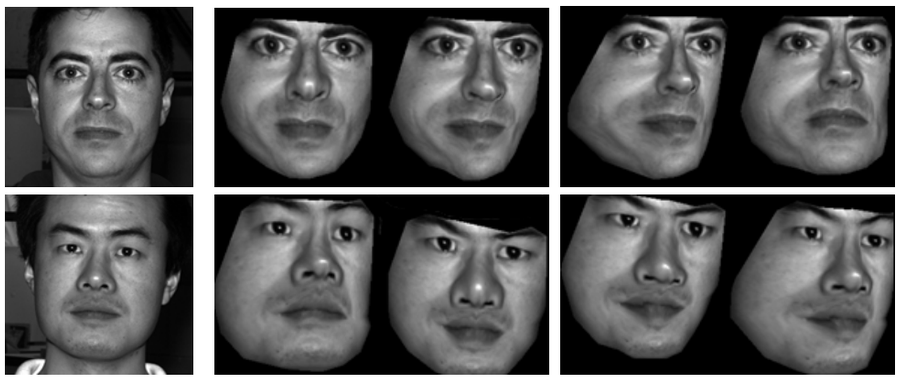
Shape of any unseen individual at pose p can now be easily obtained from frontal shape using shape-ratio vector r_p as

$$\hat{S}_{unseen}^p = r_p S_{unseen}^0. \tag{9}$$



(a) (b)

Fig. 4. Pose variations and synthetically generated landmarks. a) Texture and shape triangles variations due to pose variations, b) Synthetically generated landmark points given in the first and the second rows are generated by using the ratio matrix obtained from the landmarks in the database.



(a) (b)

Fig. 5. Synthetic pose generation from frontal face: a) Frontal face, b) Synthetically generated non-frontal faces.

Shapes in the database for $p = 8$ different poses can be synthesized from frontal-view images as,

$$\hat{S}_i^p = r_p S_i^0 \text{ , } i=1,2,\ldots,10, \text{ and } p=1,2,..,8. \tag{10}$$

AAM shape component is constructed from these aggregated shapes, $\hat{S}_i^p$ and S_i^0, by applying principal component analysis as $S = \overline{S} + Q_s s$ where $\overline{S}$ is the mean shape, Q_s contains k eigenvector of the covariance matrix corresponding to the highest k eigenvalues. Next step is to wrap each face in the training database to mean shape ($\overline{S}$) and apply the principal component analysis to the texture, this time as $T = \overline{T} + Q_t t$ where $\overline{T}$ is called as mean face. Any shape (S) and texture (T) can be steadily mapped to the AAM subspace as $s = Q_s^T \left(S - \overline{S} \right)$ and $t = Q_t^T \left(T - \overline{T} \right)$.

AAM is comprised of both shape (Q_s) and texture (Q_t) subspaces. Any change in face shape leads to a change in face texture and vice versa. Face appearance (A) is dependent on shape and textures. This dependency is expressed as $A=[\Lambda s\ t]^T$. In order to exploit the dependency between shape and texture modeled by the diagonal matrix (Λ), one further PCA is applied to the shape and texture components collectively and we obtained the combined model called appearance model as $A=Q_aA$. Any appearance is obtained by a simple multiplication as $a = Q_a^T A$.

In order to show how rich representation AAM provides us, we used the first *5* coefficients and select random points in *5*-dimensional space. The corresponding faces are plotted in Fig.6. Even this simple experiment proves that AAM trained as explained above can generate pose variations not governed by any shape ratio vector (r_p).

Fig. 6. Randomly synthesized faces from leading 5 AAM parameters.

5. Experimental Results

AAM combines the shape and texture model in one single model. The alignment algorithm (also called AAM searching) optimizes the model in the context of a test image of a face. The optimization criterion is the error occurring between a synthesized face texture and the corresponding texture of the test image.

Due to the illumination problems the error can be high and the classic searching algorithm fails. In the proposed approach, we normalize the corresponding texture in the test image just before we compute the error. We tested the proposed method on the Yale-B face dataset (Georghiades et al., 2001). The total number of images under different lighting conditions for each individual is 64. The database is portioned into four sets identified as Set 1-4. Set 1

contains face images whose light direction is less than ±20 degrees. Set 2 contains face images whose light directions are between ±20 and ±50 degrees. Set 3 contains face images whose light directions are between ±50 and ±70 degrees. Set 4 contains face images whose light directions are greater than ±70 degrees. All details about the Yale B dataset are given in (Georghiades et al., 2001). We manually labeled 4920 images. To establish the models, 73 landmarks were placed on each face image; 14 points for mouth, 12 points for nose, 9 points for left eye, 9 points for right eye, 8 points for left eyebrow, 8 points for right eyebrow and 11 points for chin. The warped images have approximately 32533 pixels inside the facial mask. We constructed a shape space to represent 95% of observed variation. Then we warped all images into the mean shape using triangulation. Using normalized textures, we constructed a 21-dimensional texture space to represent 95% of the observed variation in textures and for shapes we constructed a 12-dimensional shape space to represent 95% of the observed variation in shapes. Finally, we constructed a 15-dimensional appearance space to represent 95% of the total variation observed in the combined (shape and texture) coefficients.

Using a ground truth given by a finite set of landmarks for each example, performance can be easily calculated. A distance measure $D(x_{gt}, x)$ is computed in a leave-one-out setting, and it gives a scalar interpretation of the fit between the two shapes, i.e. the ground truth (x_{gt}) and the optimized shape (x). Two distance measures defined over landmarks are used to obtain the convergence performance. The first one is called point-to-point error, defined as the Euclidean distance between each corresponding landmark:

$$D_{pt.pt.} = \sum \sqrt{\left(x_i - x_{gt,i}\right)^2 + \left(y_i - y_{gt,i}\right)^2} \tag{11}$$

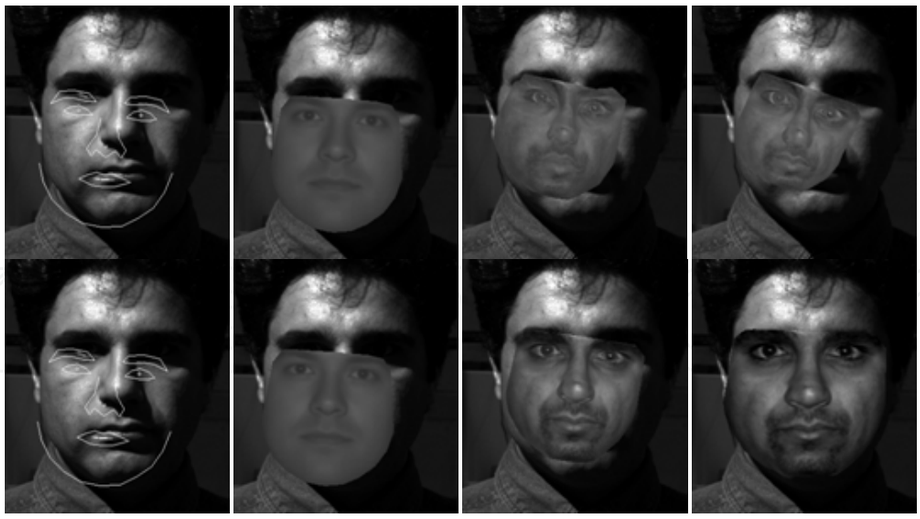
The other distance measure is called point-to-curve error, defined as the Euclidean distance between a landmark of the fitted shape (x) and the closest point on the border given as the linear spline, $r(t)=(r_x(t),r_y(t)), t\in[0,1]$, of the landmarks from the ground truth (x_{gt}):

$$D_{pt.crv.} = \frac{1}{n}\sum_{i=1}^{n} \min_t \sqrt{\left(x_i - r_x(t)\right)^2 + \left(y_i - r_y(t)\right)^2} \tag{12}$$

We calculated these errors on all images in the datasets (from Set 1 to Set 4). We conducted an experiment to see how close we fit into unseen faces at different poses.

To match a given face image with the model, an optimal vector of parameters are searched by minimizing the difference between synthetic model image and input image. Fig.7 illustrates the optimization and search procedures for fitting the model to input image. The first column of the figure is the arbitrarily illuminated unseen image from test dataset and the remaining images (columns) are the steps of the optimization. The fitting results are rendered at each iteration for classical AAM (the first row) and the proposed method (the second row).

The AAM searching is known to be very sensitive to the selection of initial configuration. We tested the proposed method against the selection of initial configuration. We translate, rotate and scale initial configurations and see how the proposed method can handle the poor initialization. We made 10 experiments for each test image with different initializations



(a) (b) (c) (d)

Fig. 7. Searching results: The first row is the classical AAM searching results and the second row is the proposed method. (a) Initial configuration (b) Mean face (c) Searching result obtained in the third iteration (d) Searching result obtained in the sixth iteration.

and took the average error. These experiments include mean-shape configuration, ±5 degrees rotation, scaling by 0.85 and 0.95, translation by 10% in x and y directions.
Table.1 summarizes the averages of point-to-point and point-to-curve errors when classical AAM search is used without any illumination normalization. Point-to-point and point-to-curve errors obtained by the proposed illumination normalization method are much less than the errors obtained by the classical AAM (Table.2).

	Yale B face database subsets			
	$Subset_1$	$Subset_2$	$Subset_3$	$Subset_4$
Pt.-pt.	4.9±0.2	11.4±0.5	19.4±0.58	36.6±1.6
Pt.-crv.	2.9±0.1	6.8±0.33	12.9±0.36	33.2±1.4

Table 1. Standard AAM fitting performance.

	Yale B face database subsets			
	$Subset_1$	$Subset_2$	$Subset_3$	$Subset_4$
Pt.-pt.	4.1±0.12	8.06±0.3	13.03±0.4	21.3±0.5
Pt.-crv.	2.4±0.08	5.24±0.23	8.76±0.3	14.7±0.4

Table.2 Proposed AAM fitting performance.

Ratio-image method is not suitable for AAM searching, at least for the first iterations of the algorithm. Let's suppose that we start searching in a position far away from the ground truth location. The model synthesizes a face that best fits the current location. Then the textures of the synthesized face and corresponding part in the test image are analyzed and an error coefficient is computed, reflecting the similarity degree of the two textures. We normalize the corresponding texture in the test image before computing the error. The main

problem with the ratio-image method is that when it is applied to a region of an image that is not face-like, the normalization result will include a lot of information of the mean face, in other words, it will be mean-face-like. Thus the error will be much smaller than the real one, and it will introduce false alarm in the searching process creating additional local minima. On the other hand, the histogram based normalization method will never change the general aspect of an image, only the pixel intensities follow a different distribution. Thus the chances of introducing false alarms are reduced using this normalization method. The ratio-image can produce very good results provided that the shape is already aligned. But this is not the case in AAM searching. We assume that the best fit returned by the searching algorithm using histogram-based normalization is a good approximation of the real face, and thus the alignment requirement is satisfied. Fig.8 summarizes the alignment results for these unseen faces.

We also analyze how the proposed alignment method affects the recognition performance. We used the following feature spaces in our experiments: PCA and LDA. Randomly selected *25* images of each person from Set 1 dataset are used in training. All datasets (Set 1 through Set 4) contain faces of all poses. The remaining faces in Set 1 dataset are used as test data. Recognition rates for two feature spaces (i.e. PCA and LDA) in Set 1-4 are plotted in Fig.10 for increasing dimensions. The recognition rates obtained when the original images are used as input to the classifier are denoted as ORG-PCA and ORG-LDA. The recognition rates obtained when the images restored by RI are used as input and are denoted as RI-PCA and RI-LDA. Finally, the recognition rates obtained when the images restored by HF are used as input and are denoted as HF-PCA and HF-LDA. PCA is known to be very sensitive to misalignment in faces. Our experimental studies also verify this behavior. When the original images are used, the PCA recognition rates for all sets are poor. LDA is more successful if dimension is closer to 9. ORG-PCA reaches to 74.36% at most, while ORG-LDA reaches to 91.26% at most in Set 1. This performance drops to 30.99% for ORG-PCA and to 41.13% for ORG-LDA in Set 4.

One important observation is that AAM alignment with histogram fitting always leads to better recognition rates in all test sets (Set 1- 4) compared to the case where original faces are used and ratio-image normalization is used right after the AAM alignment. Another advantage of the proposed method is that similar recognition performance is obtained at lower dimensions. Recognition rate for ORG-LDA is just 32.81% while LDA performance for the proposed approach (called HF-LDA) is 83.38% when the dimension is set to 3. ORG-LDA catches this rate when the dimension is set to 5.

For the challenging test set, i.e. Set 4, both ORG-LDA and ORG-PCA fail. The recognition rate is at most 30.99% for ORG-PCA and 41.13% for ORG-LDA. On the other hand, HF-PCA reaches to 76.20% at most and HF-LDA reaches to 82.68% at most. This is a significant improvement when compared to the results obtained without applying any preprocessing (41%). Note that all test sets include faces of 8 different poses selected from Yale B dataset.

6. Conclusion

In this study we developed AAM based on face alignment method which handles illumination and pose variations. The classical AAM fails to model the appearances of the same identity under different illuminations and poses. We solved this problem by inserting histogram fitting into the searching mechanism and inserting synthetically generated poses

of the same identity into the training set. From the experimental results, we showed that the proposed face restoration scheme for AAM provides higher accuracy for face alignment in point-to-point error sense. Recognition results based on PCA and LDA feature spaces showed that the proposed illumination and pose normalization outperforms the standard AAM.

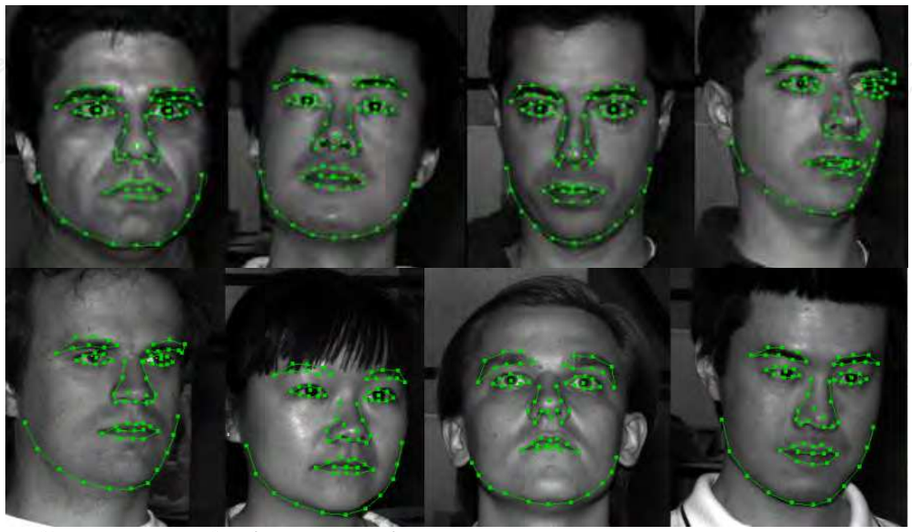


Fig. 8. Face alignment results for unseen faces.

Fig. 9. Initialization (the first row) and alignment/restoration results of the proposed method (the second row) for different pose and illumination variations.

7. Acknowledgement

This work was partly supported by the National Scientific and Research Council of Turkey; project no: EEEAG-104E121 and the State Planning Agency (DPT) of Turkey. We would like to thank Florin S. Telcean for his initial contributions to this work.

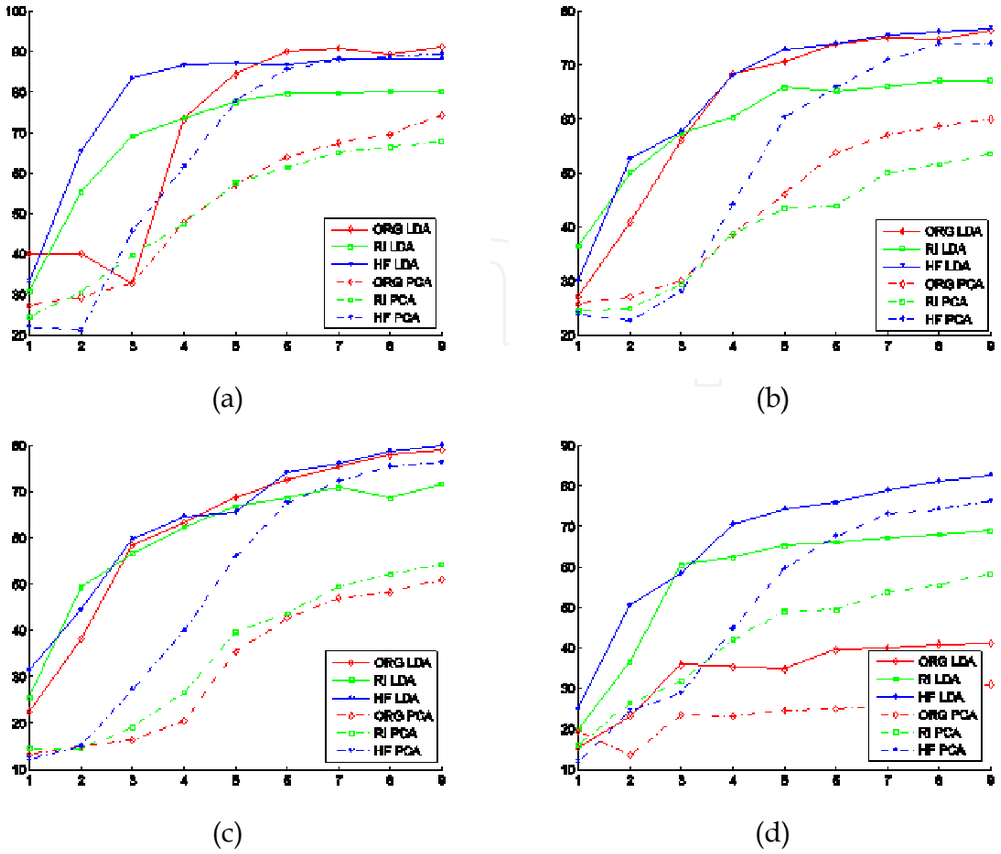


Fig. 10. PCA and LDA recognition rates for Set 1 (a), Set 2 (b), Set 3 (c), and Set 4 (d) when original face (ORG), Ratio Image (RI) and the proposed restoration (HF) are used.

8. References

Zhao, W. & Chellappa R. (2000). SFS Based View Synthesis for Robust Face Recognition. *Proc. 4th Conf. on Automatic Face and Gesture Recognition*, 2000.

Georghiades, A.S.; Belhumeur, P. N., & Kriegman, D. J. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no.6, pp.643–660, 2001.

Ramamoorthi R. (2002). Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 10, 2002.

Basri, R. & Jacobs, D. (2001). Photometric Stereo with General, Unknown Lighting. *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR), vol. 2, pp. 374-381, 2001.

Zhou S.; Aggarwal G.; Chellappa R. & Jacobs D. (2007). Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination invariant face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 230-245, 2007.

Lee J.; Moghaddam B.; Pfister H. & Machiraju R. (2005). A bilinear illumination model for robust face recognition. IEEE Conf. on ICCV, pp. 1177-1184, 2005.

Shashua A. & Riklin-Raviv T. (2001). The Quotient Image: Class-Based Re-Rendering and Recognition With Varying Illuminations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 129-139, 2001.

Wang H.; Li S. Z & Wang Y (2004). Generalized quotient image, *IEEE Proceeding Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 498-505, 2004.

Zhang Y.; Tian J.; He X. & Yang X. (2007). MQI Based Face Recognition Under Uneven Illumination, *Advances in Biometrics*, vol. 4642, pp. 290-298, 2007.

An G.; Wu J. & Ruan Q. (2008). Kernel TV-Based Quotient Image Employing Gabor Analysis and Its Application to Face Recognition", IEICE *Transactions on Information and Systems*, vol. E91-D, no. 5, pp. 1573-1576, 2008.

Sim T. & Kanade T. (2001). Combining models and exemplars for face recognition: An illuminating example, *IEEE Proceeding Conference on Computer Vision and Pattern Recognition*, *Workshop on Models versus Exemplars in Computer Vision*, 2001.

Cootes, T.F.; Taylor, C.J.; Cooper, D.H. & Graham, J. (1995). Active Shape Models-their training and application. *Computer Vision and Image Understanding*, 61(1), pp. 38-59, 1995.

Cootes, T.F.; Edwards, G. & Taylor, C.J. (2001). Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, 2001.

Stegmann, M.B.; Ersboll, B.K. & Larsen, R. (2003). FAME - A Flexible Appearance Modeling Environment. *IEEE Trans. on Medical Imaging*, vol. 22, no.10, pp.1319-1331, 2003.

Hou, X.; Li, S.; Zhang, H. & Cheng, Q. (2001). Direct appearance models. *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 828–833, 2001.

Cootes, T.F.; Wheeler G.; Walker, K. & Taylor, C. (2002). View based active appearance models. *Image and Vision Computing*, vol. 20, pp.657–664, 2002.

Blanz V. & Vetter T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence, v*ol. 25, vo.9, pp. 1063-1074, 2003.

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society*, vol.B53, no.2, pp. 285-339, 1991.

Glasbey, C.A. & Mardia, K.V. (1998). A review of image warping methods. Journal of Applied Statistics, Vol. 25 (2), 155-171, 1998.

Jebara, T. (1996). 3D Pose Estimation and Normalization for Face Recognition, B. Thesis, McGill Centre for Intelligent Machines, 1996.

Shewchuk J.R. (1996). Triangle: engineering a 2D quality mesh generator and Delaunay triangulator. *Workshop on Applied Computational Geometry. Toward Geometric Engineering*., 203–222. Springer-Verlag, 1996.

Zhao, W. & Chellappa R. (2006). *Face Processing: Advanced Modeling and Methods*. Academic Press, Elsevier, 2006.

Liu, D.H.; Lam, K.M. & Shen, L.S. (2005). Illumination invariant face recognition. *Pattern Recognition*, vol. 38, no.10, pp. 1705-1716, 2005.

Face Recognition

Edited by Milos Oravec

ISBN 978-953-307-060-5
Hard cover, 404 pages
Publisher InTech
Published online 01, April, 2010
Published in print edition April, 2010

This book aims to bring together selected recent advances, applications and original results in the area of biometric face recognition. They can be useful for researchers, engineers, graduate and postgraduate students, experts in this area and hopefully also for people interested generally in computer science, security, machine learning and artificial intelligence. Various methods, approaches and algorithms for recognition of human faces are used by authors of the chapters of this book, e.g. PCA, LDA, artificial neural networks, wavelets, curvelets, kernel methods, Gabor filters, active appearance models, 2D and 3D representations, optical correlation, hidden Markov models and others. Also a broad range of problems is covered: feature extraction and dimensionality reduction (chapters 1-4), 2D face recognition from the point of view of full system proposal (chapters 5-10), illumination and pose problems (chapters 11-13), eye movement (chapter 14), 3D face recognition (chapters 15-19) and hardware issues (chapters 19-20).

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen