Image Normalization for Face Recognition using 3D Model

Zahid Riaz, Michael Beetz and Bernd Radig

Abstract— **This paper describes an image segmentation and normalization technique using 3D point distribution model and its counterpart in 2D space. This segmentation is efficient to work for holistic image recognition algorithm. The results have been tested with face recognition application using Cohn Kanade Facial Expressions Database (CKFED). The approach follows by fitting a model to face image and registering it to a standard template. The models consist of distribution of points in 2D and 3D. We extract a set of feature vectors from normalized images using principal components analysis and using them for a binary decision tree for classification. A promising recognition rate of up to 98.75% has been achieved using 3D model and 92.93% using 2D model emphasizing the goodness of our normalization. The experiments have been performed on more than 3500 face images of the database. This algorithm is capable to work in real time in the presence of facial expressions.**

I. Introduction

Face recognition has been one of the challenging fields over last three decades. As a result, many commercially available systems exist after face recognition grand challenge (FRGC) [1] and face recognition vendor test (FRVT) [2]. Owing to computationally stronger machines and high resolution computer vision cameras, this application has made its place prominent in biometric market and is currently comparable to other biometrics as shown in Figure 1. However many researchers have solved face recognition problem in constrained environments while the outstanding challenges still exist in unconstrained environments. Face recognition in machine vision is not robust to aging factors, facial expression, poses and illumination variations. In recent decade, researchers have tried to address the problem of unconstrained face recognition but still lagging in a system working under aforementioned constraints. Existing systems which can handle more constrains at a time are either not computationally efficient to be applied directly to real environment or requires some specially sensors for data acquisition [3]. Current advancements in the field of computer vision have innovated modeling of faces as non rigid and deformable objects. We use model based approach to face recognition which is not only efficient but also produces a promising recognition rate in the presence of facial expression variations. We segment a face from one image such that all facial features coincide with the corresponding features of other library images. We use models which consist of distribution of points in 2D and 3D spaces. A model used in 2D consists of 134 anatomical landmarks corresponding to different face features [4]. Whereas a 3D model is a wireframe model called Candide-III [5]. Segmentation results are different for both models and hence producing different recognition rate for same set of images of the database.

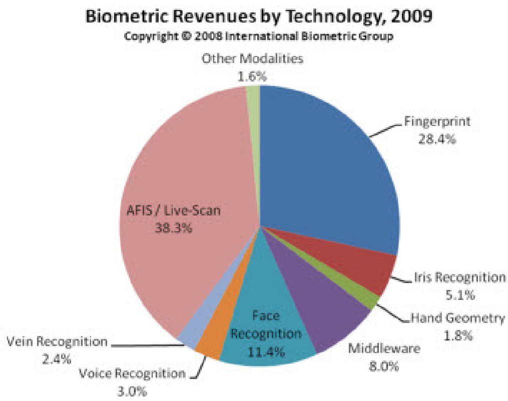

Figure 1 Biometric market

The remainder of this paper is divided in four sections. Section II describes some of the related work in this field, section III deals with our approach in detail. In section IV we describe the process of segmenting face image using appearance information from the images. In the last section, we perform experiments and compare our results. In general, our approach is same for 2D and 3D models but it is explicitly explained where required.

II. Related work

In last decade of nineteen century a revolutionary face recognition algorithm was developed by Turk and Pentland, called eigenfaces [6,7]. This technique was inspired by the earlier work of Sirovich [8]. Eigenfaces, on one hand are not robust to the variations in illuminations and poses while on the other hand are real time capable. Further it requires stronger image normalization. Keeping these pits and falls of this algorithm in view researchers have emphasis a lot on image normalization resulting in improvement of the system [9, 10]. A proposed solution to these problems in computer vision applications is to use a point distribution model (PDM). We use 3D model for face image segmentation meanwhile comparing the results to its counterpart 2D model. A well-known face model in 2D was introduced by

Manuscript received March 16, 2009.

Zahid Riaz is with the Department of Informatics, Technical University of München, 85748 Germany. (+49-89-28917753; fax: +49-89-28917757; e-mail: riaz@in.tum.de).

Michael Beetz is with the Department of Informatics, Technical University of München, 85748 Germany. (e-mail: beetz@in.tum.de).

Bernd Radig is with the Department of Informatics, Technical University of München, 85748 Germany. (e-mail: radig@in.tum.de).

Edwards et. al. [11] and is known as active appearance model (AAM). This is a generative model parameterized for shape and texture variations of a face. Shape information alone is captured using active shape model (ASM) [12]. AAM is further applied by many researchers in many applications [13, 14]. In case of 3D we use another well-known model called Candide-III [5]. This is the third version of the Candide model supported with facial action coding systems (FACS) [15] with 116 anatomical landmarks.

The problem of person identity lies under the huge area of pattern recognition. Various techniques have been used over last few decades. A formal method of classifying the faces was first proposed by Francis Galton [16] in 1888. Galton proposed collecting facial profiles as curves, finding their norm, and then classifying other profiles by their deviations from the norm. The classification was resulting in a vector of independent measure that could be compared with other vectors in the database. Traditional recognition systems have the abilities to recognise the human using various techniques like feature based recognition, face geometry based recognition, classifier design and model based methods. A term defined in [17] is Distinguishing Component Analysis (DCA) and is compared to Principal Components Analysis (PCA) attempting to capture what distinguishes one subject from the others. In [18] Luca and Roli have used the similar kind of technique, which is fusion of Linear Discriminant Analysis (LDA) and PCA is used. PCA is used to reduce the data size while LDA is used to classify the data. Similar application is performed by [19], which first used PCA and LDA in second step. The idea of combination of PCA and LDA is to improve generalization capability when few samples per class are given. PCA and LDA are two commonly used techniques for data classification and dimensionality reduction, belonging to the same class [20]. A parameterized recognition of the faces is performed by Edwards et al [21]. A segmented face is used for eigenfaces in [9] for 2D model only. They have used models parameter for recognition using LDA. In last two decades, eigenfaces are adopted by many researchers [22]. This paper focuses on using modelled based approach to segment face area out of the image and extracting texture information and then utilising this model for recognition purposes.

III. OUR APPROACH

The goal of this paper is to normalize and segment the face image addressing its efficiency and accuracy to be used for holistic face recognition techniques like eigenfaces. Further, we want to develop a system that could be able to recognize persons in real time under different variations. A model based technique followed by eigenfaces is utilized in this regard. AAM is used with shape and textural information to achieve strong recognition results but we use texture information only after image segmentation and achieve a good accuracy on benchmarked database in the presence of strong facial expressions.

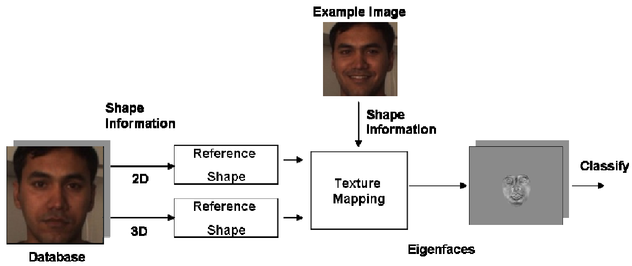

Figure 2 Our Approach

All the subjects in the database are labelled for identification. An active shape model (ASM) is fitted to all the face images in 2D and a wireframe model is used in 3D case. Shape model could be fitted using any fitting algorithm however we adopted fitting by designing local objective functions [4]. A fitted shape model on training face images is then used for defining the reference shape which is the mean shape of all the shapes in our database. Given a reference shape S_{ref} and an example image I with fitted shape S, it is suitable to apply texture mapping from example shape to reference shape using affine transformation. However prior to mapping we use planar subdivision i.e. delaunay triangulation for reference and example shape. The texture vector for corresponding image is stored as T. In case of 3D, we define 2D correspondences of the points of our wireframe model using camera parameters and perform triangulations in 2D. Instead of parametrizing our model we use this representative texture for each image to find eigenfaces. The use of eigenspace at this stage is suitable in two major aspects, 1) reducing the amount of texture to be used in classification i.e. segmenting about 12% gray values in 2D and 15% gray values information in 3D for the original image, compared to conventional eigenface approach, where we use full image texture, 2) Texture mapping on reference shape benefited in terms of shape normalization for all the available images, this makes facial expressions mould to reference shape. In this regard we can also control the facial expressions in short extent. The approach is shown in Figure 2. This established framework could be used either to parameterize the appearance model or to use the texture information for eigenface classifier. Given a texture vector T, it could be projected to the eigenspace using a linear combination:

$$T = \sum_{n} w_i \Phi_i$$

Where w_i are the weights for the example image and Φ_i are the corresponding eigenvectors over the whole space of n dimensions.

IV. FACE SEGMENTATION

A. Shape Models

Different kind of shape models have been utilized by the researchers depending upon different applications. Some are the landmark based models defining some fixed points annotated on the images and then defining the boundaries around the objects. However some rely on the contour based approach. Different contour define the shape of the object for outlining it along with covering the feature inside an object [23]. Landmark based models however provide the exact location of the features inside the object. We utilize a 2D face model as shown in Figure 3. This point distributed model (PDM) consists of 134 landmarks which prominently describe the location of individual face features. This model covers full face area except some parts of the human head like hair, forehead and ears. Shape model is projected on the image plane and fitted to the face. We used fitting of the shape on the face by training an objective function. This technique was devised by Wimmer et al [4].

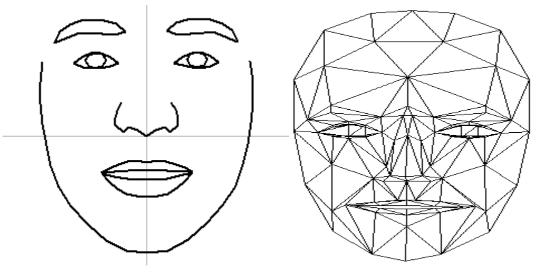

Figure 3 2D Shape model (left), 3D Shape model (right)

An ASM is parameterized using PCA to form the shape feature vector.

$$x \approx x_m + P_s b_s$$

Where the shape x is parameterized by using mean shape x_m and matrix of eigenvectors P_s to obtain the parameter vector b_s [24]. However in our case face parameters are only required for model fitting.

In case of 3D, we use the model shown in Figure 3. This model is projected to image and fitted to face. The fitting results are shown in Figure 4 and Figure 5 for both models.

B. Appearance Models

For the various instances of the same person different types of variations are required to be modelled. For example, shape deformations including both expression changes and pose variations along with the texture variations caused by illuminations, occlusions and makeups. Once we have information about the model for the example image, we can extract texture information easily.

At first, shape variation is required to be controlled in order to record the texture. This is achieved by defining a reference shape (mean shape in our case) for whole database. Figure 4 (bottom-left) shows the average shape (mean shape) for 2D model of the subject in consideration. Delaunay triangulation is used to divide the shape into a set of different facets. The delaunay triangulation is a triangulation which is equivalent to the nerve of the cells in a voronoi diagram, i.e., that triangulation of the convex hull of the points in the diagram in which every circumcircle of a triangle is an empty circle [25].

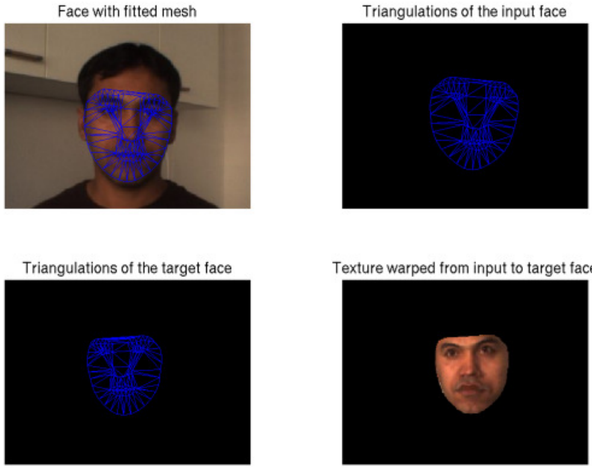

Figure 4 (Top Left) Face Fitted with Mesh, (Top Right) Mesh of the input image, (Bottom Left) Average shape where the texture is warped, (Bottom Right) Texture warping results on average face

Given a set of shape points x of the input example image and x_{avg} of the average image, we can find the texture vector g_{im} as follows:

a) Compute the pixel position in the average shape.

b) Find the relative position in example image using affine transformation.

c) Sample the texture values at the points inside convex hull of the average image forming texture vector g_{im}.

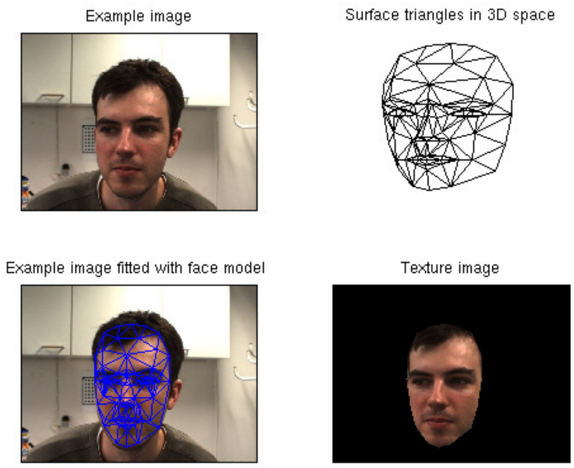

Figure 5 (Top Left) Example Image, (Top Right) Wireframe model in 3D, (Bottom Left) Model fitted to face in 2D, (Bottom Right) Extracted texture.

The texture vector is normalized to remove global lighting effects. This is performed by applying the linear transformation [24],

$$g = (g_{im} - \beta 1) / \alpha$$

Where

$$\beta = (g_{im} . 1) / n$$

$$\alpha = | g_{im} |^2 / n - \beta$$

Where, g is the texture vector obtained after light variations adjustment. Varying textural parameters cause changes in the image similar to eigenfaces. Face images segmented after appearance model are shown in Figure 4 and Figure 5 for 2D and 3D face models respectively. We repeat same process for 3D by knowing the 2D correspondences of the points of our model.

V. Eigenfaces for Face Recognition

Eigenfaces approach is taken from information theory and consists of calculating Principal Components Analysis (PCA). Using PCA for face recognition is to express the large 1-D vector of pixels constructed from 2-D facial image into the eigenspace projection.

The idea of the eigenfaces was given by Turk and Pentland in early 1990's. Recognition is performed by projecting a new image into the subspace spanned by the eigenfaces ("face space") and then classifying the face by comparing its position in face space with positions of known individuals. This is a holistic approach representing the image as a linear combination of the eigenfaces. The extracted features do not represent local face features rather they contain global face information [6,7].

In mathematical terms, we wish to find the principal components of the distribution of faces, or the eigenvectors of the covariance matrix of the set of face images. These eigenvectors can be thought of as a set of features, which together characterize the variation between the face images. Each image location contributes more or less to each eigenvector, so that these vectors can be displayed as a sort of ghostly face, which are called the eigenfaces. Figure 6 shows the eigenfaces in our scenario for both types of the models The amount of texture is reduced up to 12% of the original image for 2D whereas 15% approximately for 3D model.

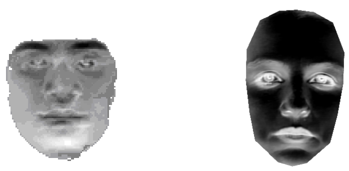

Figure 6 Eigenfaces for segmented frontal face images with 2D model (left), with 3D model (right)

VI. Experimental Evaluations

Keeping in view the facial expressional changes in daily life, we try to recognize a person under varying facial expression. In this regard we choose Cohn-Kanade-Facial-Expression database (CKFE-DB) [26]. CKFE-DB contains 488 short image sequences for persons performing the six universal facial expressions. Each sequence shows a neutral face at the beginning and then develops into the peak expression. Furthermore, a set of action units (AUs) has been manually specified by licensed Facial Expressions Coding System (FACS)-experts for each sequence. The subjects in this database were said to produce exaggerated expressions, hence expressions are not natural.

We use binary decision trees (BDT) as classifier with 10 fold cross validation in weka [27]. We experiment on a subset of aforementioned dataset consisting of more than 3500 images of 49 different persons. A recognition rate of 97.75% was obtained on these image sequences in case of 3D model. We train our classifier with same set of images and classifier with same specification and achieved a recognition rate of 92.93%. Figure 7 shows true positive rate (TPR) and false positive rate (FPR) for models used in this paper.

VII. Conclusions

This paper discusses about a technique to segment the face images and then using eigenfaces for recognition purpose. In image segmentation phase, each image is mapped to a pre-defined shape and hence segment and normalizes our shape information. Since faces contain facial expressions so after textural mapping the expressions mould to the expression of the target shape. This helps in transforming all the images towards unique expression of the reference shape and training the classifier for the optimum results. Eigenface approach is well-known and has computational ability to run efficiently in real time. We used a benchmark database to show our result, however this system can be applied in real time for daily life applications. In future, this system could be coupled with other sources of variations like lighting, poses and occlusions.

Acknowledgement

This work is mainly supported by cluster of excellence Cognition for Technical Systems (CoTeSys) at Technical University of Munich, Germany, for human robot joint activities. Further we are thankful to Higher Education Commission (HEC) of Pakistan and German Academic Exchange Service (DAAD).

References

[1] P. J. Philips et. al. "Overview of face recognition grand challenge", IEEE Conference on Computer Vision and Pattern Recognition, 2005.

[2] P. J. Philips et. al., "FRVT 2006 and ICE 2006 large-scale results", National Institute of Standard and Technology, 2006.

[3] Blanz, V., Vetter, Face Recognition Based on Fitting a 3D Morphable Model, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.25 no. 9, pp 1063 - 1074, 2003.

[4] Wimmer M, Stulp F, Tschechne S, and Radig B, "Learning Robust Objective Functions for Model Fitting in Image Understanding Applications", In Proceedings of the 17th British Machine Vision Conference, pp 1159–1168, BMVA, Edinburgh, UK, 2006.

[5] J. Ahlberg, *An Experiment on 3D Face Model Adaptation using the Active Appearance Algorithm*, Image Coding Group, Deptt of Electric Engineering, Linköping University.

[6] Turk M.A., Pentland A.P., "Face Recognition using Eigenfaces", IEEE Conference on Computer Vision and Pattern Recognition, 1991, pp 586-591.

[7] Turk M.A., Pentland A.P., “Eigenfaces for Recognition”, Journal of Cognitive Neuroscience **3** (1): 1991, pp 71–86.

[8] L. Sirovich and M. Kirby, “Low-dimensional procedure for the characterization of human faces”, J. Opt. Soc. Am. A, Vol. 4, No. 3, March 1987, pp 519-524

[9] Z. Riaz, M. Beetz and B. Radig, “Shape invariant recognition of segmented human faces using eigenfaces”, IEEE International Multitopic Conference, Pakistan, 2008.

[10] Moghaddam B., Nastar, C. and Pentland A., “A bayesian image similarity for direct image matching”, International Conference on Pattern Recognition, Austria, 1996.

[11] Edwards, G.J., Taylor, C.J., Cootes, T.F., “Interpreting Face Images using Active Appearance Models”, Proceedings of International Conference on Automatic Face and Gesture Recognition (1998) 300-305.

[12] T.F. Cootes et al, “Active shape models – their training and applications”, Computer Vision and Image Understanding 1995.

[13] Stegmann M.B., Active Appearance Models: Theory Extensions and Cases, Master Thesis, Technical University of Denmark, 2000.

[14] J. Xiao et al, “Real-time combined 2D+3D active appearance models”, Proceedings of IEEE conference on Computer Vision and Pattern Recogntion, 2004.

[15] Paul Ekman and Wallace Friesen. The Facial Action Coding System: A Technique for The Measurement of Facial Movement. Consulting Psychologists Press, San Francisco, 1978

[16] William A. Barrett, “A Survey of Face Recognition Algorithms and Testing Results”, Proceeding of IEEE, 1998.

[17] Tsuhan Chen, Yufeng Jessie Hsu, Xiaoming Liu and Wende Zhang, “Principal Components Analysis and its Variants for Biometric”, International Conference of Image Processing, Feb, 2002.

[18] Gian Luca Marcialis and Fabio Roli, “Fusion of LDA and PCA for Face Recognition”, Department of Electrical and Electronic Engineering - University of Cagliari, Piazza d’Armi, Italy.

[19] Wenyi Zhao, Arvindh Krishnaswamy, Rama Chellappa, Denial L, Swets and John Weng,“Discriminant Analysis of Principal Components for Face Recognition”, Proceedings of International Conference on Automatic Face and Gesture Recognition, 1998, pp. 336-341.

[20] S. Balakrishnama, A. Ganapathiraju, “Linear Discriminant Analysis: A Brief Tutorial”, Institute for Signal and Information Processing, Mississippi State University, USA.

[21] G.J. Edwards, T.F. Cootes and C.J. Taylor, “Face Recognition using Active Appearance Models”, in Proc. European Conference on Computer Vision 1998, vol. 2, pp- 581-695, Springer 1998.

[22] W. Zhao, R. Chellapa, “A. Rosenfeld and P.J. Philips, Face Recognition: A Literature Survey”, UMD CFAR Technical Report CAR-TR-948, 2000.

[23] Gupta, H., R. Chowdhury, A. K., Chellappa, R., Contour-based 3D Face Modeling from a Monocular Video, British Machine Vision Conference 2004.

[24] Stan Z. Li and Anil K. Jain, Handbook of Face recognition, Springer 2005.

[25] Okabe, A.; Boots, B.; and Sugihara, K. Spatial Tessellations: Concepts and Applications of Voronoi Diagrams. New York: Wiley, 1992.

[26] Kanade, T., Cohn, J.F., Tian, Y. (2000). Comprehensive database for facial expression analysis, Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG’00), Grenoble, France, 46-53.

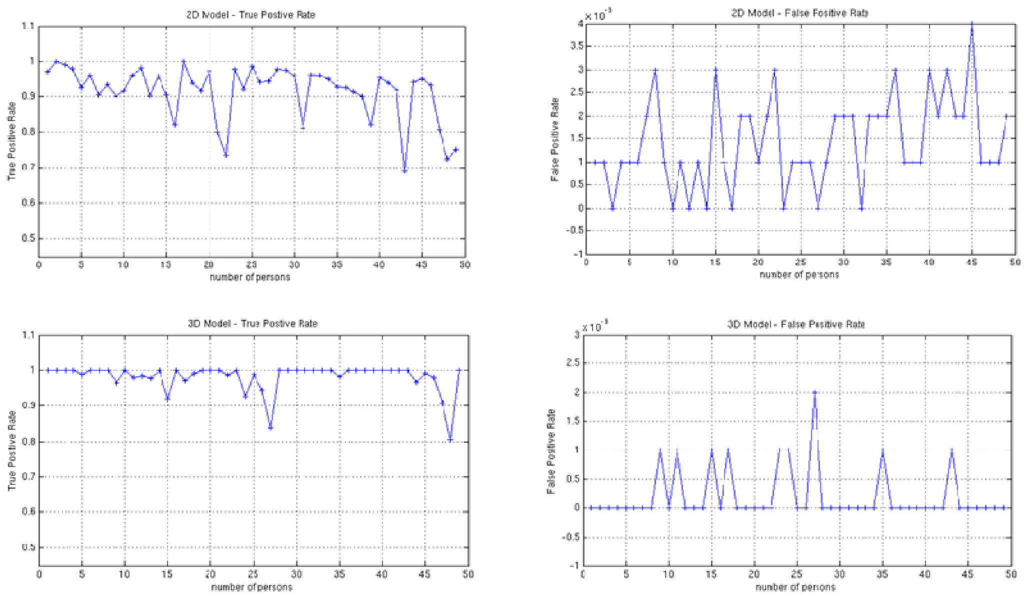

Figure 7 Classification results in the form of true positives and false positive for 2D and 3D model