Support Vector Machine Components - Based Face Recognition Technique using 3D Morphable Modeling Method

Mukundhan Srinivasan and Nivas Ravichandran

***Abstract*—We present a novel solution towards the problem of pose and illumination variation of face detection (FD) and face recognition (FR). In this paper, two advanced method are used to provide pose and illumination invariant FR. The 3D morphable model is implemented to generate 3D face images from our very own training database. This process requires a set of three input face images with varying pose and illumination constraints. The resulting 3D model is then used to train the Support Vector Machine (SVM) component-based FR. SVM component-based 3D model has promising results yielding close to 92.6% accuracy when tested on three training face images of each subject under test.**

***Index Terms*—Face recognition, 3D model, support vector machine (SVM), component-based recognition.**

I. Introduction

Over the past two decades, numerous FR researches, papers and studies have been carried out in the field of Computer Vision (CV) [1]. There are several real-time applications like biometrics, surveillance, security access, Human Computer Interaction (HCI), robotic vision that demand a robust, accurate and simply trainable face recognition systems. The availability of cheap and yet so competent systems have led to rapid development and commercialization of FR systems. Despite these achievements, however, external parameters such as pose variation, discrepancy in illumination, facial expression, gender recognition and twins' recognition are still a paradox. Among many developed method, the component-based approach have shown appreciable results in various recognition tests such as FD [2], [3] and FR [4]-[7].

In this paper, a system is described in which 3D morphable face models are generated. These generated face modes are inputs to the training stage of the component based classifier. An image set of three images of a test subject are used to synthesis the 3D morphable model. Once the 3D face models of all the subjects in the training database are computed, we generate arbitrary synthetic face images under varying pose and illumination to train the component-based recognition system.

The outline of this paper is as follows: Section II revels some related work in this area, Section III describes the methodology, Section IV results are discussed and Section V while 6 deals with the conclusion and future work.

II. Related Work

In [1], a SVM based FR system is discussed in which the face images are decomposed into a set of components that are either related or interconnected by a mathematical model. This approach from [1] is compared with the neural network face recognition system and the results favor the SVM component model with higher percentage of accuracy. Different poses of the major part of the face that is, the head leads to change in position of all other facial components. The variation is compensated by the geometric model implemented. In experiments pertaining to our paper show results that, the SVM model consistently surpasses the other models in terms of efficiency and accuracy.

III. Methodology

A. Generating Face Models in 3D

The primary objective is to generate a robust 3D Face Model that can be propagated into a set of SVM classifiers that perform the face recognition. In order o achieve this; we generate a model based on three training images of each test subject's face image. The three views used here are: (1) Frontal or Straight pose of the face image (2) The other two face images are profile images with rotations in opposite directions with respect to the frontal view. The three views are explicitly shown in Fig. 1.

Fig. 1. (a) Frontal view (b&c) semi profile view

The fundamental concept underlying the 3D morphable models is that, from a database of relatively large number of 3D face models any generic random face can be generated by morphing attributes like facial expression, eye movement, facial hair, movement of eyebrows, epicanthic fold, position of ear and many more parameter. Our database of 97 students was initially generated by capturing the three different input

Manuscript received November 7, 2012; revised January 28, 2013.

The authors are with the Department of Electronics & Communication Engineering, College of Engineering, Alpha Group of Institutions, Chennai, TN India (e-mail: mukundhan@ ieee.org, nivasravichandran@ieee.org).

DOI: 10.7763/IJFCC.2013.V2.218

face images (i.e. frontal view & left, right profile views). Once the frontal, left and right profile images are loaded the feature points are set to generate a 3D model. Through processes like flow computation a relation between the 3D model and the input face images are established.

In the 3D modeling of face images, we can obtain a dataset of three results; the first one being the texture-removed 3D image, secondly a tween model and at last a mesh model. Fig. 2 describes the results clearly. All the three results are very useful to compute many parameters for FR. Example, a wire/mesh model can be used to calculate metrics like the inter-eye distance, or calculate the area of the triangle formed by the feature points of the two eyes and the nose. The tween model can be used in eyeball tracking and etc.

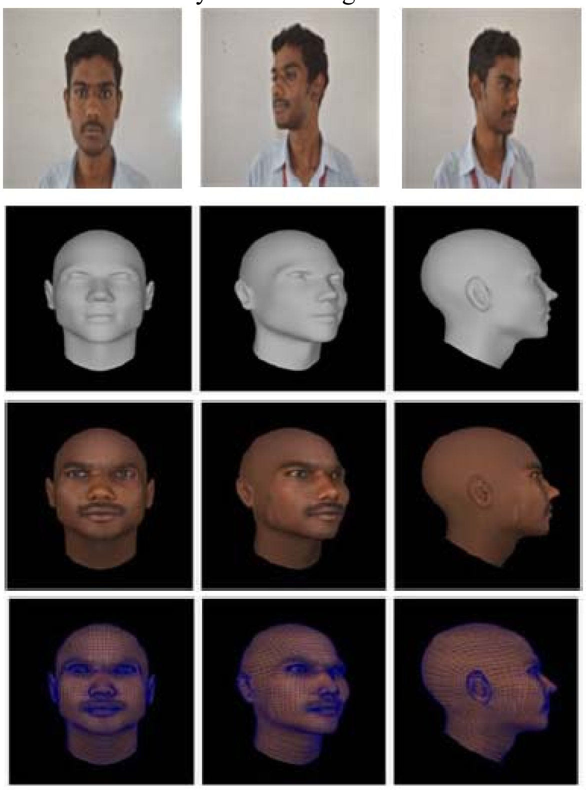

Fig. 2. 3D generated face model

Using the 3D models, synthetic images such as the ones in Fig. 2 can easily be created by rendering the models. The 3D morphable model also provides the full 3D correspondence between the head models, which allows for automatic extraction of facial components.

B. Component based Approach Using SVM

There are many other approaches that are highly sensitive to image variation caused by rotation in facial angle. The SVM component-based model avoids such issues by recognizing independent components of the face image. For small rotations, the changes in the components are relatively small compared to the changes in the whole face pattern. Changes in the 2-D locations of the components due to pose changes are accounted for by a learned, flexible face model.

C. Face Detection

To obtain stable results, a two-level, component-based face detector which is described in detail in [8]. In this following section we give a brief overview of the system. The principles of the component-based detection system are illustrated in Fig. 3. On the first level, component classifiers independently detected facial components. On the second level, a mathematically configuration classifier performs the final face detection by combining the results of the component-based classifiers.

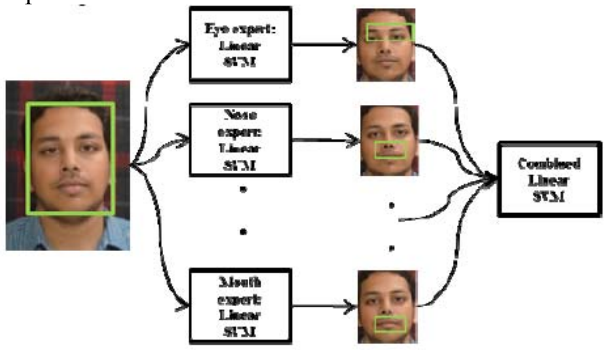

Fig. 3. Structure of the proposed system

Given a 32 ×32 window, the maximum continuous outputs of the component classifiers within rectangular search regions around the expected positions of the components were used as inputs to the mathematical configuration classifier. The search regions have been calculated from the mean and standard deviation of the components' locations in the training images. The 14 facial components used in the detection system are shown in Fig. 4. The shapes and positions of the components have been automatically determined from the training data in order to provide maximum discrimination between face and non-face images.

Fig. 4. (a, b) Fourteen facial components used for FD

Fig. 5. Combined components for recognition.

We trained 14 linear SVMs on the component data and applied them to the whole training set in order to generate the training data for the geometrical classifier.

D. Face Recognition

To train the face recognizer we first ran the component-based detector over each image in the training set and extracted the components. From the 14 original components we kept 9 for face recognition, removing those that either contained lesser features vectors (e.g., area around the cheeks) or strongly overlapped with other components. The 9 selected components are shown in Fig. 4b. Examples of the component-based face detector applied to images of the training set are shown in figure 5. To generate the input to our face recognition classifier we normalized each of the components in size and combined their gray values into a single feature vector considering it as the tenth component

The normalization used here is the linear normalization of grey scale facial images is performed according to the formula:

$$I_N = (I - Min)\frac{Max' - Min'}{Max - Min} + Min' \quad (1)$$

The normalization of inner face image is included for better recognition results.

IV. Results

A test set was created by taking facial images of the four people in the database. The subjects were asked to rotate their faces in depth and the lighting conditions were changed by moving a light source around the subject. The test set consisted of three images of each person under various pose and illumination conditions.

The component-based face recognition system was compared to a NN-based face recognition system; both systems were trained and tested on the same images.

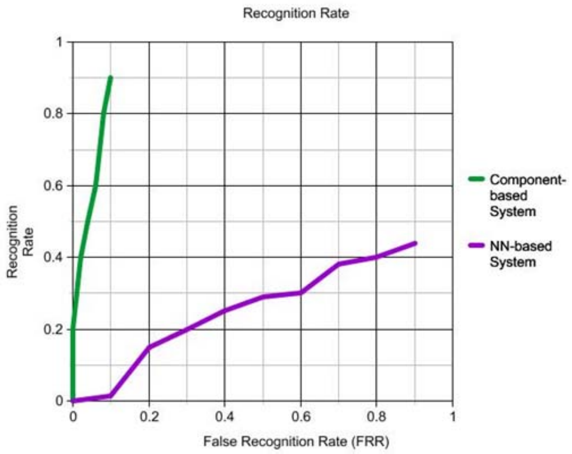

Fig. 6. ROC plot of the FRR vs. recognition rate for both the component and NN- based system

In contrast to the component-based classifiers, the input vector to the whole face detector and recognizer consisted of the linearly normalized grey values from the entire 32×32 facial section. The resulting ROC curves of Neural Network and component- based recognition on the test set can be seen in Fig. 6. The component-based system achieved recognition of 92%, which is approximately some 30% above the recognition rate of the NN system.

This large discrepancy in results can be attributed to two main factors: First, the components of a face vary less under rotation than the whole face pattern, explaining why the component-based recognition is more robust against pose changes. Second, in contrast to the training data, the backgrounds in the test images were non-uniform. Component-based recognition only used face parts as input features for the classifier.

V. Conclusion

In this paper, we propose a 3D model which is based on a component-based approach using the Support Vector Machine (SVM) model. The same test dataset is used for a Neural Network system which has a lesser rate of recognition.

The component-based approach yields about 92% efficiency which is about 30% more than the NN system. The component approach gives better results as the variation do not have much implication on the component learning.

VI. Future Work

We intend to test on a much larger and standard database to check if these results concur. We would also like to inspect the efficiency of HOG over linear normalization techniques to augment the efficiency

Acknowledgment

The authors would like to thanks all those involved in helping us write this paper. We would also like to thank the subject who helped us in compiling a database for testing.

References

[1] R. Chellapa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," in *Proc. IEEE*, vol. 83, no. 5, 1995, pp. 705–741.

[2] H. Schneiderman and T. Kanade, "A statistical method for 3D object detection applied to faces and cars," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 746-751, 2000.

[3] B. Heisele, T. Serre, M. Pontil, and T. Poggio, "Component-based face detection," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition,* vol. 1, pp. 657-662, Hawaii, 2001.

[4] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, 1993.

[5] L. Wiskott, "Labeled Graphs and Dynamic Link Matching for Face Recognition and Scene Analysis," PhD thesis, Ruhr-UniversitÄat Bochum, Bochum, Germany, 1995.

[6] A.V. Neifan and M. H. Hayes, "An embedded HMM-based approach for face detection and recognition," in *Proc. IEEE International Conference on Acoustics*, Speech, and Signal Processing, vol. 6, pp. 3553-3556, 1999.

[7] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: global versus component-based approach," in *Proc. 8th International Conference on Computer Vision*, vol. 2, pp. 688-694, Vancouver, 2001.

[8] B. Heisele, T. Serre, M. Pontil, T. Vetter, and T. Poggio, "Categorization by learning and combining object parts," *Neural Information Processing Systems (NIPS)*, Vancouver, 2001.

[9] A. Lanitis, C. Taylor, and T. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 19, no. 7, 1997, pp. 743–756.

[10] K. Jonsson, J. Matas, J. Kittler, and Y. Li, "Learning support vectors for face verification and recognition," in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 208–213.

[11] T. Poggio and S. Edelman, "A network that learns to recognize 3-D objects," *Nature*, vol. 343, 1990, pp. 163–266.

[12] A. Pentland, B. Mogghadam, and T. Starner, "View-based and modular eigenspaces for face recognition," Technical Report, vol. 245, MIT Media Laboratory, Cambridge, 1994.

[13] H. Schneiderman and T. Kanade, "A statistical method for 3d object detection applied to faces and cars," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2000, pp. 746–751.

[14] R. Brunelli and T. Poggio, "Face recognition: features versus templates," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 15, no. 10, 1993, pp. 1042–1052.

[15] D. J. Beymer, "Face recognition under varying pose," A.I. Memo 1461, Center for Biological and Computational Learning, M.I.T., Cambridge, MA, 1993.

[16] L. Wiskott, J. M. Fellous, N. Kr€uger, and C. V. D. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 19, no. 7, 1997, pp. 775–779.

[17] A. Nefian and M. Hayes, "An embedded hmm-based approach for face detection and recognition," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 6, 1999, pp. 3553–3556.

[18] A. M. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 24, no. 6, 2002, pp. 748–763.

[19] V. Vapnik, *Statistical Learning Theory*, John Wiley and Sons, New York, 1998.

[20] R. Rifkin, "Everything old is new again: a fresh look at historical approaches in machine learning," Ph.D. thesis, M.I.T., 2002.

[21] C. Cortes and V. Vapnik, "Support vector networks," *Mach. Learning*, vol. 20, 1995, pp. 1–25.

[22] B. Sch€olkopf, C. Burges, and V. Vapnik, "Extracting support data for a given task," in: U. Fayyad, R. Uthurusamy (Eds.), *Proc. First Int. Conf. on Knowledge Discovery and Data Mining*, AAAI Press, Menlo Park, CA, 1995.

[23] M. Pontil and A. Verri, "Support vector machines for 3-d object recognition," *IEEE Trans. Pattern Anal. Mach. Intell*, 1998, pp. 637–646.

[24] G. Guodong, S. Li, and C. Kapluk, "Face recognition by support vector machines," in: *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2000, pp. 196–201.

[25] J. Platt, N. Cristianini, and J. S. Taylor, "Large margin dags for multiclass classification," *Adv. Neural Inform. Process. Systems*.

[26] C. Nakajima, M. Pontil, and T. Poggio, "People recognition and pose estimation in image sequences," in *Proc. IEEE-INNS-ENNS International Joint Conf. on Neural Networks*, 2000, vol. 4, pp. 4189–4195.

[27] B. Heisele, T. Poggio, and M. Pontil, "Face detection in still gray images," AI Memo 1687, Center for Biological and Computational Learning, MIT, Cambridge, MA, 2000.

[28] K. K. Sung, "Learning and example selection for object and pattern recognition," Ph.D. thesis, MIT, Artificial Intelligence Laboratory and Center for Biological and Computational Learning, Cambridge, MA, 1996.

[29] V. Blanz and T. Vetter, "A morphable model for synthesis of 3D faces," in: *Comput. Graphics Proc. SIGGRAPH*, Los Angeles, 1999, pp. 187–194.

Mukundhan Srinivasan is pursuing Bachelor of Engineering with major in Electronics and Communication Engineering from Alpha College of Engineering, Anna University, Chennai, TN India. Mukundhan also obtained a Bachelor of Science (associate) degree in Computer Science from NIIT, Chennai, TN India.

He has published two papers in internal journals and three papers in international conferences. His current research interests are Image Processing, Pattern Recognition, Machine Learning, Computer Vision, and application of Robotics. Mr. Srinivasan is a student member in IEEE, ACM, ISTE, ACCS, and CSI

Nivas Ravichandran is with the Electronics and Communication Engineering, Jeppiaar Engineering College, Anna University, Chennai, Tamil Nadu, India, 2013. He is the Hub Leader - Chennai, IEEE Madras Student Network. He is very passionate about Image Processing, Signal Processing, Pattern Recognition and Computer Vision. Mr. Ravichandran, IEEE is an active student volunteer. He has contributed and volunteered a lot towards IEEE and Student Activities.