Robust Technique of Age and Expression Invariant Feature Extraction

Prof. Md Yousuf Ahmed[1] Prof. Feroza D Mirajkar[2]
[1,2]AssistantProfessor
[1,2]Department of Electronics and Communication Engineering
[1,2]KBN College of Engineering Kalaburagi

Abstract— The development of a robust technique to extract the facial features that is invariant to extrinsic (head, pose, lightning conditions) and intrinsic (aging, facial expression) source of variability is one of the greatest challenges of current research in this field. Designing such a robust technique can overcome the problems of missing child, multiple enrollments where subjects purposely try to hide their identity. In this paper, a robust technique is developed which extracts the features invariant to aging and expression variation. The two local descriptors SIFT and MLBP are used to extract the local features of the image. The random sampling techniques, random subspace and bagging are used to construct the multiple classifiers of lower dimensionality. The final decision is made using the simple score sum based rule.

Key words: Face recognition, age invariance, local feature representation, SIFT, MLBP, Multi-Feature Discriminant Analysis, Discriminative Model

I. INTRODUCTION

Faces are probably the most common biometric identifier that is used to recognize people. Among the various biometric security systems based on finger print, iris, voice or speech, signature, etc., face recognition seems to be the most universal, non-intrusive, and accessible system. It can be used efficiently for mass scanning which is quite difficult in case of other biometrics, and also increases user-friendliness in human-computer interaction. Automatic recognition of people is a challenging problem which has received much attention during the recent years due to its potential applications in different fields such as law enforcement, security applications or video indexing.

The research in age related face recognition has increased since 2002. However, related published work is still very limited in both quantity and quality. Most of the approaches mainly focused on either face age progression simulation [5] or age estimation [6]. The approach showed that, due to the temporal property of age progression, face images with aging features may display some sequential patterns with low-dimensional distributions. It also demonstrated that such aging patterns can be effectively extracted from a Discriminant subspace learning algorithm and visualized as distinct manifold structures. Through the manifold method of analysis on face images, the dimensionality redundancy of the original image space can be significantly reduced with subspace learning.

A. *Motivation:*

Aging is an unavoidable natural process during the lifespan of a person and it also affects the performance of a face recognition system. The major goal of this research is to design a face recognition system that is robust to expression and aging variation. The growth and development of face from birth to adulthood, the greatest change is the craniofacial growth. In the younger age groups, there are large texture changes and minor shape changes in older age groups. The local features inherently possess spatial locality and orientation selectivity, which allows the local feature representations to be robust to aging, illumination, and expression variations.

B. *Issues and Challenges:*

1) *Expression Variation:*

In face recognition, the problem of expression-invariant representation is of particular importance as most of today's face recognition systems are sensitive to facial expression variations. The development of robust face recognition algorithm insensitive to facial expression is one of the greatest challenges of current research in this field. One possibility is to use only regions of the face the least susceptible to variations due to facial expression [1]. For example, one can remove the mouth region as it varies greatly with expression. Yet, practice shows that there is no large subset of the face that is perfectly rigid across a broad range of expressions [2]. Another possibility is to add different expressions of each subject into the gallery. The main limitation is the large number of instances of each subject due to the richness of all the possible expression. Moreover, it is practically impossible to capture the subtle individual differences of expressions in different subjects. Figure 1.1 shows the different expressions of the subject image taken from ORL database.

Fig. 1.1: Expressions of a subject in ORL database

2) Age variation:

Aging related changes on the face appear in a number of different ways: i) wrinkles and speckles, ii) weight loss and gain, and iii) change in shape of face primitives (e.g., sagged eyes, cheeks, or mouth). All these aging related variations degrade face recognition performance. These variations could be learned and artificially introduced or removed in a face image to improve face recognition performance. Even though it is possible to update the template images as the subject ages, template updating is not always possible in cases of i) missing child, ii) screening, and iii) multiple enrollment problems where subjects are either not available or purposely trying to hide their identity. Multiple enrollment detection for issuing government documents such as driver licenses and passports is a major problem that various government and law enforcement agencies face in the facial databases that they maintain[3]. Therefore, facial aging has become an important research problem in face recognition.

II. LITERATURE SURVEY

One of the first papers related to face aging from digital images belongs to Kwon et al. (1993) [7]. In this method, 47 high resolution images of a face are used to classify the images into one of three age groups: babies, young adults or senior ones. This approach was based on geometric ratios of key face features and wrinkle analysis. However, the database was limited to 47 high resolution photos and such ideal image quality would be very difficult to acquire in practical applications.

Gandhi [8] designed a support vector machine based age estimation technique and extended the image based surface detail transfer approach to simulate aging effects on faces. The approach discussed the training of a Support Vector Regression Machine to perform age prediction on labeled frontal images. The images were collected from the Internet and compensated for pose, illumination and expression by our normalizing scheme.

Sethuram et al. [5] also built a high accuracy face aging model based on AAMs, support vector machines (SVMs) and Monte-Carlo simulation. The approach conducted two experiments. In experiment 1, it has been shown that the accuracy of face recognition goes down when probe faces age. In experiment 2, the probe faces are first artificially aged to the same age of the gallery by using the face aging model. Then, the face recognition algorithm is applied to get a higher accuracy than the ones in experiment1.

Meanwhile Geng et al. [9] introduced an Aging pattern Subspace (AGES) on the assumption that similar faces age in similar ways for all individuals. Their basic idea is to model the aging pattern, which is defined as a sequence of a particular individual's face images sorted in time order, by constructing a representative subspace. The proper aging pattern for a previously unseen face image is determined by the projection in the subspace that can reconstruct the face image with a minimum reconstruction error, while the position of the face image in that aging pattern will then indicate its age.

Park et al [9] presented an approach to age invariant face recognition by using a 3D generative model for face aging. In their method, in order to compensate for the age effect, probe face images are first transformed to the same age as the gallery image by using the trained 3D aging model. Then, FaceVACS, a commercial face recognition engine, was used to evaluate the identification results.

Zhao et al. [3] have proposed to combine the local binary pattern representation with Kernel Fisher Discriminant Analysis in order to improve the face verification performance of LBP and they also mentioned that the performance of combining LBP histogram (LBPH) with Linear Discriminant Analysis method is worse than the LBP histogram itself.

Shan et al. showed that LBPH with the LDA method outperforms LBPH itself. The difference between these two systems is the use of similarity measure in which Zhao's measure is a Euclidean metric and ours is a normalized correlation. As Kittler have shown that the normalized correlation can achieve better performance in the LDA space.

III. PROPOSED SYSTEM

A. Introduction:

This paper implements the learning algorithm that has the capability to address the aging variations and also handle the other intra-user variations (e.g., pose, illumination, expression). Compared to other methods in literature survey, this method differs in both feature representation and classification. Compared to the global appearance features, the local feature representations are robust to aging, illumination and expression variation. This model consists of two components: densely sampled local feature description (SIFT & MLBP) [10] and Multi-Feature Discriminant Analysis (MFDA).

In this approach a patch-based local feature representation scheme is used. Since it is difficult to characterize the entire face image by a single image descriptor, hence the input face image is divided into a set of overlapping patches with each patch represented by an appropriate image descriptor. In order to ensure local consistency, a 50% overlap between the adjacent patches is used. Since, both SIFT-based local features and MLBP-based local features span a high dimensional feature space, a dimensionality reduction algorithm Multi-Feature Discriminant Analysis (MFDA) is used as a classifier. MFDA is an extension and improvement over LDA using multiple features sampled at two different feature and sample space. LDA based classifiers are constructed and then combined to form a robust algorithm via a fusion rule.

B. Flowchart:

The figure 5.1 shows the flow diagram for the proposed approach. The steps involved in the framework are, Data acquisition, preprocessing of the given input image, applying local descriptors to the patches, reducing the dimensions of the feature space in order to boost the performance of the system and finally by using the simple score-sum based rule.

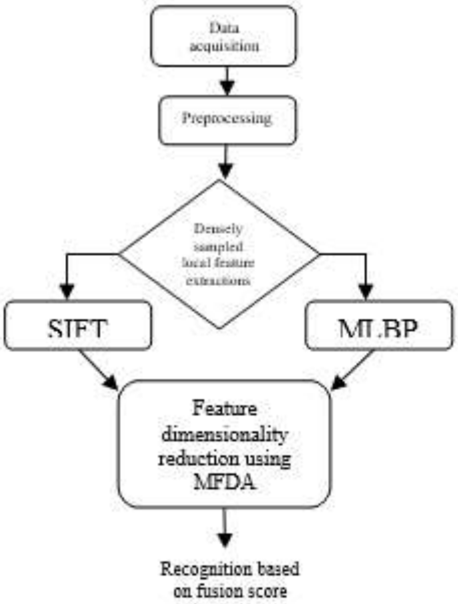


Fig 3.1: Design Steps

Given a face image X x Y, it is divided into r x r patches each overlapping by s pixels. The number of horizontal and vertical patches obtained is,

$$H = \frac{X - s}{s} + 1$$

$$V = \frac{Y - r}{s} + 1$$

For each of these H x V patches a d- dimensional feature vector is computed. These image features are then concatenated to form H x V x d dimensional feature vector for a given face image.

C. Face Database:

The face images used for the experimentation are obtained from FGNET aging and ORL database.

The FG-NET Aging Database [4] is widely used for research of age related facial image analysis. The database contains 1002 images from 82 subjects, over large age ranges. Consequently, there is an average of 12 images per subject in the FG-NET aging database. In order to verify the effective of the proposed algorithm, all the images of the database are chosen and divided into training and the test data separately.

The figure 3.2 shows the example images from the FG-Net aging database.

DATA BASE	#sub	Male	Female	Total # Image	Age Range
FG-NET	82	48	34	1002	0-69
ORL	40	36	4	400	18-81

Table 1:

Fig. 3.2: Example images of a subject in FG-NET aging database

The ORL database contains a set of face images of 40 distinct subjects, with 10 images per subject. The age of the subjects ranges from 18 to 81, with the majority of the subjects being aged between 20 and 35. There are 4 female and 36 male subjects. Some subjects are captured with and without glasses. The figure 3.3 shows example face images of the ORL database.

Fig. 3.3: Example images from ORL database

D. Preprocessing:

The face images acquired from the database are normalized to 150 x 200 pixels and then divided into either 88 overlapping patches (for a patch size of 32 x 32) or 408 overlapping patches (for a patch size of 16 x 16) as shown in the figure 3.2. Each patch is represented by a 128-dimensional SIFT feature vector and 236-dimensional MLBP feature vector.

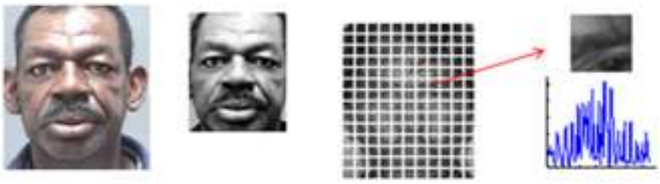


Fig. 3.2: Formation of overlapping patches

E. Local Feature Descriptors:

In the proposed system, the SIFT feature descriptors are sampled densely from the entire facial image by placing a rectangular grid on the face. This allows the discriminatory information in the form of the distribution of the edge direction in the face.

The Multi-scale Local Binary Pattern (MLBP)) is a simple but powerful texture representation of a face image. This multi-resolution representation based LBP can be obtained by varying the sample radius, R, and combining the LBP images. In the proposed approach the LBP descriptors are computed at four different radii {1, 3, 5, 7}. The studies in local binary pattern have suggested for texture classification, the accuracy of multi-scale LBP is better than that of the single scale local binary pattern method.

F. Multi-Feature Discriminant Analysis:

MFDA is an extension and improvement of LDA. The Discriminant analysis is used to reduce the dimensionality of the resulting feature vector comprising of both SIFT and MLBP. In order to better address the dimensionality problem, both random subspace and bagging schemes have been utilized. First, in order to reduce the feature dimensionality, we apply the random subspace technique to sample the feature space to generate multiple subspaces with lower dimensionalities.

Second, in order to utilize the classification boundary information where LDA fails, the specific sample pairs are selected from different classes to better estimate the between-class scatter matrix and the Discriminant subspace. By combing the random subspace and bagging techniques, a random sampling based classification framework, called MFDA is implemented as shown in fig. 3.4.

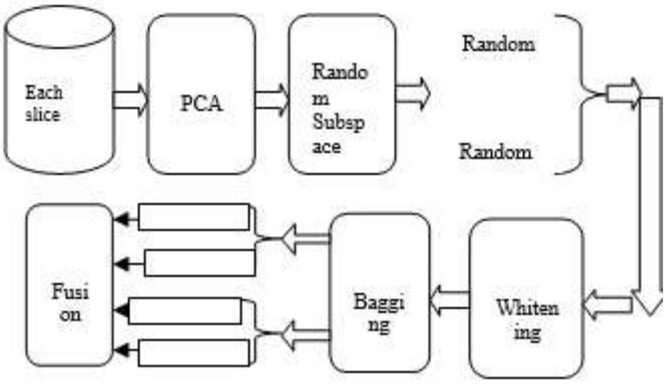


Fig. 3.4: Entire framework

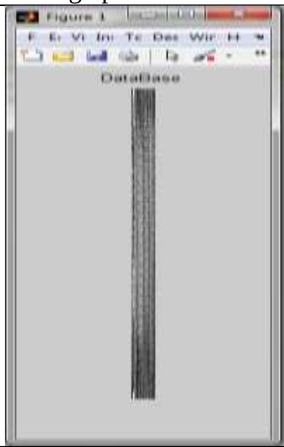
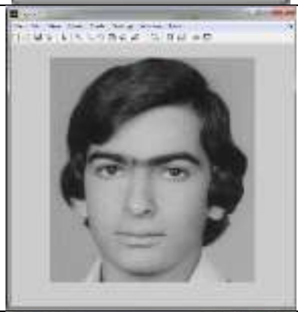
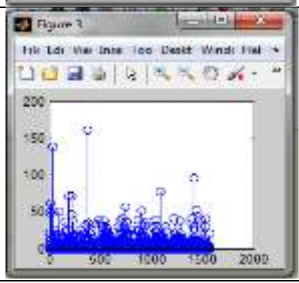
IV. RESULTS AND DISCUSSION

This section discuss about the databases used in the experiment and the results at each step of the execution.

Since the FG-NET database has large variation in expression, pose, and illumination among the images. All the images of the database have been utilized and each image has been preprocessed before extracting its features. The following steps show the preprocessing operation followed.

1) Rotating the face image so that it is aligned with the vertical face orientation.
2) Scaling the face image so that the distance between the two eyes is the same for all the face images.
3) Cropping the image in order to remove the background and the hair region.
4) Applying histogram equalization.

The table 4.1 shows the face image before and after preprocessing operation.

Database	
Input image before preprocess	
After resizing	
After histogram equalization	
SIFT features	

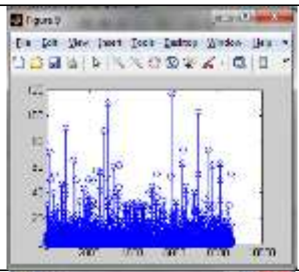
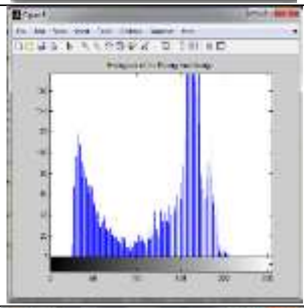
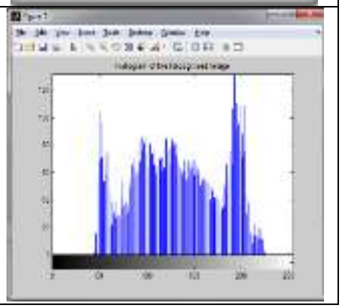
MLBP features	
Final Matching Test 1:	
Histogram of recognized image	
Test 2:	
Histogram of recognized image	

Table 4.1: Experiment on FG-NET database

In the testing stage, a total of 70 slices are obtained for each test image. The trained subspace classifiers are used to obtain the classification outputs of these slices. These outputs are normalized using min-max score normalization scheme. To make the final decision, a simple score-sum based fusion rule is utilized. As shown in table 4.1, the input image is at the age of 14 years and the recognized image is at the age of 30 years of the same subject. It also shows a recognized image at an age of 30years through an input image at an age of 5 years of the same subject.

The figure 4.3 shows the recognition performance of the proposed system. Experimenting on the FG-NET was very challenging task, because this aging database contains much larger age gaps. The largest age gap is 45 years.

The second reason is that the number of subjects in the FG-NET database is very limited, which makes learning very difficult. The proposed method was applied to all the images of the database without grouping them into specific age range. For the first round, the 80% images were randomly selected as the training and the remaining as the testing dataset. Here, we achieved a recognition accuracy of 46.5%.

In the second experiment, all the images of the FG-NET database were selected for training and testing. An increase of 2.5% in the recognition accuracy was observed.

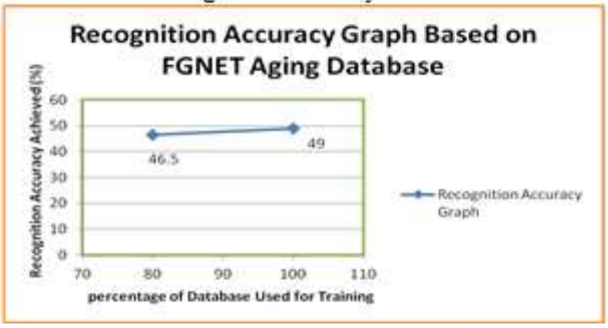
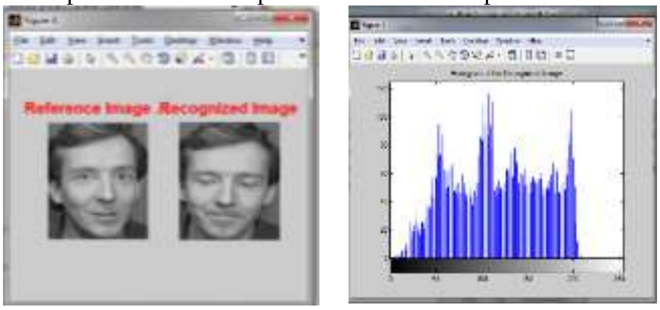


Fig. 4.3: Recognition accuracy achieved for the FG-NET database with two sets of training data i.e. 80% and 100% of the database.

The figure 4.4 shows the final output of the proposed system experimented on ORL database. It can be observed that, the proposed system works well in the presence of the expression and also the pose variation.



(a) (b)

Fig. 4.4: Expression variation (a) Input and the recognized image in the presence of expression variation, (b) Histogram of the recognized image.

The figure 4.5 shows the performance of the system on the ORL database. The robustness of the project has been tested in the three rounds of experiments.

Experiment I: In this experiment, 60% of the total images from the ORL database had kept for training and the remaining 40% had used for testing. The recognition accuracy achieved was 92.5%, with an error rate of 7.5%.

Experiment II: Here, 80% of total images in the ORL database were used for training and the remaining 20% for testing. The performance of the system improved to 95.75% from 92.5%. Hence it can be said that the recognition performance will be improved if the size of the training data increases.

Experiment III: In this experiment all the images of the ORL database was utilized for training. The recognition accuracy now achieved was 100%.

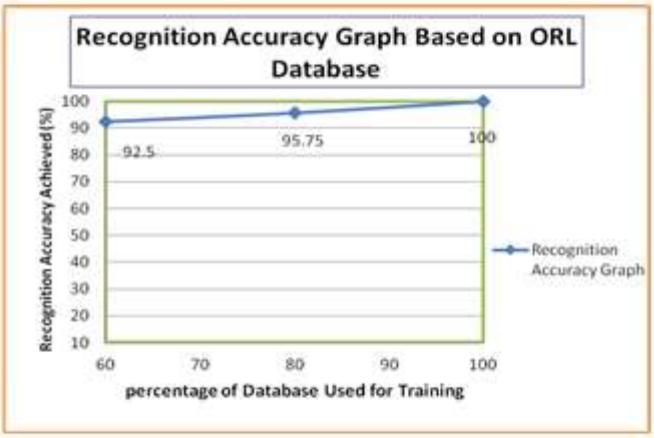


Fig. 4.5: Recognition accuracy achieved for the ORL database with three sets of training data i.e. 60%, 80% and 100% of the database.

V. CONCLUSION

In this paper, the problem of age invariant and expression invariant face recognition is considered. The proposed technique for extracting the features of the face that are invariant to changes in expression and age can be used to develop a robust face recognition system. This technique addresses the face aging problem that does not depend on the generating the aging model. The SIFT and MLBP features extracted by dividing the face image into a number of overlapping patches can withstand the changes in the persons age and also facial expressions. The proposed technique can also withstand the large variations in illumination and pose in the training set of the subjects. Here each image is represented with a patch-based local feature representation scheme, since it is difficult to characterize the entire face image by a single image descriptor. For matching the set of large number of SIFT and MLBP local features effectively and efficiently, a multi-feature Discriminant analysis (MFDA) algorithm is developed for dimensionality reduction. In MFDA, local descriptors are combined to construct a robust decision rule by a random subspace fusion model. Being a discriminative approach, the proposed method requires no prior age knowledge and does not rely on age estimation. The robustness of the system was experimented successfully on the two public domain databases i.e. ORL and the FG-NET aging database. Hence it can be concluded that the proposed technique is able to address the facial aging and the expression variation problem that occur in the face recognition system with improvement over the 3D aging model.

REFERENCES

[1] C. Chua, F. Han, and Y. Ho, "3D human face recognition using point signature," in International Conference on Automatic Face and Gesture Recognition, 2000, pp. 233–238.
[2] K. Chang, K. Bowyer, and P. Flynn, "Multiple nose regions matching for 3D face recognition under varying facial expression," preprint.
[3] Jiali Zhao, Haitao Wang, Haibing Ren, and Seok-Cheol Kee. Lbp. Discriminative analysis for face verification. In Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on, volume 3, pages 167–167, 20-26 June 2005.
[4] FG-NET Aging Database, http://www.fgnet.rsunit.com/.
[5] A.Sethuram, E.Patterson, K.Ricanek, and A.Rawls. "Improvements and Performance Evaluation Concerning Synthetic Age Progression and Face Recognition Affected by Adult Aging" in the 3rd IEEE Int'l Conf. on Biometrics, 2009.
[6] Y.Fu and T.Huang. "Human age estimation with regression on discriminative aging manifold", in IEEE Trans. on Multimedia, volume 10, pages 578–584, 2008.
[7] Y.Kwon and N.Lobo. Locating facial features for age classification, in the Int'l Society for Optical Engineering, volume 2055, pages 62–72, 1993.
[8] M.Gandhi. "A method for automatic synthesis of aged human facial images", Master's thesis, McGill University, September 2004.
[9] U. Park, Y. Tong, and A. K. Jain, "Age Invariant Face Recognition, " IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 5, pp. 947-954, May 2010.
[10] Lowe D.G.: Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision.