Age Invariant Face Recognition Using Graph Matching

Gayathri Mahalingam and Chandra Kambhamettu
Video/Image Modeling and Synthesis (VIMS) Laboratory
Department of Computer Science, University of Delaware, Newark, DE
{mahaling, chandra}@cis.udel.edu

***Abstract*— In this paper, we present a graph based face representation for efficient age invariant face recognition. The graph contains information on the appearance and geometry of facial feature points. An age model is learned for each individual and a graph space is built using the set of feature descriptors extracted from each face image. A two-stage method for matching is developed, where the first stage involves a Maximum a Posteriori solution based on PCA factorization to efficiently prune the search space and select very few candidate model sets. A simple deterministic algorithm which exploits the topology of the graphs is used for matching in the second stage. The experimental results on the FGnet database show that the proposed method is robust to age variations and provides better performance than existing techniques.**

I. INTRODUCTION

In recent years, face recognition across aging has gained attention from computer vision researchers and has been looked into using age estimation and aging models. Facial aging is attributed by changes in facial features, shape and texture and other biological factors like weight loss/gain, facial hair, etc. A detailed survey of contributions from both psychologists and computer scientists on facial aging is given in [27], [28].

Ling *et al.* [16], [17] studied how age differences affect the face recognition performance in a real passport photo verification task. The authors proposed a non-generative approach in which they defined a face operator, derived based on the image gradient orientations derived from multiple resolutions and then used support vector machines to perform face verification across age progression. Kwon and Lobo [21] proposed a theory and practical computations based on the anthropometry of the face and the density of the wrinkles to classify age from facial images. Lanitis *et al.* [5]-[7] proposed an approach which adopts active appearance model (AAM) technique for age estimation. They developed a model which combines the shape and intensity information to represent the face images. Geng *et al.* [14], [15] learned a subspace of aging pattern based on the assumption that similar faces age in similar ways.

Guo *et al.* [19] proposed an age manifold learning scheme for age estimation. The face aging features are extracted and a locally adjusted robust regressor is designed for learning and prediction of human ages. Fu and Huang [29] also proposed a manifold learning technique in which a low-dimensional manifold is learnt from a set of age-separated face images, and linear and quadratic regression functions are applied on the low-dimensional feature vectors from the respective manifolds in face age estimation. Ramanathan and Chellapa [10], [11] proposed a craniofacial growth model that characterized the shape variations in human faces across age variations. Authors observed that the growth parameter k for different facial features across age can be adapted in the model to characterize the facial growth. Ramanathan and Chellapa [12] also proposed a two step approach for modeling aging in adults, which comprised of a shape and texture variation model. The formulation of shape variations is performed by constructing physical models which characterizes the functionalities of the facial muscles. Drygajlo *et al.* [9] demonstrated the application of Q-stack classifier to perform face verification across age progression. Mahalingam and Kambhamettu [13] proposed a probabilistic approach for face verification across aging, in which facial features descriptors are extracted from hierarchical representation of the face images, and are used in a probabilistic model for verification purposes.

Park *et al.* [2], [8] designed an aging simulation technique that learns the aging patterns of shape and the texture based on PCA coefficients. A 3D morphable model is used to model the aging variations from a set of 2D face images. Kristina [22] also proposed an algorithm for face recognition and growth prediction using a 3D morphable face model. Suo *et al.* [3] proposed a dynamic model for simulating the aging process. This model is a grammatical face model augmented with age and hair features. The model represents all the images by a multi-layer And-Or graph which includes the changes in hair style and shape, deformations and aging effects of facial components, and wrinkles appearance at various zones. Udeni and Dharmaratne [23] proposed a face recognition approach where they warped the facial images of different ages using the facial features which were extracted from the images using face anthropometry.

Singh *et al.* [1] used an age transformation algorithm which registers the gallery and probe face images in polar coordinates domain and minimizes the variations in facial features caused due to aging. Wang *et al.* [18] obtained the texture and shape information using PCA and used this in the reconstruction of shape and texture at any particular age. Tiddeman *et al.* [30] proposed a wavelet transformation based age simulation technique to prototype the composite face images. Burt and Perrett [4] proposed an age simulation algorithm using shape and texture information to create composite face images for various age groups. They analyzed and measured the facial cues that are affected by age variations.

Biswas *et al.* [20] proposed a coherency in facial feature drifts in images of different ages and used that to measure the performance of face verification.

In this paper, we present a graph based feature representation of the face images, and build a probabilistic aging model for each individual using Gaussian Mixture Model (GMM) which incorporates both the shape and texture information. A simple graph construction algorithm is presented which uses the feature points of an image as vertices, and their corresponding feature descriptors as labels. Matching is performed in two stages. In the first stage, a Maximum a posteriori solution is computed using the aging model of the individuals to effectively reduce the search space and identify potential individuals for the second stage. In the second stage, a simple deterministic graph matching algorithm that exploits the spatial similarity between the graphs is proposed. In section II, we describe our approach in feature point extraction, feature descriptor extraction, graph construction, aging model, and matching.

II. Face Image Representation

In this section, we describe our approach in representing the face images. In our approach, the face image is represented by a graph which is constructed using the facial feature points as vertices. The vertices are labeled by their corresponding feature descriptors which are extracted using the Local Binary Pattern (LBP) [25], [26]. Every face is distinguished not by the properties of individual features, but by the contextual relative location and comparative appearance of these features. Hence it is important to identify those features that are conceptually common in every face such as eye corners, nose, mouth, etc. In our approach, the feature points are extracted using a modified Local Feature Analysis (LFA) technique [24]. We then define a procedure to construct a graph using the feature points extracted from the image.

A. Feature Point Extraction

A modified statistical Local Feature Analysis (LFA) [24] technique is used to extract a set of feature points from each image. The Local Feature Analysis (LFA) proposed by Penev and Atick [24] constructs kernels, which are used as basis vectors for feature extraction. Given a set of n d-dimensional images $x_1, ..., x_n$, Penev and Atick [24] compute the covariance matrix C, from the zero-mean matrix X of the n vectorized images as follows:

$$C = XX^T \tag{1}$$

The eigenvalues and eigenvectors of the covariance matrix C are computed using PCA and the first k largest eigenvalues, $\lambda_1, \lambda_2, ..., \lambda_k$ and their associated eigenvectors $\psi_1, ..., \psi_k$ are used to define the kernel K,

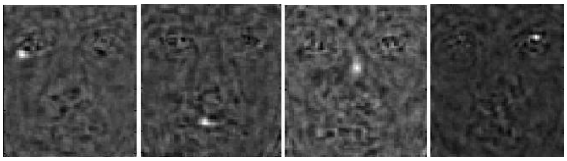
$$K = \Psi \Lambda \Psi^T \tag{2}$$

where $\Psi = [\psi_1, ..., \psi_k], \Lambda = diag(\frac{1}{\sqrt{\lambda_r}})$.

The rows of K contain the kernels. These kernels have spatially local properties and are "topographic" in the sense that the kernels are indexed by spatial location of the pixels in the image. The kernel matrix K transforms the image matrix X to the LFA output $O = KX^T$. LFA constructs n kernels where n is the number of pixels in the images. Since n outputs are described by $p << n$ linearly independent variables, there are residual correlations in the output. Penev and Atick [24] proposed a sparsification algorithm for reducing the dimensionality of the representation by choosing a subset M of kernels that generates an output as decorrelated as possible. In our approach, we use the Fisher's linear discriminant method [32] to choose a subset of kernels that produces an output as decorrelated as possible. Fisher score is a measure of discriminant power which estimates how well different class of data are separated from each other, and is measured as the ratio of variance between the classes to the variance within the classes. Given the LFA output $O = [o_1...o_n]$ for c classes, with each class having n_i samples in the subset χ_i, the Fisher score of the x^{th} kernel, $J(x)$ is given by

$$J(x) = \frac{\sum_{i=1}^{c} n_i (m_i(x) - m(x))^2}{\sum_{i=1}^{c} \sum_{o \in \chi_i} (o(x) - m_i(x))^2}. \tag{3}$$

where $m(x) = \frac{1}{n} \sum_{i=1}^{c} n_i m_i(x)$ and $m_i(x) = \frac{1}{n_i} \sum_{o \in \chi_i} o(x)$. Since spatial location of the kernels correspond to the spatial location of the pixels in the image, those kernels that has higher Fisher score correspond to the feature points in the image. The kernels that correspond to high Fisher scores are chosen to represent the most discriminative features of the image, and are used in our system for training and testing purposes. Figure 1(a) shows the kernels at spatial locations of the right eye, mouth, nose. Figure 1(b) shows the set of feature points extracted using the LFA procedure.



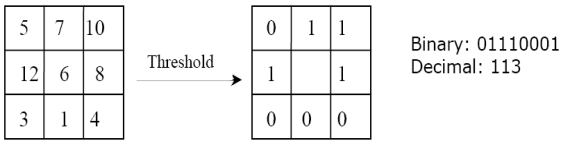
(a) $K(x, y)$ at four locations on the face

(b) Top 100 features with highest Fisher scores

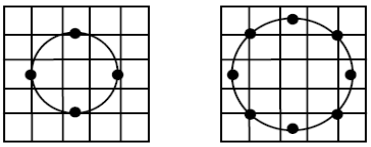
Fig. 1. 1(a) The kernels $K(x, y)$ at right eye, mouth, nose, and left eye position (the white dots). 1(b) Top 100 features corresponding to highest Fisher scores (best viewed in color).

B. Feature Description with Local Binary Pattern

A feature descriptor is constructed for each feature point extracted from an image using Local Binary Pattern (LBP). The original LBP operator proposed by Ojala *et al.* [25] is a simple but very efficient and powerful operator for texture description. The operator labels the pixels of an image by



(a) The Basic LBP operator



(b) $(4,1)$ and $(8,2)$ circular LBP operator

Fig. 2. 2(a) The basic LBP operator. 2(b) $(4,1)$ and $(8,2)$ circular LBP operator.

thresholding the $n \times n$ neighborhood of each pixel with the value of the center pixel, and considering the result value as a binary number. Figure 2(a) shows an example of the basic LBP operator and figure 2(b) shows two circular LBP operators with different radii. The histogram of the labels of the pixels of the image can be used as a texture descriptor. The gray-scale invariance is achieved by considering a local neighborhood for each pixel, and scale invariance is achieved by considering just the sign of the differences in the pixel values instead of their exact values. The LBP operator with P sampling points on a circular neighborhood of radius R is given by,

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p. \tag{4}$$

where

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{5}$$

and g_p, g_c are the intensities of the neighborhood pixels and the center pixel in the window.

An extension to the original operator was also introduced by Ojala *et al.* [26], which uses the property called *uniform patterns* according to which a LBP is called uniform if there exist at most two bitwise transitions from 0 to 1 or vice versa. Uniform patterns can reduce the dimension of the LBP significantly which is advantageous for face recognition. In our experiments, we use the $LBP^{u2}_{P,R}$ which represents a uniform LBP operator with a local neighborhood region of P pixels in a radius R to extract the feature descriptor of each feature point in a window of 5×5 centered by the feature point. In our experiments, we use a value of 8 and 2 for P and R respectively.

C. Graph Construction

We represent each face image by a graph using the feature points as the vertices. The most distinctive property of a graph is its appearance, which is computed from the description vectors of the vertices of the graph. Graph geometry, i.e., the way the vertices of a graph are arranged spatially could vary with every graph and play an important role in discriminating the graphs of different face images. In our approach, the graph geometry is defined by constructing a graph with constraints imposed on the length of the edges between a vertex and its neighbors.

Considering that we extract around n feature points from each face image, at least $n!$ graphs can be generated for each image. Evaluating this number of graphs for each probe image would be very computationally expensive. Also, generating the probabilistic appearance model for the training data set could be computationally expensive when the training data set increases in its size. Hence, a graph generating procedure that imposes constraints on the length of the edges between a vertex and its neighbors is used in our system. At each iteration, vertices and edges are added to the graph in a Breadth-first search manner. Adding the vertices and the edges in a Breadth-first search manner with neighbors of a vertex within a spatial neighborhood distance generates a unique graph with any vertex picked as an initial vertex in the graph generation process. This procedure is efficient since it generates a graph in which a vertex and its neighbors are sufficiently close to each other. This also reduces the computation time during both the training and the testing stage. Also, these geometrical properties can be utilized in the graph matching process for faster matching. The procedure to generate a graph given a set of vertices is given in Algorithm 1.

Algorithm 1 Algorithm for Graph Construction

INPUT: Set of vertices $\{v | v \in V, visited(v) = 0\}$, an empty queue Q
Pick a random vertex $v \in V$;
Add the vertex v to the end of the queue Q;
while $\exists u \in V$,such that $visited(u) = 0$ **do**
  Pick a vertex u from the front of the queue Q;
  if $visited(u) = 0$ **then**
    Find the neighbors N of u, such that $\forall n \in N$, $distance(u, n) < \delta$;
    Add N to the end of the queue Q;
    Set $visited(u) = 1$;
  end if;
end while;

The graph is an effective representation of the spatial relationship between the feature points of an image. It effectively represents the inherent shape changes of a face and also provides a simple, but powerful matching technique to compare graphs.

D. Age Model

Face of humans at a younger age undergoes shape changes, while face of adults undergo textural changes than shape changes. Hence it is appropriate to include the shape and textural changes in an aging model of an individual. In our approach, we learn an age model using the model graphs, extracted from the training images of an individual. Given N individuals and M training face images, the algorithm to learn the model is described in Algorithm 2.

Algorithm 2 Algorithm to Construct Age Model

Initialize N age model sets;
for each training image I_c^j, (j^{th} image of the c^{th} individual) **do**
  Extract the feature points (as described in Section II-A);
  Compute feature descriptors for each feature point (as described in Section II-B);
  Construct Image graph (as described in Section II-C);
  Include the graph in the age model of the c^{th} individual;
end for;

The procedure to construct the age model for each individual using their model sets from their model graphs are explained in detail below.

Given a graph $G(V, E, F)$, where V is the set of vertices in the graph, E is the set of edges in the graph, and F is the set of description vectors for the vertices of the graph; the probability of G belonging to a model set (individual) k is given by,

$$P_k = \max_n P(G|\Phi_n) \tag{6}$$

where P_k is the Maximum a Posterior (MAP) probability, and Φ_n is the age model for the n^{th} subject. In our approach, the age model Φ_n is constructed by estimating the joint probability distribution of the appearance of the graphs of all the images of a subject. We estimate the joint probability distribution of the appearance of the graphs using the d-dimensional feature descriptors that describe the graphs.

In our approach, the joint probability distribution of the appearance of the graphs is modeled using Gaussian Mixture Model (GMM) [31]. GMMs can be used as a parametric estimation technique to estimate the parameters of the aging model. The GMM can efficiently represent heterogeneous data, and the dominant patterns in the data are captured by the component distributions which are Gaussians in a GMM.

Given a training set of N subjects and each subject having at least one image in the training database, the set of feature descriptors for each subject which are used in modeling the joint likelihood will be $(m \times f) \times d$ distribution, where m is the number of images of the subject, f is the total number of feature descriptors extracted from all the images of the same subject, and d is the dimension of the feature descriptor vector (in our case, it is 59 reduced to 20). PCA is applied to reduce the dimension of the feature descriptors of all the images in the training set in order to make the age model estimation more accurate and tractable.

The age model for each subject in the database is modeled as a GMM with K Gaussian components. The feature vectors from all the images of the individual form the set of feature vectors F and is used for training the GMM of that individual. Mathematically, a GMM is defined as:

$$P(F|\Theta) = \sum_{i=1}^{K} w_i N(X|\mu_i, \sigma_i) \tag{7}$$

where

$$N(X|\mu_i, \sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp^{-\frac{(X-\mu_i)^2}{2\sigma^2}} \tag{8}$$

and $\Theta = w_i, \mu_i, \sigma_{i\ i=1}^{2\ K}$ are the parameters of the model, which includes the weight w_i, the mean μ_i, and the variances σ_i^2 of the K Gaussian components. In order to maximize the likelihood function $P(F|\Theta)$, the model parameters are re-estimated using the Expectation-Maximization (EM) technique [33]. For more details about the EM algorithm, see [33].

E. Matching and Recognition

In the testing stage, feature points are extracted from the probe image using the technique explained in Section II-A, and the feature descriptors are computed as explained in Section II-B. A graph G is constructed with the set of feature points obtained for the probe image. The MAP probability is computed in the first stage of matching and is used to effectively prune the search space by selecting the age model of those individuals which are likely to match with the testing image. In the second stage, we use a simple deterministic algorithm to match the graph G with the set of graphs in the training data set. The procedure is given as follows;

Algorithm 3 Algorithm for Graph Matching

INPUT: Test graph G, and the set of training graphs $\{H_c^i | i = 1, ..., n_i, c = 1, ..., n\}$, where H_c^i is the graph of the i^{th} image of the c^{th} class
for each class $c \in C$ **do**
  for each graph $H_c^i \in c$ **do**
    Pick a vertex $v \in G$ with a spatial neighborhood W;
    Search the training graph H_c^i in the corresponding spatial neighborhood W and select the best matching vertex u of the training graph such that,

$$S_{uv} = \frac{f_v \cdot f_u}{|f_v||f_u|} \tag{9}$$

    where f_u and f_v are the feature vectors of u and v respectively, and S_{uv} is the similarity score between u and v;
    Repeat the above step with neighbors of v and so on until all the vertices have been matched;
  end for;
end for;
The classes with high average similarity scores are selected.

III. EXPERIMENTS

We evaluated the performance of the proposed algorithms on FG-NET database and compared the face recognition accuracy with the method proposed by Park *et al.* [2], since authors in this work have evaluated their algorithm with various techniques and have shown good performance than the existing ones. [2] constructs a 3D aging model and

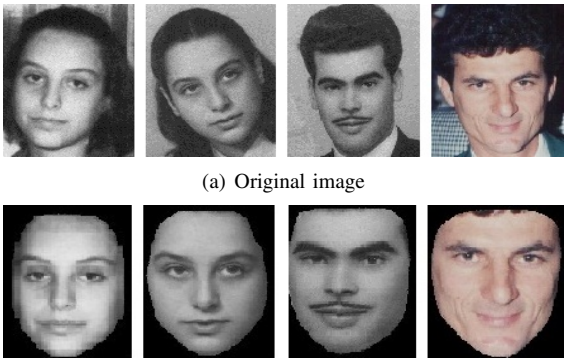
uses it for age-invariant face recognition. We use the FG-NET database [34] for model construction and for testing purposes. The FG-NET aging database includes images of 82 subjects with a total of 1002 images with an average of 12 images per subject. The age of the subjects range from 0 (less than 12 months) - 69.

We performed two experiments to evaluate the performance of the aging model with respect to aging in adults and with the inclusion of the aging effects of young faces in the aging model. For the first experiment, we took the images in the age range from 18 years to 69 years, and the images for each individual were equally divided between the training and the testing set. The training set however included the younger age face images of an individual, and the testing set included the older age face images of the individual. This setup is to exactly mimic the application of passport verification in Homeland security.

All the images were normalized using histogram equalization technique, and resized to 72×60 for computational efficiency. The face region is extracted from the training dataset and rectified using the eye coordinates. We perform a pose correction to the non-frontal face images using Active Appearance Model (AAM) technique as proposed by Cootes *et al.* [35]. The images from the FG-NET database are annotated with 68 points, and we use a generic model to fit these points and calculate the pose of the face, and perform pose correction by warping the image onto the model. Figure 3 shows pose corrected images from the database.

From our experiments, we deduced that an optimum number of feature points to describe a face image is 150, and hence we choose the top 150 feature points for each image in the training set, and their corresponding feature descriptors were computed using uniform LBP. A GMM was constructed for each individual with 10 Gaussian components and using the set of feature descriptors from all the images of that individual. A graph was constructed for each image in the training set using the feature points extracted from the image as the vertices of the graph and labeling them with their corresponding feature descriptors. The maximum Euclidean distance between two vertices was deduced to be half the width of the face image, to ensure a connected graph.

In the testing stage, a graph is constructed for each probe image by extracting the feature points and by computing their corresponding feature descriptors. Dimensionality reduction is done to the feature descriptors to reduce it to a length of 20. The likelihood score of the probe image graph belonging to each age model is computed as explained in section II-D. We choose the top 10% of the models with the highest likelihood scores as the potential list of individuals to be used in the second stage of matching. The graph of the probe image is matched using the procedure explained in section II-E with the set of graphs from the reduced search space. Average of the similarity scores are computed from the probe-gallery image pairs, where the gallery of images are from the chosen subset of individuals from the first stage of matching. The recognition result is based on the average of the similarity scores.



(a) Original image

(b) Pose corrected image

Fig. 3. Pose correction using AAM. 3(a) Original image. 3(b) Pose corrected image.

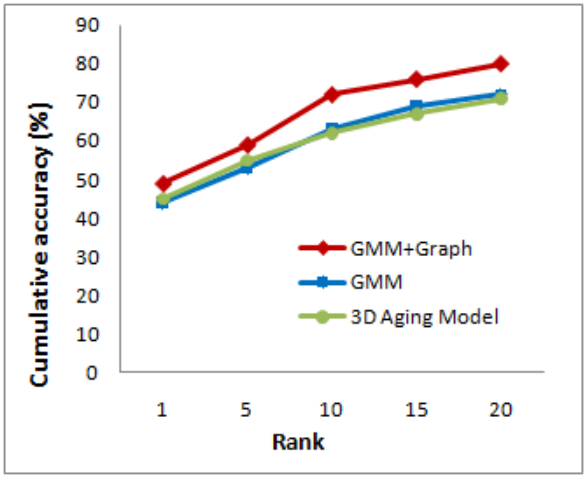


Fig. 4. CMC Curve for age in range [18, 69] for FG-NET database

The performance of the algorithm is evaluated using the Cumulative Match Characteristic (CMC) curve. Figure 4 shows the performance of three methods: GMM with Graphs, GMM, and the 3D aging model technique [2]. The results indicate that our algorithm performs well in comparison to the 3D aging model technique, especially the performance is significantly improved for rank 10 to 20 recognition. It can also be seen that the graph representation and matching significantly improved the performance of the age model in recognition.

In our second experiment, the training set included images of individuals from ages 0 to 30, and the rest of the images of the individual (with higher ages) were included in the testing set. The performance of the algorithm is again represented using the CMC curve and is shown in Figure 5. The results indicate that shape and texture information from young face images incorporated in the age model affect the performance of the system significantly. This is due to the fact that aging in young faces undergo shape changes, while aging in adults is mainly due to textural changes. Large shape variations between the young age and older age face image of the same individual have significant effect on the performance of the algorithm. Figure 5 illustrates this effect, as it can be seen

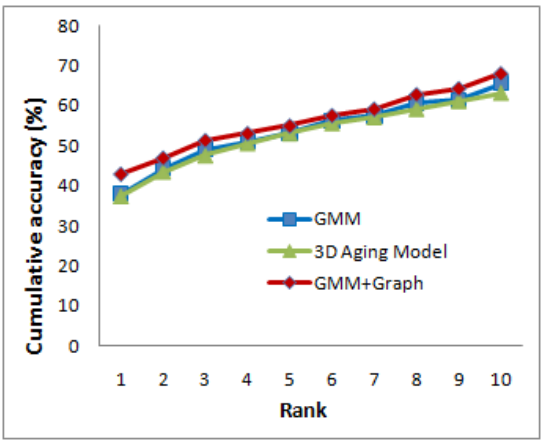


Fig. 5. CMC Curve for age in range [0, 69] for FG-NET database

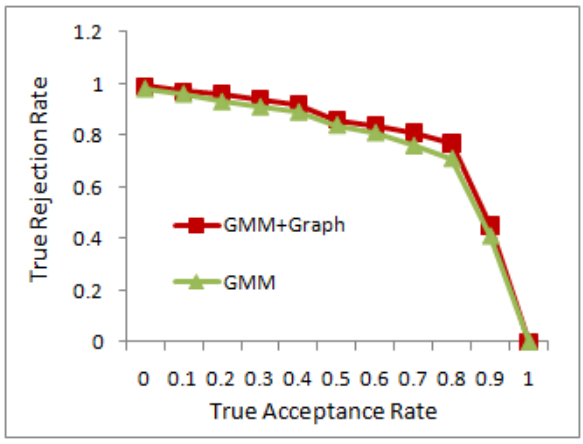


Fig. 6. TAR-TRR Curve for age in range [18, 69] for FG-NET database

that the performance of the approaches are nearly equal with a little improvement in the performance by the GMM+Graph method.

A. Performance Evaluation

We evaluated the performance of the proposed method by comparing the recognition rates after the first stage of matching and after the second stage of matching. The True Acceptance Rate (TAR) - True Rejection Rate (TRR) curves are generated by varying the threshold value for recognition. The parameters TAR and TRR are defined as,

$$TAR = \frac{\#\text{truly accepted intra-personal pairs}}{\#\text{total intra-personal pairs}} \quad (10)$$

$$TRR = \frac{\#\text{truly rejected inter-personal pairs}}{\#\text{total inter-personal pairs}} \quad (11)$$

where an image pair is truly accepted as intra-personal if the images are from the same subject and the image pair is said to be truly rejected extra-personal pair if the images are from different subjects. The *equal error rate* (EER), defined as the error rate when a solution has the same TRR and TAR is also used to measure the performance in addition to TAR-TRR curves. Each image pair included an image from the training set and the testing set to form the intra-personal and inter-personal pairs.

The TAR-TRR curve is computed for the two subsets of the FG-NET database. Figures 6 and 7 show the TAR-TRR curves for the two subsets of FG-NET database. There are several observations that can be made from these figures. The first observation is that the recognition performance is improved with the introduction of the two stage matching process. It is also evident from the figures that the graph based representation effectively improves the performance of the system, as the inclusion of the second stage of matching has improved the recognition rate. The second observation is that the performance of the system is improved in the case of the subset $[18, 69]$ than the subset $[0, 69]$, which clearly shows that the effect of facial shape changes in children play an important role in the recognition. This is also evident from the equal error rates shown in the table I. The error

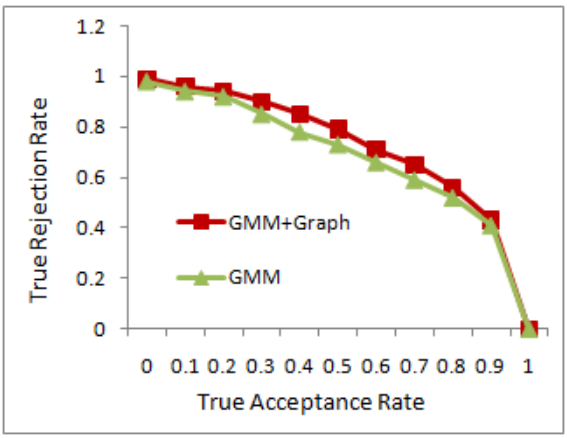


Fig. 7. TAR-TRR Curve for age in range [0, 69] for FG-NET database

rates indicate that the spatial relationship between the facial features and their appearance can effectively be incorporated in an age invariant face recognition system and thus improve its recognition accuracy. The facial shape changes in children can be addressed by generating a mean graph which would effectively represent the shape changes using the feature points.

IV. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

In this paper, we presented a graph based image representation and an aging model constructed using GMM for each individual to model their age variations mainly in shape and texture. A modified Local Feature Analysis that uses Fisher score to extract the feature points has been used effectively to extract feature points. Uniform LBP operator is applied to

TABLE I
EQUAL ERROR RATES FROM EXPERIMENTS ON THE SUBSETS IN AGE RANGE [12,69] AND [0,69] OF FG-NET DATABASE

Age Range	GMM	GMM+Graph
$[18, 69]$	27.2	25.4
$[0, 69]$	31.3	29.2

each feature point to compute a feature descriptor for each feature point, and is used in the graph representation. A two stage approach for recognition has been proposed in which a simple deterministic algorithm that exploits the topology of the graphs is proposed for efficient graph matching between the probe image and the gallery image. The experimental results indicate that the combination of aging model and the graph representation perform well in age invariant face recognition. Thus, an effective representation of the spatial relationship between the feature points of an image can improve the performance of a face recognition system across age progression.

In our future work, we would like to test our algorithm on the MORPH database [36], and also with face images that involve disguise, expression, etc. in addition to age variations.

REFERENCES

[1] Singh, R., Vatsa, M., Noore, A., and Singh, S.K., "Age Transformation for Improving Face Recognition Performance", Springer-Verlag Berlin Heidelberg, 2007.

[2] Park, U., Tong, Y., Jain, AK., "Age Invariant Face Recognition", IEEE Trans Pattern Analysis and Machine Intelligence, 32(5):947-54, 2010.

[3] Jl. Suo, F. Min, S. C. Zhu, S. G. Shan, and X. L. Chen, "A Multi-Resolution Dynamic Model for Face Aging Simulation", Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), June, 2007.

[4] D. M. Burt and D. I. Perrett, "Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and color information," In Proceedings of Royal Society London - Series B, 259:137-143, 1995.

[5] A. Lanitis, C. J. Taylor, T. F. Cootes, "Toward Automatic Simulation of aging effects on face images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(4):442-455, 2002.

[6] A. Lanitis, "On the significance of different facial parts for automatic age estimation," In Proceedings of International Conference on Digital Signal Processing, 2:1027-1030, 2002.

[7] A. Lanitis, C. Draganova, C. Christodoulou, "Comparing different classifiers for automatic age estimations," IEEE Transactions on Systems, Man, and Cybernetics, 34(1):621-628, 2004.

[8] U. Park, Y. Tong, and A. K. Jain, "Face Recognition with Temporal Invariance: A 3D Aging Model", In Proceedings of International Conference on Automatic Face and Gesture Recognition, 2008.

[9] A. Drygajlo, W, Li, and K. Zhu, "Q-Stack aging model for face verification," 17th European Signal Processing Conference, 2009.

[10] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," In Proceedings of IEEE Computer Vision and Pattern Recognition, 1:387-394, 2006.

[11] N. Ramanathan and R. Chellappa, "Face verification across age progression," IEEE Transactions on Image Processing, 15(11):3349-3362, 2006.

[12] N. Ramanathan, R. Chellapa, "Modeling shape and textural variations in aging adult faces," IEEE International Conference on Automatic Face and Gesture, Amsterdam, Netherlands, 2008.

[13] G. Mahalingam, C. Kambhamettu, "Face identification across age progression," International Conference on Image and Video Processing in Computer Vision, 2010.

[14] X. Geng and Z. H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(12):2234-2240, 2007.

[15] X. Geng, Z. H. Zhang, G. Li, H. Dai, "Learning from facial aging patterns for automatic age estimation," Proceedings of the 14th ACM International Conference on Multimedia, 2006, 307-316.

[16] H. Ling, S. Soatto, N. Ramanathan, and D. Jacobs, "A study of face recognition as people age," In Proc. IEEE International Conference on Computer Vision, 1-8, 2007.

[17] H. Ling, S. Soatto, N. Ramanathan, and D. Jacobs, "Face Verification across age progression using Discriminative Methods," IEEE Transactions on Information Forensics and Security, 5(1):82-91, 2010.

[18] J. Wang, Y. Shang, G. Su, and X. Lin, "Age simulation for face recognition," In Proc. International Conference on Pattern recognition, 913-916, 2006.

[19] Guo, G., Fu, Y., Dyer, C. R., Huang, T. S., "Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression," IEEE Transactions on Image Processing, 17(7), 2008.

[20] S. Biswas, G. Aggarwal, R. Chellapa, "A non-generative approach for face recognition across aging," IEEE Second International Conference on Biometrics: Theory , Application and Systems, 2008.

[21] Y. H. Kwon, N. da Vitoria Lobo, "Age classification from facial images," Computer Vision and Image Understanding, 74:1-21, 1999.

[22] K.Scherbaum, "Face Recognition and Growth Prediction using a 3D Morphable Face Model," Master's Thesis, Saarland University, 2007.

[23] U. Jayasinghe, and A. Dharmaratne, "Matching Facial Images using age related morphing changes", World Academy of Science, Engineering and Technology, 60, 2009.

[24] P. Penev and J. Atick, "Local feature analysis: A general statistical theory for object representation," Network: Computation in Neural Systems, 7(30: 477-500, 1996.

[25] Ojala, T., Pietikainen, M., Harwood, D., "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition 51-59, 1996.

[26] Ojala, T., Pietikainen, M., Maenpaa, T., "A generalized local binary pattern operator for multi-resolution gray scale and rotation invariant texture classification," Second International Conference on Advances in Pattern Recognition, Rio de Janeiro, Brazil (2001) 397-406.

[27] N. Ramanathan, R. Chellapa, and Soma Biswas, "Computational methods for modeling facial aging: A survey," Journal of Visual Languages and Computing, 20(3):131-144, 2009.

[28] W. Zhao, R. Chellapa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys, 35(4):399-458, 2003.

[29] Y. Fu, T. S. Huang, "Human age estimation with regression on discriminative aging manifold," IEEE Transactions on Multimedia, 10(4):578-584, 2008.

[30] B. Tiddeman, M. Burt, and D. Perrett, "Prototyping and transforming facial textures for perception research," IEEE Computer Graphics and Applications, 21(5): 42-50, 2001.

[31] McLachlan, J. Peel, D., "Finite mixture models," 2000.

[32] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7): 711-720, 1997.

[33] Redner, R. A., Walker, H. F., "Mixture densities, maximum likelihood and the EM algorithm," SIAM Review 26 195-239, 1984.

[34] FGnet Database, http://www.fgnet.rsunit.com/

[35] T. F. Cootes, G. J. Edwards, C. J. Taylor, "Active Appearance Models," European Conference on Computer Vision, 1998.

[36] Karl Ricanek Jr and Tamirat Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression," IEEE 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, April 2006, 341-345.