Age Sensitivity of Face Recognition Algorithms

D H B Pg Hj Mohd Yassin, S Hoque, F. Deravi

School of Engineering and Digital Arts
University of Kent
Canterbury, United Kingdom.
Email: {dhbp3; s.hoque; f.deravi}@kent.ac.uk

Abstract—This paper investigates the performance degradation of facial recognition systems due to the influence of age. A comparative analysis of verification performance is conducted for four subspace projection techniques combined with four different distance metrics. The experimental results based on a subset of the MORPH-II database show that the choice of subspace projection technique and associated distance metric can have a significant impact on the performance of the face recognition system for particular age groups.

Keywords—face recognition, facial aging, MORPH database;

I. INTRODUCTION

Face recognition has been a very popular modality for biometric applications for decades. A variety of algorithms have been reported in the literature not only to improve accuracy and computational complexity in a tightly controlled scenario but also to address various real-world challenges in the process, such as, variations in pose[1], illumination[2][3], and so on. Among these factors, biological aging contributes to the most significant variations particularly to the face and can degrade the face recognition system performance dramatically. Factors such as ethnicity, lifestyle, gender, living environment and physiological changes that occur with time are some of the uncontrolled factors in human aging [4]. In truth, human aging is not an overnight phenomenon but as a rule is gradual occurring over a period of years.

Nowadays, in many practical systems (e.g., passport control, etc.), the time intervals between two acquired images can lead up to several years. Therefore, there is an urgent need to understand the process of aging and its implications on various biometric systems. Conversely in the recent studies on face recognition, aging factor is generally overlooked. Most of the reported studies in relation to face aging focused on age estimation [5], simulating the aging process for each individual in face appearance as different age [6] and craniofacial aging [7].

In a recent work, Guo *et al* [8] reported face recognition accuracies versus age intervals using PCA approach with elastic bunch graph matching (EBGM) technique. They concluded that the face recognition performance does not degrade linearly with respect to the years of age difference and the use of soft biometrics can improve the accuracy.

Ling *et al* [9] reported an empirical study on how age differences affect recognition performance. Their experimental results show that, although the aging process adds difficulty to the recognition task, it did not surpass the impact due to illumination or expression variations.

Biswas *et al* [4] argued that facial appearance changes in a coherent manner as people age. They suggested measures to capture this coherency in feature drifts. Illustrations and experimental results showed the efficacy of such an approach for matching faces across age progression.

Age sensitivity of biometric systems has also been reported for a number of other modalities. For example, in a study using iris biometrics, Fairhurst *et al* [10] concluded that the physical ageing effect on iris-based biometrics is primarily the result of the physiology of pupil dilation responsiveness decreasing with age and but this degrades the recognition performance for younger individuals.

Modi *et al* [11] reported that fingerprint image quality across age group showed a significant variation not only due to the user interactions but also due to age related physiological conditions such as loss of firmness and moisture. Age related medical conditions may also have an impact. These subsequently affected the system performances, most noticeably for the 62 and older age groups.

In this work, we investigated various face verification systems and variations in their performances when in smaller age groups. Four different feature selection criteria and four distance metrics were studied for the classification algorithm. A comparative analysis is then conducted to evaluate the effects of aging on system accuracy. A subset of the MORPH-II database [12] has been used for the experiments reported here.

II. SYSTEM SETUP

The face recognition system comprises four modules such as image acquisition, preprocessing and normalization, feature extraction, and classification as shown in Figure 1. The classification stage comprises a feature projection scheme combined a distance metric to generate the decision. The database block holds all the enrolled templates. More details of the individual modules are discussed below.
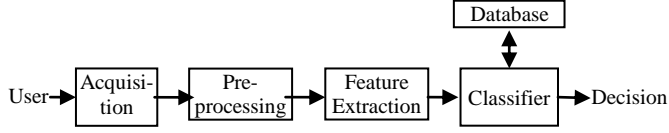
Fig 1. Basic modules of a face recognition system

A. *Preprocessing and Normalisation.* The original RGB face images are first converted to grayscale images. The location of the eye centres are detected manually. The images are then scaled and rotated so that the distance between eye centres are same for all and head tilt is corrected. The resulting images are then cropped to 200x260 pixels (the recommended dimension by NIST is width x width/0.75 [13]) in such a way that faces are centred in the cropped image. Finally the pixel intensities are normalized.

B. *Feature Extraction.* Gabor filters are used to extract features from the face images. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, it can be seen as the product of a complex sinusoid and a Gaussian envelop, as shown in equation (1).

$$G_{\theta_n, f_i, \sigma_x, \sigma_y}(x, y) = \exp\left[-\left\{\frac{x_{\theta_n}^2}{\sigma_x^2} + \frac{y_{\theta_n}^2}{\sigma_y^2}\right\}\right] \cdot \exp(2\pi f_i x_{\theta_n} + \varphi), \quad (1)$$

where,
$$\begin{bmatrix} x_{\theta_n} \\ y_{\theta_n} \end{bmatrix} = \begin{bmatrix} \sin\theta_n & \cos\theta_n \\ -\cos\theta_n & \sin\theta_n \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

and σ_x, σ_y represent the standard deviations of the Gaussian envelope along the two axes, f_i provides the central frequency of the sinusoidal wave at an angle θ_n.

A set of 40 Gabor filters with different frequencies and orientations were used for the feature extraction here. The feature vectors consists of the Gabor magnitudes extracted.

C. *Projection algorithms.* For limited training samples, it is often difficult to model class distributions well if the feature vectors are too large. Use of a feature reduction scheme can reduce the complexity of the task and hence improve the performance of a system. Many techniques have been proposed where high dimensional feature vectors can be projected to a lower dimensional space while minimally compromising (rather often improving) attributes such as information content, discriminability, etc. Four such schemes have been investigated in this study.

Principal Component Analysis (PCA): PCA uses an orthogonal linear transformation to convert a feature vector (possibly with correlated elements) into a set of uncorrelated values in a lower dimensional space. Details of PCA technique can be found in [14].

Linear Discriminant Analysis (LDA): LDA also uses a linear transform based on Fisher criteria of maximizing the ratio of the within-class scatter matrix and the between-class scatter matrix [15]. While PCA ignores the class labels in the determination of the transformation function, LDA optimizes the class separability.

Kernel Fisher Analysis (KFA): An implicit non-linear mapping is used to project the feature vector into a high dimensional feature space where patterns are linearly separable. LDA is then performed in the new feature space [16].

Kernel Principal Component Analysis (KPCA): This process is similar to KFA except that PCA is performed in the new feature space [17].

D. *Distance Metrics.* Identity of the user are verified by comparing the transformed feature vector from a live sample with the ones already stored in the template database. Distances between two vectors can be treated as a measure of dissimalrity of the two biometric samples. Various distance metrics can be found in the literature of which four metrics have been chosen for this study. Mathematical description of these metrics is given below.

Euclidean distance: It is the most commonly used metric based on the Pythagorean formula. If $u = (u_1, u_2, \ldots, u_N)$ and $v = (v_1, v_2, \ldots, v_N)$ are two feature vectors, the Euclidean distance is given by

$$D_{Euc}(u, v) = \sqrt{\sum_{i=1}^{N} (u_i - v_i)^2} \quad (2)$$

City block distance: In this metric, the distance between two vectors is the sum of the absolute differences of their elements. Mathematically, the distance between two vectors u and v is given by,

$$D_{Ctb}(u, v) = \sum_{i=1}^{N} |u_i - v_i| \quad (3)$$

Cosine distance: Cosine similarity is measure of similarity based on the cosine of the angle between two vectors. Two vectors of same orientation has a cosine similarity value of 1 which reduces to -1 as the angle between the vectors increase to 180°. The mathematical expression is shown in equation (4) where a minus sign is introduced to convert the measure to dissimilarity metric in line with the other metrics here. Hence the cosine distance between two vector u and v is given by

$$D_{\cos}(u, v) = -\frac{u \cdot v}{\|u\|\|v\|} = -\frac{\sum_{i=1}^{N} u_i v_i}{\sqrt{\sum_{i=1}^{N} u_i^2} \cdot \sqrt{\sum_{i=1}^{N} v_i^2}} \quad (4)$$

Mahalanobis Cosine: It is similar to cosine distance but computed in the Mahalanobis space as shown in equation (5).

$$D_{\text{MahCos}}(u,v) = -\frac{m \cdot n}{\|m\|\|n\|} \quad \text{where, } m_i = \frac{u_i}{\sigma_i}, n_i = \frac{v_i}{\sigma_i} \quad (5)$$

Here σ_i is the standard deviation of the ith dimension in the Mahanalobis space.

III. EXPERIMENTAL SETUP

A subset of images from the MORPH-II database [12] has been adopted to test the performance of the face recognition system using various feature selection-distance metric combinations. MORPH-II original database consists of 55,132 facial images acquired from 13,618 subjects. Population age range is 16-67 when first enrolled although most of the images are from 20-49 year olds. Data was acquired over a period of 5 years but not everybody provided samples every year. For this study, the population is split into 5 age bands (≤19, 20-29, 30-39, 40-49, >50) depending on their first enrollment age. 20 subjects were randomly picked for each group ensuring that there are equal gender distributions in each group. There are 6 images from each person. Of these 6 images, 3 were used for enrollment and the rest for verification.

For comparative performance analysis, the verification scenario has been implemented in this study. In a biometric verification scenario, False Accept Rate (FAR) denotes the portion of the impostors accepted as genuine whereas False Reject rate (FRR) denotes the proportion of genuine users rejected as impostor. The FAR and FRR depend on the operational parameters. The phenomenon when FAR equals FRR, these errors are termed as Equal Error Rates (EER). We will be reporting all these for our comparative analysis here.

IV. RESULTS AND DISCUSSION

Equal Error Rates were calculated for a range of system parameters and are reported in this section. The preliminary experiment gathered all the 100 subjects together in one group and Table 1 summarises the overall equal error rates (EER) for various combinations of selection algorithms with different distance metrics.

TABLE I.   COMPARATIVE EER (IN %) WITHOUT ANY AGE GROUPING

Feature Projection algorithm	Distance metric			
	Euclidean	*CityBlock*	*Cosine*	*Mahalanobis-Cosine*
PCA	12.9	20.2	13.6	14.2
LDA	8.1	8.7	4.0	4.3
KPCA	16.8	24.3	14.4	49.5
KFA	20.8	17.1	17.2	16.1

From the observations, it is evident that LDA performed consistently well. Among the distance metrics, cosine distance almost always produced the lowest EER except when combined with KFA. Even then the margin is only about 1%. Mahalanobis-cosine distance also produced very low EER except when combined with KPCA. The best EER of 4.0% was achieved with the LDA-Cosine Distance combination. It is therefore evident that the choice of projection algorithm and distance metric has a significant impact on the achievable EER from a face recognition system.

Table 2 shows the EER for all possible combinations of system parameters. It can readily be observed that, even for the same combination of selection algorithm and distance metric, the EER for each age group varied quite drastically. The LDA-Cosine Distance combination produced the lowest EER of 4% when there was no age grouping; but with the age grouping incorporated, the EER values for the groups are significantly lower (the lowest being 1.4% for the 20-29 age band) except for the ≤19 group where the EER increased to 6.3%. Even for the KPC-Mahalanobis/Cosine combination which produced a poor 49.5% EER, after age grouping the EER values ranged between 13.3% and 25.5%. It is also noticed that no particular combination produced the best EER for all the age bands (the lowest EER for any age band is shown in bold in Table 2) suggesting that, for optimum performance, different schemes should be used for different age groups. It can also be noticed that, of all the age groups, the ≤19 group showed the highest EER suggesting that this age group is experiencing most variation in their facial appearances.

TABLE II.   COMPARATIVE EER (IN %) FOR DIFFERENT AGE GROUPS

Projection Algorithm	Distance Metric	Age Group				
		≤19	20-29	30-39	40-49	≥50
PCA	*Euclidean*	18.0	21.3	17.2	29.6	10.4
	CityBlock	14.2	20.3	20.3	32.4	11.3
	Cosine	12.9	17.3	15.1	28.3	5.5
	Maha-Cosine	15.9	18.6	17.8	22.5	9.2
LDA	*Euclidean*	9.3	5.6	9.0	15.7	5.0
	CityBlock	8.3	5.8	10.7	17.5	2.6
	Cosine	6.9	1.4	**1.7**	**1.6**	2.7
	Maha-Cosine	6.3	**1.2**	10.3	8.6	3.3
KPCA	*Euclidean*	22.9	27.6	17.6	11.6	13.4
	CityBlock	20.0	24.6	31.0	24.5	20.5
	Cosine	20.4	22.0	17.6	13.5	8.9
	Maha-Cosine	19.1	21.6	25.5	16.7	13.3
KFA	*Euclidean*	6.8	8.1	9.8	7.9	5.0
	CityBlock	5.7	8.1	7.6	8.2	7.8
	Cosine	**3.6**	2.5	8.9	7.1	2.6
	Maha-Cosine	8.2	16.8	9.0	6.1	**1.1**

Equal Error Rates (EER) present a snapshot of the system performance for a particular operating threshold. To show the general trend in the variations of FAR-FRR over a range of thresholds, Figure 2 shows the DET-curve for the KPCA-Cosine combination. It can be seen that ≥50 age group consistently performed better for all operating thresholds. The 40-49 age group performance was comparable to those of 'No Grouping'. The performances for the remaining age groups were comparatively less accurate.
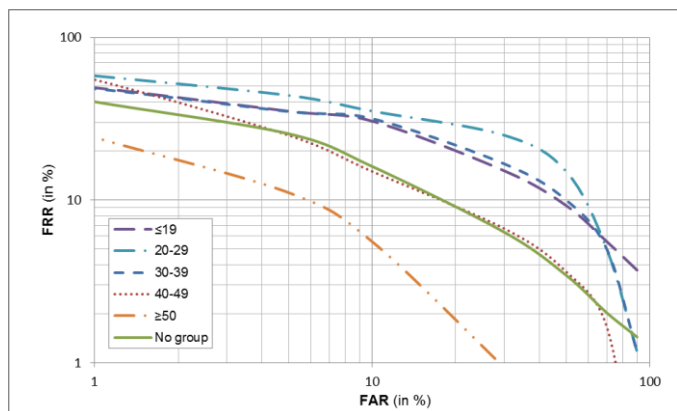

Fig 2. DET curve for the KPCA-Cosine combination

V. CONCLUSION AND FUTURE DIRECTION

The aim of this study was to investigate and analyze the relative performance of different subspace projection algorithms combined with various distance metrics for different age groups. The experiments used a facial image database containing face images acquired over a five year period. It was observed that system performance improved significantly when age groupings were introduced and separate classification schemes were deployed for each age group. It was also observed that for this dataset the EER was highest for the younger population (≤19 year olds). Further work will explore how an adaptive system can be designed to harness this age dependency for maximum recognition performance.

ACKNOWLEDGMENT

REFERENCES

[1] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Generative models for recognition under variable pose and illumination" In Proceedings of the Fourth International Conference on Face and Gesture Recognition, pp.277-284, 2000

[2] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.19(7), pp.721-732, 1997.

[3] D. Jacobs, P. Belhumeur, and R. Basri, "Comparing images under variable illumination," In Proc of the IEEE Conference on Computer Vision and Pattern Recognition, pp 610-617, 1998.

[4] S. Biswas, G. Aggarwal and R.Chellappa, "A non-generative approach form face recognition across aging," In Proc of the IEEE Conference on Biometrics: Theory, Application and System (BTAS), 2008.

[5] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging pattern," IEEE Trans. on PAMI, vol. 29(12), pp.2234-2240, 2007.

[6] N. Ramanathan and R. Chellappa. "Modelling age progression in young face," In Proc of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 387-394, 2006.

[7] K. Ricanek, "Facial Aging Update & Facial Modeling," http://biometrics.org/bc2012/presentations/FaceTechnology/Ricanek420.pdf

[8] G. Guo, G. Mu and K.Ricanek, "Cross-Age face recognition on a very large database: The performance versus age intervals and improvement using soft biometric traits," In Proc of the 20th Int Conference on Pattern Recognition, 2010.

[9] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs. "A study of face recognition as people age," In Proc of 11th IEEE Intl. Conf. on Computer Vision, 2007.

[10] M. Fairhurst and M. Erbilek, "Analysis of physical ageing effect in iris biometrics," IET Computer Vision, vol. 5(6), pp.358-366, 2011.

[11] S.K. Modi, S.J., Elliott, J. Whetsone, H. Kim, "Impact of age groups on fingerprint recognition performance," In Proc of the 5th IEEE Workshop on Automatic Identification Advanced Technologies, pp. 19–23, 2007.

[12] K. Ricanek and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," In Proc of the 7th IEEE Conf. on Automatic face and Gesture Recognition, pages 341–345, 2006.

[13] P. Griffin, "Understanding the face image format standards," ANSI/NIST Workshop, Gaithersburg, MD, USA, 2005.

[14] K.C. Chung, S.C. Kee and S.R. Kim, "Face recognition using principal component analysis of gabor filter responses," In Proc of the IAPR Workshop on Machine Vision Applications, 2000.

[15] X. Wang and X. Tang, "Random Sampling LDA for Face Recognition," In Proc. of the Conference on Computer Vision and Pattern Recognition, 2004.

[16] Y. Li, S. Gong, and H. Liddell, "Recognizing trajectories of facial identities using kernel discriminant analysis," Image and Vision Computing, vol. 21(13-14), pp.1077–1086, 2003.

[17] X. Xie and K. Lam, "Gabor-Based Kernel PCA with Doubly Nonlinear Mapping for Face Recognition with a Single Face Image," IEEE Trans on Image Processing, vol.15(9), 2006.