2010 International Conference on Pattern Recognition

Cross-Age Face Recognition on A Very Large Database: The Performance Versus Age Intervals and Improvement Using Soft Biometric Traits

Guodong Guo
West Virginia Univ.
guodong.guo@mail.wvu.edu

Guowang Mu
NC Central Univ.
gmu@nccu.edu

Karl Ricanek
Univ. of NC Wilmington
ricanekk@uncw.edu

Abstract

Facial aging can degrade the face recognition performance dramatically. Traditional face recognition studies focus on dealing with pose, illumination, and expression (PIE) changes. Considering a large span of age difference, the influence of facial aging could be very significant compared to the PIE variations. How big the aging influence could be? What is the relation between recognition accuracy and age intervals? Can soft biometrics be used to improve the face recognition performance under age variations? In this paper we address all these issues. First, we investigate the face recognition performance degradation with respect to age intervals between the probe and gallery images on a very large database which contains about 55,000 face images of more than 13,000 individuals. Second, we study if soft biometric traits, e.g., race, gender, height, and weight, could be used to improve the cross-age face recognition accuracies, and how useful each of them could be.

1. Introduction

Facial aging refers to the problem in face recognition where the time difference between the face images of the same person is large. The time separation between two face images could be several years or tens of years. Facial aging is one of the major reasons of performance degradation in face recognition. Traditional work on face recognition focuses on dealing with pose, illumination, and expression (PIE) variations, and the CMU-PIE [14] and FERET [11] are usually adopted for experimental evaluation. Compared to the PIE changes, facial aging can influence the face recognition performance very significantly. How big the aging influence could be in face recognition? No previous work has quantitatively verified the aging influence, especially on a large database. Here for the first time we investigate the aging influence on face recognition in terms of age intervals using a very large database.

In comparison with the large number of publications on face recognition related to PIE variations, only a few studies were reported on cross-age face recognition. A typical way for cross-age face recognition is to simulate the aging process for each individual and render new face images at different ages for matching [6] [12] [3] [17] [10] [9]. A problem of aging simulation based face recognition approaches is that the synthesized face images at new ages might be distorted because there are many uncontrolled factors in human aging, such as health, life style, living location, and weather conditions. Further, it is not clear whether the age simulation based methods are applicable to a large database. Another category of approach is to extract facial features that are not sensitive to aging variations, called gradient orientation pyramid [7]. The authors performed face recognition experiments on a passport database containing 1,824 subjects. The maximum age gap is about 9 to 11 years in their experiments. It is not clear about the performance of the method on a larger database containing larger age gaps, e.g., 20 or 30 years.

In this paper, we want to investigate the performance of cross-age face recognition on a large database containing large age spans. We also want to study if soft biometric traits can be used to improve the accuracy, which has not be examined before in the context of cross-age face recognition, although some soft biometric features were used previously for fingerprint [4] or face recognition [5, 8]. Our main contributions:

- Study the face recognition accuracies versus age intervals using different approaches.
- Study whether soft biometric traits are useful for cross-age face recognition and how useful.
- Present a benchmark result of cross-age face recognition on a very large database. This may inspire further research in the future.

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.828

Table 1. Recognition accuracies of different approaches vs. age intervals between the probe (Pr.) and gallery (Ga.) faces, and the percentage of improvement using soft biometrics.

Age span between Pr. and Ga.	Face Recognition Accuracy					
	PCA	PCA+SOFT	EBGM	EBGM+SOFT	PCA+EBGM	PCA+EBGM+SOFT
(0, 1]	46.23%	49.79%	33.36%	36.42%	50.58%	54.12%
(1, 2]	42.31%	45.76%	28.86%	31.55%	46.25%	49.80%
(2, 3]	42.64%	46.39%	30.82%	32.83%	47.39%	49.29%
(3, 4]	46.97%	50.72%	33.52%	37.42%	51.19%	51.43%
(4, 5]	44.74%	48.76%	30.98%	34.61%	50.29%	50.67%
(5, 10]	45.78%	49.12%	30.45%	33.01%	51.47%	53.24%
(10, 15]	41.72%	46.55%	26.55%	27.59%	48.28%	51.38%
(15, 32]	29.79%	34.04%	17.02%	19.15%	34.04%	34.04%
Overall	44.58%	48.15%	31.56%	34.41%	48.89%	52.09%
Soft Improve	-	8.0%	-	9.0%	-	6.5%

2. The Database

The MORPH database [13] was adopted in our research. It is a large database containing two sections, I and II. Since MORPH-I is too small (1,690 face images), we chose to use MORPH-II for our study. The FGNET [1] is also too small. The original MORPH-II has 55,132 face images of 13,617 subjects. There are 457 subjects having only one image in the database. We removed those subjects, and used individuals with at least two face images for face recognition. As a result, we have 54,675 face images of 13,160 subjects for our experiments. The age range is from 16 to 67 years in MORPH-II. And the age intervals for the subjects range from less than one year to more than 30 years.

Few work has been reported on MORPH-II, although MORPH-I [13] was used in previous works. To our knowledge, this is the first time to perform cross-age face recognition using this very large MORPH-II database with 13,160 subjects of 54,675 face images. No previous work has used so many subjects in any face recognition study.

The face images in MORPH-II were detected and aligned with eye centers, and then cropped, resized to 60×60, histogram equalized, and intensity normalized. We only used the gray level images for face recognition.

3. Primary Features

Eigenfaces [15] [16] are a classical technique for face representation using the principal component analysis (PCA) method. The elastic bunch graph matching (EBGM) [18] is another technique for practical face recognition and has been used in many commercial face recognition systems. The key idea of EBGM is to extract features using Gabor filters with various scales and orientations at a couple of fiducial points. These two methods and their variations were used extensively in FERET evaluations [11]. We use both of them to extract the primary features in our cross-age face recognition. The implementation in the CSU Face Identification Evaluation System [2] was adopted. Through evaluating the performance of cross-age face recognition on the large database of MORPH-II, our main purpose is to investigate how face recognition is affected by the age differences between the query and gallery faces, and benchmark the results for future research.

Given a probe image p_i $(i = 1, 2, \cdots, m)$, its distance to a gallery image g_j $(j = 1, 2, \cdots, n)$ is computed based on a primary feature. Denote the distance by d_{ij}. The similarity between p_i and g_j using the primary feature is s_{ij}, and $s_{ij} = -d_{ij}$. For probe image p_i, we sort the s_{ij}'s in the descending order, and pick the K-th value to normalize and obtain a probabilistic measure of the similarity between p_i and g_j,

$$P_{ij} = \begin{cases} \frac{s_{ij}-s_{iK}}{s_{i1}-s_{iK}}, & if \;\; s_{ij} \geq s_{iK} \\ 0, & otherwise \end{cases} \tag{1}$$

In our experiments, we chose $K = 1000$. The purpose of using the probabilistic measure is to combine the primary feature with soft features.

4. Soft Biometric Traits

For soft biometrics, we used race, gender, height, and weight. Some soft biometric traits could be extracted from the face images, such as race and gender, while others have to be recorded in the acquisition process, such as weight and height. In order to obtain a fair evaluation of these soft traits, we used the recorded data with the face database.

For soft biometric features, we also change them into probabilistic values. In the following, "G" is for gender, "R" for race, "W" for weight, and "H" for height.

$$P(G(p_i)|G(g_j)) = \begin{cases} 1, & if \;\; G(p_i) = G(g_j) \\ 0, & otherwise \end{cases} \tag{2}$$

$$P(R(p_i)|R(g_j)) = \begin{cases} 1, & if \;\; R(p_i) = R(g_j) \\ 0, & otherwise \end{cases} \tag{3}$$

$$p(H(p_i)|H(g_j)) = \gamma_1 \exp\left(-\frac{(H(p_i) - H(g_j))^2}{2\sigma_h^2}\right) \tag{4}$$

$$p(W(p_i)|W(g_j)) = \gamma_2 \exp\left(-\frac{(W(p_i) - W(g_j))^2}{2\sigma_w^2}\right) \tag{5}$$

We chose $\sigma_h = 4.5$ and $\sigma_w = 40$ in our experiments.

To combine the primary and soft biometric features, we used a probabilistic fusion scheme similar to [4], although the primary feature was fingerprint in [4]. The final similarity measure is

$$\begin{aligned} S(i,j) &= \alpha_{PCA} \cdot \log(P_{ij,PCA} + \varepsilon) \\ &+ \alpha_{EBGM} \cdot \log(P_{ij,EBGM} + \varepsilon) \\ &+ \beta_G \cdot \log(P(G(p_i)|G(g_j)) + \varepsilon) \\ &+ \beta_R \cdot \log(P(R(p_i)|R(g_j)) + \varepsilon) \\ &+ \beta_H \cdot \log(p(H(p_i)|H(g_j))) \\ &+ \beta_W \cdot \log(p(W(p_i)|W(g_j))) \end{aligned} \tag{6}$$

A small value of ε was used to avoid the log of zero. Considering the accuracy difference of PCA and EBGM methods (see experiments), we chose $\alpha_{PCA} = 0.7$ and $\alpha_{EBGM} = 0.3$. Since the probability calculation for gender and race is different from that for height and weight, and the different contributions of height and weight (see experiments), we set $\beta_G = 1$, $\beta_R = 1$, $\beta_H = 0.5$, $\beta_W = 0.3$, and $\gamma_1 = \gamma_2 = 1$. Keeping some coefficients while setting the remaining to zero, we can study the performance of each primary feature and the contribution of each soft feature, and various combinations of the primary and soft features.

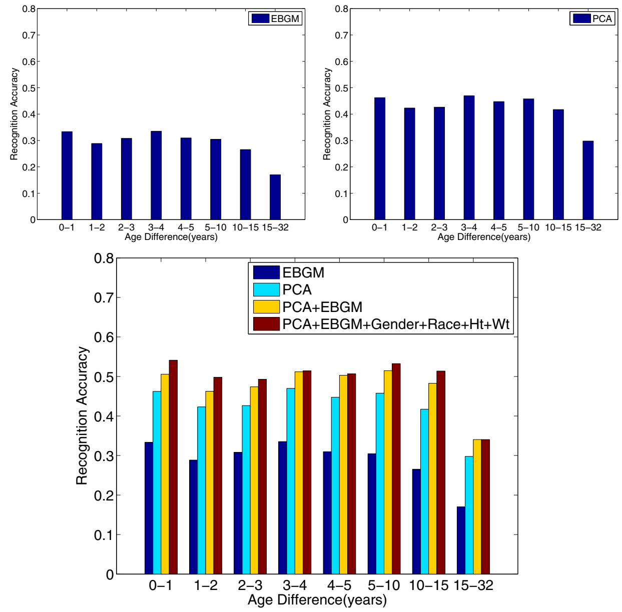

Figure 1. Face recognition vs. age intervals: (1) EBGM, (2) PCA, and (3) Put together of PCA, EBGM, PCA+EBGM, and PCA+EBGM+SOFT to compare.

5. Experimental Results

We first study the face recognition performance versus the age difference between the probe and gallery images. The recognition accuracies are shown in Table 1 and partially displayed in Figure 1. "SOFT" means using all soft features together. From the results, we can observe that the face recognition accuracy is pretty low on the large database containing aging variations, e.g., 31% to 48%, compared to the usually reported accuracies of 70% to 90% in other databases. Please notice that the recognition accuracies are not very different in the age range of 0 to 15 years. In other words, the recognition accuracies do not degrade linearly with respect to age intervals. When the age differences are greater than 15 years, the accuracies become significantly lower than those within 15 years.

Next, we investigate how useful each soft biometric trait is to improve the face recognition accuracies over the primary features. Because of the space limit, we only display the soft features applied to the primary feature PCA in Figure 2. Basically, the gender, race, and height contribute a little bit more than the weight. But all soft features should be combined together (see the last one) in order to have a significant improvement over the primary features. The same observations were obtained for the primary feature EBGM and the combi-

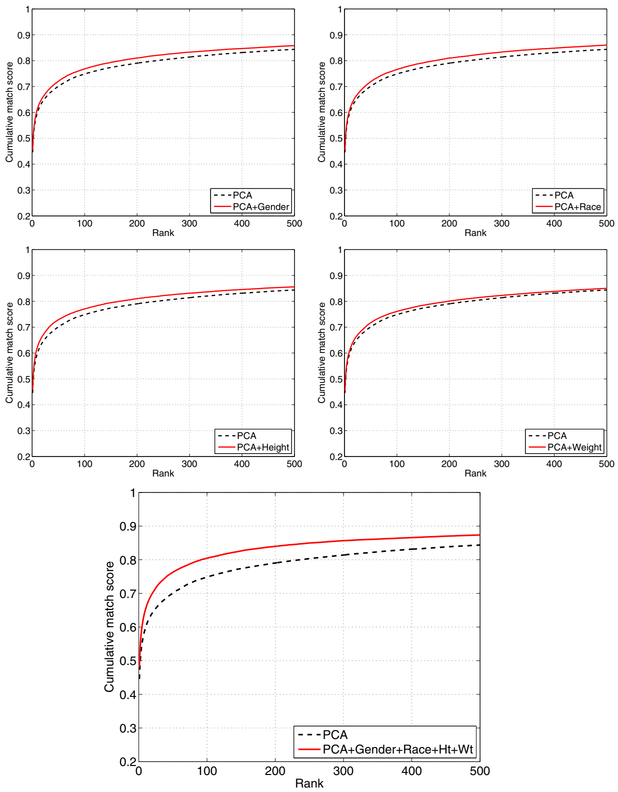

Figure 2. Improving face recognition using each soft biometric trait and their combinations for primary feature PCA.

nation of primary features, PCA and EBGM. So in practice, one should take all available soft biometric traits together instead of using just one or two.

6. Conclusions

We have studied the problem of face recognition influenced by aging variations. The face recognition performance does not degrade linearly with respect to the years of age difference. When the age intervals are larger than 15 years, the performance becomes much lower than within 15 years. Using soft biometrics can improve the accuracy of each primary feature or the combination of the primary features we explored. All soft biometric traits combined together can improve the accuracy of cross-age face recognition significantly. Overall, the performance is still very low in large-scale cross-age face recognition. More efforts are needed.

References

[1] The fg-net aging database. In *http://www.fgnet.rsunit.com/*.

[2] D. Bolme, J. R. Beveridge, and M. T. . B. Draper. The csu face identification evaluation system: its purpose, features and structure. In *Conf. on Vision Systems*, pages 304–311, 2003.

[3] X. Geng, Z.-H. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE Trans. on PAMI*, 29(12):2234–2240, 2007.

[4] A. K. Jain, S. Dass, and K. Nandakumar. Soft biometric traits for personal recognition systems. *Lecture Notes in Computer Science, Springer*, 3072:1–40, 2004.

[5] A. K. Jain, K. Nandakumar, X. Lu, and U. Park. Integrating faces, fingerprints, and soft biometric traits for user recognition. In *Proc. of the Biometric Authentication Workshop*, pages 259–269, 2004.

[6] A. Lanitis, C. J. Taylor, and T. F. Cootes. Toward automatic simulation of aging effects on face images. *IEEE Trans. on Pattern Anal. Mach. Intell.*, 34(1):442–455, 2002.

[7] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs. A study of face recognition as people age. In *Intl. Conf. on Computer Vision*, 2007.

[8] G. Marcialis, F. Rolia, and D. Muntonia. Group-specific face verification using soft biometrics. *Journal of Visual Languages & Computing*, 20:101–109, 2009.

[9] U. Park, Y. Tong, and A. K. Jain. Face recognition with temporal invariance: A 3d aging model. In *Intl. Conf. on Automatic Face and Gesture Recognition*, 2008.

[10] E. Patterson, K. Ricanek, M. Albert, and E. Boone. Automatic representation of adult aging in facial images. In *Proc. Intl. Conf. on Visulization, Imaging, and Image Processing*, pages 171–176, 2006.

[11] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The feret evaluation methodology for face recognition algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(10):1090–1104, 2000.

[12] N. Ramanathan and R. Chellappa. Modeling age progression in young faces. In *IEEE CVPR*, pages 387–394, 2006.

[13] K. Ricanek and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *IEEE Conf. on AFGR*, pages 341–345, 2006.

[14] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression database. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(12):1615–1618, 2003.

[15] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. of the Optical Society of America A*, (4):519–524, 1987.

[16] M. Turk and A. Pentland. Eigenfaces for recognition. *J. of Cognitive Neuroscience*, 3(1):71–86, 1991.

[17] J. Wang, Y. Shang, G. Su, and X. Lin. Age simulation for face recognition. In *Intl. Conf. on Pattern Recognition*, pages 913–916, 2006.

[18] L. Wiskott, J. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):775–779, 1997.