Age Classification from Facial Images: Is Frontalization Necessary?

A. Báez-Suárez[1,2](✉), C. Nikou[1](✉), J.A. Nolazco-Flores[2](✉), and I.A. Kakadiaris[1](✉)

[1] Computational Biomedicine Lab, Department of Computer Science, University of Houston, Houston, TX, USA
{cnikou,ikakadia}@central.uh.edu

[2] Department of Computer Science, ITESM Campus Monterrey, Monterrey, Mexico
basuam@gmail.com, jnolazco@itesm.mx

Abstract. In the majority of the methods proposed for age classification from facial images, the preprocessing steps consist of alignment and illumination correction followed by the extraction of features, which are forwarded to a classifier to estimate the age group of the person in the image. In this work, we argue that face frontalization, which is the correction of the pitch, yaw, and roll angles of the headpose in the 3D space, should be an integral part of any such algorithm as it unveils more discriminative features. Specifically, we propose a method for age classification which integrates a frontalization algorithm before feature extraction. Numerical experiments on the widely used FGnet Aging Database confirmed the importance of face frontalization achieving an average increment in accuracy of 4.43%.

1 Introduction

Human age classification from facial images has become an active research topic in computer vision and pattern recognition due to its applications to demographic analysis, electronic customer relationship management, and video security surveillance. Nonetheless, human age classification is challenging because of: (i) the aging process, which is complicated, irreversible and uncontrollable [1]; (ii) changes in apparent age due to facial hair (beards and mustaches) and makeup; and (iii) the difficulty in collecting complete and sufficient training data [2].

The aging process and apparent age have been modeled via geometrical and textural features which have been employed to describe the craniofacial development that occurs as a series of overlapping events in human chronological age [3]. Specifically, geometrical features consist of ratios [4] or models created from fiducial points such as active appearance models [5] and the textural features consist of filters which are able to find patterns in the skin. The most common textural features are local binary patterns (LBP) [6], Gradient Orientation Pyramid (GOP) [7], and biologically-inspired features (BIF) [8]. However, the discriminative information that is extracted using geometrical and textural features depends on the pose of the head with respect to the camera.

G. Bebis et al. (Eds.): ISVC 2016, Part I, LNCS 10072, pp. 769–778, 2016.
DOI: 10.1007/978-3-319-50835-1_69

Kwon and Lobo [4] categorized facial images into three age groups: babies, young adults, and senior adults, using six ratios of distances between primary facial components (e.g., eyes, nose, mouth, chin). Geng *et al.* [9] proposed an automatic age estimation method named AGing pattErn Subspace (AGES) which, instead of taking each facial image as a single point in the aging pattern, models each aging pattern as a sequence of samples. Age estimation is then accomplished by minimizing the reconstruction error. Guo *et al.* [8] extracted biologically-inspired features which proved to be effective as they are still employed in the age estimation task. Kilinc and Akgul [10] extracted 34 ratios and textural features such as BIF [8] and LBP [6] which are fused in eight different ways, confirming that classifiers perform better if both geometric and textural features are provided. Recently, Levi and Hassner [11] developed a shallow convolutional neural network for age and gender classification. The deep-learning architecture was designed to avoid overfitting due to the limitation of training data by increasing the size of the training set using cropped versions of the training images.

Although much work has been done on the task of human age classification, most of the published work focuses on frontal images. Mirzaei and Toygar [12] studied the influence of gender on the age classification task, but their work is constrained to frontal images free of glasses, mustache, and beard. Levi and Hassner [11] also constrained their work to in-plane aligned images when feeding their convolutional neural network (CNN). Liu *et al.* [13] described two new geometrical features (CirFace and Angle) which are Gaussianly distributed along certain age ranges and are defined on frontal images. Nevertheless, current databases, such as the FGnet Aging Database [14], contain images in the wild with variations in head pose.

In this work, we present an algorithm for Frontalized fACial Image Age cLassification (FACIAL). A method that addresses the age classification task employing a frontalization technique [15] which normalizes pose, size, and alignment before the feature extraction procedure. The importance of the frontalization technique is examined in different image resolutions and in different age groups.

Our first contribution is proposing the use of a frontalization algorithm, which normalizes the facial images reducing the variation due to the position of the head and increases the classification accuracy. The second contribution is the study of BIF and GOP features for different image resolutions. Therefore, the effect of frontalization can be quantified in the presence of different features and image resolutions. Finally, new age groups are defined following the human development [16], which result in improved accuracy. A comparison with the age groups proposed in Liu *et al.* [17] justifies the proposed age grouping.

The rest of the paper is organized as follows: In Sect. 2 we introduce the overall method. In Sect. 3 we discuss important implementation details and present the experimental evaluation. In Sect. 4 we summarize our findings.

2 Age Classification Using Frontalized Facial Images

The steps of FACIAL are illustrated in Fig. 1. Standard pre-processing converts images from color to grayscale and, after locating the center of the eyes, the images are rotated, scaled, cropped and aligned. In our method, frontalization [15] registers a 2D facial image onto a 3D facial model where pose, size, and alignment are normalized.

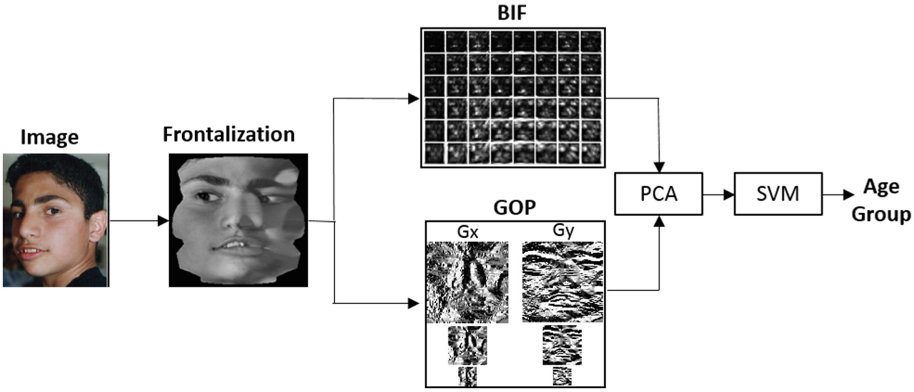

Fig. 1. Depiction of a block diagram of FACIAL. Facial images are pre-processed in order to have the same pose, size, and alignment. Then, features are extracted (BIF/GOP) and PCA is employed to reduce the dimensionality. Finally, the age group is determined through an SVM classifier.

In frontalization, a 3D Annotated Face Model (AFM) [18] is reconstructed from single 2D images to lift the 2D facial appearance [19] to a canonical 2D space (geometry image space) [18]. To construct the final representation, a mapping between the original image and the geometry image space is performed. It has been reported that face alignment [20] has an impact on the estimation of age. In this work, we study how frontalization improves the performance of the age classification. Representative examples of frontalized images are depicted in Fig. 2.

With regard to image resolution, there is no preferred standard size. In many works [8,20,21], an image resolution of 60×60 has been used, in Liu *et al.* [13] the image resolution is 180×150 and in Levi and Hassner [11] the image resolution is 256×256.

In this work, two features are employed: BIF [8] and GOP [7]. BIF features are derived from a feedforward model of the primate visual object recognition pathway (HMAX Model [22]). The method alternates between layers called simple (S) and complex (C) cell units. BIF features are built following the procedure in Guo *et al.* [8] and are chosen because several state-of-the-art results have been achieved [2,23–25] using these features. GOP features are insensitive to

Fig. 2. Representative examples of non-frontalized and frontalized images under various initial head pose orientations. The first column indicates the dominant orientation of the non-frontalized image.

illumination changes and the pyramid provides robustness at different scales [7]. They are based on normalizing the gradient vector at each pixel and concatenating the results for both directions. From the work presented in Liu *et al.* [13], it is observed that frontal faces are better described using GOP features which improve the classification accuracy.

Dimensionality reduction of the feature vectors is accomplished through principal component analysis (PCA) [26]. Then, a one-versus-all multiclass support vector machine (SVM) [27] is trained to classify the features, following the age groups defined in Table 1. The overall methodology is described in Algorithm 1. The FGnet Aging Database [14] was used for testing the methodology.

An ideal definition of the age groups would follow the human craniofacial development discussed by Shu *et al.* [3]. However, there is no one-to-one correspondence between age groups and craniofacial state. Therefore, we propose to define the age groups based on the aging process defined by Armstrong [16], as aging and craniofacial development are correlated. Armstrong [16] described overlapping age groups, but in our work, non-overlapping age groups are chosen to set crisp boundaries with respect to craniofacial development.

Table 1. Age groups used by Liu *et al.* [17] and the age groups proposed in this paper.

Author	Group name	Age groups									
Liu *et al.* [17]	Liu-3	0–3	4–19	20–69	-	-	-	-	-	-	-
	Liu-4	0–5	6–12	13–21	22–69	-	-	-	-	-	-
	Liu-5	0–4	5–10	11–15	16–29	30–69	-	-	-	-	-
	Liu-6	0–4	5–9	10–14	15–29	30–49	50–69	-	-	-	-
	Liu-7	0–4	5–9	10–14	15–19	20–25	26–35	36–69	-	-	-
Proposed Groups	FACIAL-7	0–3	4–6	7–8	9–11	12–19	20–35	36–50	-	-	-
Liu *et al.* [17]	Liu-8	0–4	5–9	10–14	15–19	20–29	30–39	40–49	50–69	-	-
	Liu-9	0–4	5–9	10–14	15–19	20–29	30–35	36–41	42–49	50–69	-
	Liu-10	0–4	5–9	10–14	15–19	20–29	30–34	35–39	40–44	45–49	50–69

Algorithm 1. Frontalized fACial Image Age cLassification (FACIAL).

Input: Facial Image
Output: Age Group Label
1: Frontalize the input image
2: Extract BIF or GOP features from the frontalized image
3: Reduce the dimensionality of the features using PCA
4: Classify the reduced features using SVM

3 Experimental Results

To evaluate FACIAL, we have used the publicly available FGnet Aging Database [14] which contains 1,002 color and gray-scale facial images with large variations in illumination, pose, and expression (the pictures are taken in the wild). The age range is from 0 to 69 years old with chronological aging images available for each subject as each person has between 6 to 18 facial images at different ages. It comprises 34 female and 48 male subjects for a total of 82 subjects. The proposed age groups for the FGnet Aging Database are presented in Table 2 where the number of images and the number of subjects per age group are reported.

Table 2. FGnet DB [14] summary.

Age groups	#Images	#Subjects	#Images/Subject
0–3	151	68	2.2
4–6	123	66	1.9
7–8	72	54	1.3
9–11	98	59	1.7
12–19	266	79	3.4
20–35	201	55	3.7
36–50	70	30	2.3
51–80	21	8	2.6

FGnet Aging Database [14] contains images with interpupillary distances (IPD) ranging from 78 to 200 pixels. For comparison purposes, we selected two image resolutions (60×60 and 150×150) having an IPD of 32 and 80 pixels, respectively. BIF and GOP features were extracted. BIF were created with 12 scales and 8 orientations resulting in a feature dimensionality of 6,752 and 45,424 for the 60×60 and 150×150 image resolutions, respectively. GOP were built with two and three scales resulting in a feature dimensionality of 9,450 and 59,860 for the 60×60 and 150×150 image resolutions, respectively. The features are normalized to have zero-mean and unit standard deviation. Next, dimensionality reduction was performed through PCA retaining 95% of the variance. Furthermore, to determine the parameter of the SVM linear kernel classifier, a 5-fold cross-validation was applied for the cost parameter in the range of $[2^{-5}, 2^{-3}, ..., 2^{15}]$ and the area under the ROC curve was the criterion for the selection of the best model. In particular, the LibSVM library [28] was used to train the classifier. Finally, Leave-One-Person-Out (LOPO) was employed for testing and comparison.

Representative results from FACIAL are depicted in Fig. 3. The variation in pose is normalized for the frontalization technique delivering a better

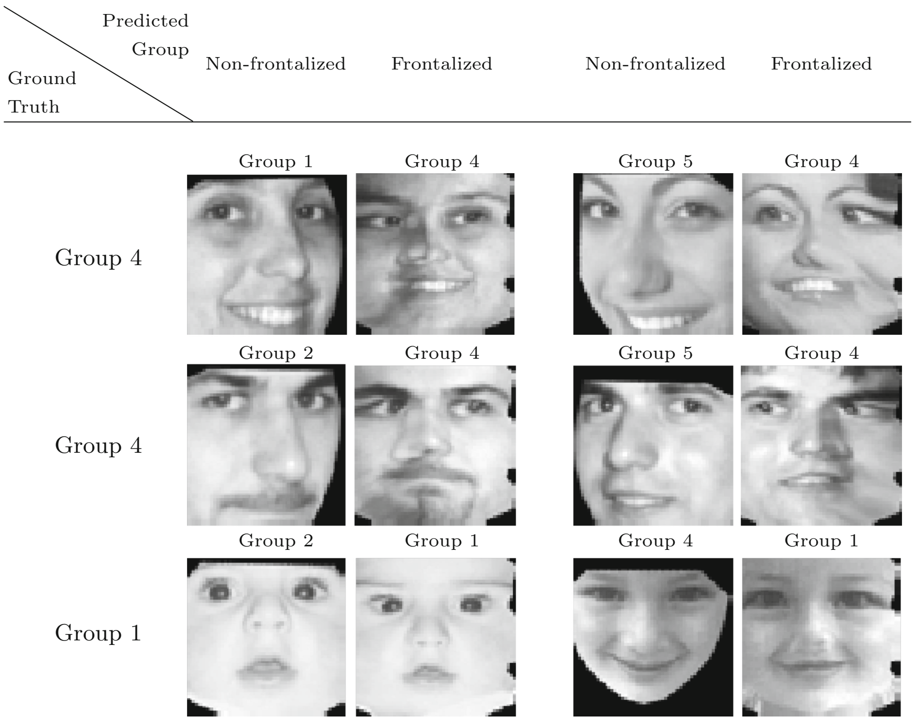

Fig. 3. Representative results from FACIAL applied to images from the FGnet Aging Database [14], where frontalized images are correctly classified with respect to their non-frontalized counterpart. Five age groups were considered.

Table 3. Classification accuracy (%) on the FGnet Aging Database using images with a resolution of 60×60 and 150×150, and the age groups proposed by Liu *et al.* [17]. Results for both BIF and GOP features are presented.

#Groups	Size 60×60				Size 150×150			
	Non-frontalized		Frontalized		Non-frontalized		Frontalized	
	BIF	GOP	BIF	GOP	BIF	GOP	BIF	GOP
3	73.65	**74.35**	74.85	**75.25**	**73.75**	73.25	**77.54**	77.35
4	**61.08**	58.78	**65.67**	64.07	**62.87**	60.68	**66.37**	65.57
5	**54.49**	50.90	**58.88**	57.68	**55.69**	52.99	**57.39**	55.59
6	**53.09**	51.10	52.89	**55.09**	**52.99**	52.79	56.19	**56.99**
7	**42.22**	39.92	42.51	**43.11**	**42.81**	41.22	**47.21**	47.01
8	39.72	**40.42**	40.22	**41.82**	**41.52**	41.12	45.51	**45.71**
9	**39.92**	39.12	40.02	**41.32**	**41.12**	40.52	**45.01**	44.91
10	38.32	**39.12**	39.72	**41.62**	**40.92**	40.22	**45.21**	45.11

classification performance. The experiments were designed to examine the impact of frontalization, image resolution, and age groups definition. The classification using the age groups proposed by Liu *et al.* [17] are considered as baseline.

Table 3 summarizes the results of age classification using the non-frontalized and frontalized images. On average, when the image resolution is 60×60 and images are frontalized, the classification accuracy increases along the different splits by 1.53% for BIF and 3.28% for GOP. When the image resolution is 150×150, the classification accuracy is increased by 3.60% for BIF and 4.43% for GOP. From both of these results, we may conclude that frontalization improves the performance in age classification. It may also be observed that the frontalized version of GOP achieves higher accuracy than the non-frontalized version. In the case of BIF features, the non-frontalized version and the frontalized version have similar performance.

Table 3 also highlights the difference in performance when the image resolution increases. On average, the classification accuracy increases along the different splits by 1.15% for BIF and 1.14% for GOP when the non-frontalized images are compared and by 3.21% for BIF and 2.29% for GOP when the frontalized images are compared. In the case of 60×60 images, the frontalized version of GOP is superior to the non-frontalized versions of BIF and GOP. On the other hand, for the 150×150 images the non-frontalized and frontalized versions of BIF are better.

Table 4 summarizes the performance achieved when the age groups are defined following the human age development based on Armstrong [16]. It is worth mentioning that to avoid imbalance in the data set, the 21 images of the last group in Table 2 were not used because the number of images in that group is very low with respect to the number of images in the other groups.

Table 4. Comparison of classification accuracy (%) using the Liu-7 group (1,002 images) and FACIAL-7 group (981 images).

Groups by	Groups	Size 60×60				Size 150×150			
		Non-frontalized		Frontalized		Non-frontalized		Frontalized	
		BIF	GOP	BIF	GOP	BIF	GOP	BIF	GOP
Liu *et al.* [17]	Liu-7	42.22	39.92	42.51	43.11	42.81	41.22	47.21	47.01
FACIAL	FACIAL-7	44.55	42.30	48.11	46.99	43.22	44.04	49.54	47.50

Table 4 also summarizes the performance when the age groups are defined following Liu *et al.* [17]. It may be observed that the proposed splitting achieves better classification results in all the configurations: non-frontalized, frontalized, 60×60, and 150×150. The average accuracy increment is 2.36% and 4.74% in images with a resolution of 60×60 non-frontalized and frontalized, respectively. And the average accuracy increment is 1.62% and 1.41% in images with a resolution of 150×50 non-frontalized and frontalized, respectively.

4 Conclusions

In this work, we presented a methodology for age classification which uses frontalized images. We concluded that using face frontalization before feature extraction is beneficial. The method was tested on the FGnet Aging Database [14]. As it was demonstrated, the frontalization increases the age classification accuracy in cases with different image resolution, features (BIF and GOP) and age groups. In addition, a new age grouping based on the human aging process was proposed and evaluated.

Acknowledgments. This work has been funded in part by the Mexican National Council for Science and Technology (CONACYT) scholarship 328083 and by the UH Hugh Roy and Lillie Cranz Cullen Endowment Fund. The authors acknowledge the use of the Maxwell/Opuntia Cluster and the support of the Center of Advanced Computing and Data Systems at the University of Houston to carry out the research presented herein. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of the sponsors.

References

1. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.: Face verification across age progression using discriminative methods. IEEE Trans. Inf. Forensics Secur. **5**, 82–91 (2010)
2. Geng, X., Yin, C., Zhou, Z.H.: Facial age estimation by learning from label distributions. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 2401–2412 (2013)
3. Shu, X., Xie, G.S., Li, Z., Tang, J.: Age progression: current technologies and applications. Neurocomputing **208**, 249–261 (2016)

4. Kwon, Y., da Vitoria Lobo, N.: Age classification from facial images. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, pp. 762–767 (1994)
5. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Proceedings of European Conference on Computer Vision, Freiburg, Germany, vol. 2, pp. 484–498 (1998)
6. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. Pattern Recogn. **29**, 51–59 (1996)
7. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.: A study of face recognition as people age. In: Proceedings of IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, pp. 1–8 (2007)
8. Guo, G., Mu, G., Fu, Y., Huang, T.: Human age estimation using bio-inspired features. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, pp. 112–119 (2009)
9. Geng, X., Zhou, Z.H., Smith-Miles, K.: Automatic age estimation based on facial aging patterns. IEEE Trans. Pattern Anal. Mach. Intell. **29**, 2234–2240 (2007)
10. Kilinc, M., Akgul, Y.S.: Human age estimation via geometric and textural features. In: Proceedings of International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Rome, Italy, pp. 531–538 (2012)
11. Levi, G., Hassner, T.: Age and gender classification using convolutional neural networks. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops, Boston, Massachusetts (2015)
12. Mirzaei, F., Toygar, O.: Facial age classification using subpattern-based approaches. In: Proceedings of International Conference on Image Processing, Computer Vision, and Pattern Recognition, Las Vegas, NV (2011)
13. Liu, K.H., Yan, S., Kuo, C.C.J.: Age group classification via structured fusion of uncertainty-driven shape features and selected surface features. In: Proceedings of IEEE Winter Conference on Applications of Computer Vision, Steamboat Springs, CO, pp. 445–452 (2014)
14. Lanitis, A.: FG-NET aging database face and gesture recognition (2012). http://grail.cs.washington.edu/aging/FGNET.zip. Accessed 21 Oct 2016
15. Kakadiaris, I.A., Toderici, G., Evangelopoulos, G., Passalis, G., Chu, D., Zhao, X., Shah, S.K., Theoharis, T.: 3D–2D face recognition with pose and illumination normalization. Computer Vision and Image Understanding (2016, in press). http://dx.doi.org/10.1016/j.cviu.2016.04.012
16. Armstrong, T.P.: The Human Odyssey: Navigating the Twelve Stages of Life. Sterling, New York (2008)
17. Liu, K.H., Yan, S., Kuo, C.C.: Age estimation via grouping and decision fusion. IEEE Trans. Inf. Forensics Secur. **10**, 2408–2423 (2015)
18. Kakadiaris, I.A., Passalis, G., Toderici, G., Murtuza, M.N., Lu, Y., Karampatziakis, N., Theoharis, T.: Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach. IEEE Trans. Pattern Anal. Mach. Intell. **29**, 640–649 (2007)
19. Toderici, G., Passalis, G., Zafeiriou, S., Tzimiropoulos, G., Petrou, M., Theoharis, T., Kakadiaris, I.: Bidirectional relighting for 3D-aided 2D face recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, pp. 2721–2728 (2010)
20. Wang, H.L., Wang, J.G., Yau, W.Y., Chua, X.L., Tan, Y.P.: Effects of facial alignment for age estimation. In: Proceedings of International Conference on Control Automation Robotics Vision, Singapore, pp. 644–647 (2010)

21. Han, H., Otto, C., Liu, X., Jain, A.: Demographic estimation from face images: human vs. machine performance. IEEE Trans. Pattern Anal. Mach. Intell. **37**, 1148–1161 (2014)
22. Riesenhuber, M., Poggio, T.: Hierarchical models of object recognition in cortex. Nat. Neurosci. **2**, 1019–1025 (1999)
23. Guo, G., Wang, X.: A study on human age estimation under facial expression changes. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, pp. 2547–2553 (2012)
24. Zhang, C., Guo, G.: Age estimation with expression changes using multiple aging subspaces. In: Proceedings of IEEE International Conference on Biometrics: Theory Applications and Systems, Washington DC, pp. 1–6 (2013)
25. Zhang, C., Guo, G.: Exploiting unlabeled ages for aging pattern analysis on a large database. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops, Portland, OR, pp. 458–464 (2013)
26. Webb, A.: Statistical Pattern Recognition, 2nd edn. Wiley, New York (2002)
27. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: Proceedings of Annual Workshop on Computational Learning Theory, Pittsburgh, PA, pp. 144–152 (1992)
28. Chang, C.C., Lin, C.J.: LibSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**, 1–27 (2011)