HUMAN AGE ESTIMATION USING ENHANCED BIO-INSPIRED FEATURES (EBIF)

Mohamed Y.El Dib and Motaz El-Saban
{mohamed.y.eldib,motaz.elsaban}@gmail.com
Faculty of Computers and Information, Cairo University, Cairo, Egypt

ABSTRACT

The Aging process is a non-reversible process, causing human face characteristics change with time as hair whitening, muscles drop and wrinkles. Recently, age estimation from facial images has emerged as a hot research area. One of the most successful works is based on biologically inspired features (BIF). In this paper we extend BIF by incorporating fine details facial features, automatic initialization using active shape models and analyzing a more complete facial area by including the forehead details. Besides, we combine regression-based and classification-based models and test them experimentally on standard datasets showing the superiority of our proposed algorithm (extended BIF – EBIF) over the state-of-the-art.

***Index Terms*—**Automatic age estimation, Facial feature extraction, ASM.

1. INTRODUCTION

Automatic age estimation from facial images has recently emerged as a technology with multiple interesting applications such as targeted advertisement, law enforcement for certain types of drug usages and entertainment scenarios. The age estimation problem is particularly challenging as age depends on many factors, some of them are visual and many others are non-visual such as ethnic background, living style, working environment, health condition and social life. Even the visual features that can help in estimating age such as people’s facial features are affected by pose, lighting and imaging conditions. Due to its numerous challenges, there has been relatively little work, to date, concerning automatic age estimation, despite having potential useful applications.

In this paper, we present a fully automated age estimation approach based on the recently proposed and promising work using bio-inspired features (BIF) [6]. Our main contributions include (1) automatic localization of facial landmarks for the input faces for the first time with the BIF method [6] using Active Shape Models (ASM) [7] ;(2) utilizing micro facial features to reveal facial details in the forehead leading to a significant increase in the overall accuracy over the state-of-the-art and (3) increasing the number of shape features through inclusion of the forehead shape features.

The paper is organized as follows. In section 2; we present the related work. Section 3 gives an overview of the proposed age estimation framework. Section 4 presents a detailed implementation for our proposed algorithm. Section 5 describes our results on the publicly available FG-NET face aging database [4] and the MORPH database [14]. Finally, section 6 draws some conclusions and discusses our future work.

2. RELATED WORK

There are several existing works on the facial aging progress, originating from psychological and biological studies. However, most of them aim at simulating the aging effects on human faces [1] (i.e. simulate how the face would look like at a certain age), which is the inverse procedure of age estimation. Age estimation approaches fall into two categories: a) classification-based [2][5][3][11] and 2) regression-based [11][9][6][12][13][1]. An example classification-based work is by [5] where an anthropometric model has been for age classification based on cranio-facial development theory and skin wrinkle analysis, with human faces finally classified into three groups: babies, young adults and senior adults. Geng et al. [2] which defines The AGingpattEn Subspace (AGES) method. AGES learns a subspace representation of aging sequences and estimating age by projecting the test face into the subspace.

In the second category of works, age estimation is viewed as a regression problem where facial features are extracted by the active appearance models (AAMs) [8] that incorporate shape and appearance information together. An input face image is then represented by a set of fitted model parameters. The regression coefficients are estimated from training data with an assumption of the regression function such as a quadratic model (QM) [3]. Yan et al. [12], [13] also dealt with age uncertainty by formulating a semi-definite programming problem [13] or utilizing an EM-based algorithm [12] where they adopted traditional discriminative methods, using image intensities directly or other features exhaustively extracted from images. Suo et al. [1] presented compositional and dynamic models which decompose a face into parts and representing the face aging process dynamics as a first-order Markov chain on sparse graphs. Guo et al. [6] investigated the biologically inspired features (BIF) for human age estimation from faces using a regression-based approach BIF has shown very promising results compared to the literature.

3. SYSTEM OVERIEW

The proposed algorithm for age estimation is divided into five steps. First the facial landmarks for the face image are detected automatically using ASM (as opposed to the case of BIF [6] where this step was manually performed). The image is cropped to just the area covering a fixed number of points generated from the ASM step (several numbers of points was tested experimentally). Then, the cropped face is filtered by a set of Gabor functions at different orientations and scales. The filtered outputs undergo a feature dimensionality reduction step by just keeping the maximum (MAX) and standard deviations (STD) of the Gabor filtered outputs. Finally, both a classification-based and regression-based models were used in the training phase (SVM and SVR in this case) to produce the final age model estimator. The complete algorithm block diagram is illustrated in Fig 1.

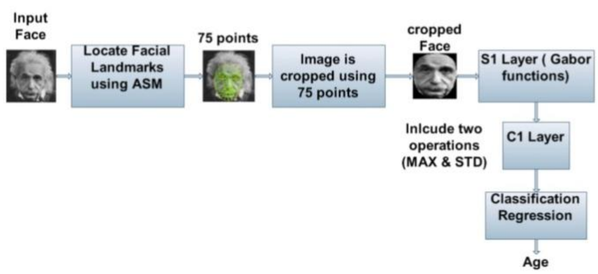


Figure 1. Age estimation framework

4. DETAILED IMPLEMENTATION

4.1 Accurate face localization

Images of the face can demonstrate a wide degree of variation in both shape and texture. Appearance variations are caused by differences between individuals, the deformation of an individual face due to changes in expression and speaking, and variations in lighting. Typically we would like to locate the features of a face in order to perform age estimation on the detected shape features.

In this step, we aim at accurately localizing the facial region to extract features only from the relevant parts of the input image. This localization step was manually performed in [6] which limit its practical usage. In this work, we explore the use of Active Shape Models [7] for the automatic localization of facial landmark points; ASM has two main stage, namely training and fitting. In the training stage, we manually locate landmark points for hundreds of images [10] in such a way that each landmark represents a distinguishable point present on every example image. We tried 75 and 68 points were provided by [10][4] respectively. We use 75 points to cover the whole face including the forehead. Then, we build a statistical model of shape based on these annotated images.

After shape model building, points on the incoming face image are fitted. First, we detect the face in the input image. With the detected face position, we initialize the shape points and do image alignment fitting to get the detected shape features automatically. Finally the input image is cropped to just the area covered by the ASM fitted landmark points. In this work, we use 75 points for landmark points as opposed to 68 in [6] as this includes features from the forehead region which contributed in an increase of accuracy as shown in the reported experiments, The difference between using 68 and 75 landmark points is illustrated in Fig 2.

68

75

Figure 2. 68 and 75 points samples from FG-NET

4.2 Texture-based face representation

Texture features have proven to be distinctive for the task of age estimation from facial images [6]. Particularly, the use of Gabor has proven to be successful as in [6, 15]. We follow the same approach here. The cropped image, output from the ASM model fitting, is filtered by a family of Gabor functions with 8 orientations and 16 scales. Gabor functions for a particular scale (sigma) and orientation (theta) are described by the equations:

$$G(x, y) = \exp\left(-\frac{(X^2+\gamma^2 Y^2)}{2\sigma^2}\right) \times \cos\left(\frac{2\pi}{\lambda} X\right) \qquad (1)$$

Where $\boldsymbol{X} = x\cos\theta + y\sin\theta$ and $\boldsymbol{Y} = -x\sin\theta + y\cos\theta$ and θ varies between 0 and π. The parameters λ(wavelength), σ(effective width), and γ(aspect ratio) =0.3 are based on the work in [6],with the addition of a filter with a smaller size (3x3) capable of revealing facial details in the forehead area (with parameters shown in Table 1)as shown in Fig 3.

Feature dimensionality reduction parameters			Gabor filters parameters		
Scale band S	Pool. grid	Overlap Δ_S	Filter size s	σ	λ
Band 1	6 x 6	3	3×3	1.5	2.3

Table.1 Feature dimensionality reduction and Gabor filters parameters.

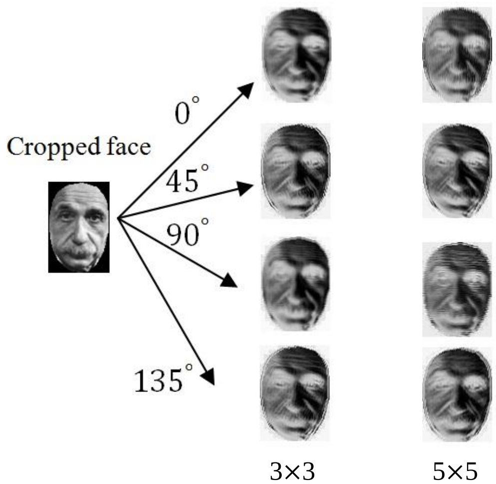


Figure 3.Gabor filtered results at band 1 (two scales with filer sizes 3×3 and 5×5) at four orientations. Note that in the case of 3x3 filter forehead features start to be more visible.

Gabor filtered outputs can serve as candidate features for the age estimation problem. However, there are of a very high dimension leading to difficulties in training. Besides there are redundancies in the Gabor filter outputs. Hence, a usual adopted scheme is to summarize the outputs of the Gabor filters using some statistics measure. Here, we adopt the ones used in [6] and proven to work quite well, namely the maximum "**MAX**" and standard deviations "**STD**" with a variation on the MAX definition by avoiding image subsampling to keep the local variations which might be important for characterizing the facial details (e.g., wrinkles, crease and muscles drop).

4.2. Classification or Regression

Age estimation can be treated as a classification problem, when each age is considered as a class label. Alternatively, age estimation can be treated as a regression problem where each age is considered as a regression value. In our experiments, we use both SVR and SVM methods for age estimation on the FG-NET [4] and MORPH [14] standard databases. The RBF SVR can address the three limitations of the traditional quadratic regression model [16] :(1) the simple quadratic function may not model properly the complex aging process, especially for a large span of years, e.g., 0-70.; (2) the least square estimation is sensitive to outliers that come from incorrect labels in collecting a large image database; and (3) the least square estimate criterion only minimizes the empirical risk which may not generalize well for unseen examples [11][16].

5. EXPERIMENTS

We used two measures to evaluate age estimation performance: (1) Mean Absolute Error (MAE); (2) Cumulative score (CS).The MAE is defined as the average of the absolute errors between the estimated ages and the ground truth ages, MAE$=\sum_{k=1}^{N}\frac{|\hat{l}_k-l_k|}{N}$ where l_k is the ground truth age for the test image k, $\hat{l}_k$ is the estimated age, and N is the total number of test images. The cumulative score CS(j) is defined as $\frac{N_{e\le j}}{N}\times 100\%$ where $N_{e\le j}$ is the number of test images on which the age estimation makes an absolute error no higher than *j* years.

5.1. Datasets

We use the FG-NET [4] aging database which is publically available and the MORPH [14] database to evaluate the performance of our approach. The FG-NET contains 1002 face images of 82 subjects with ages ranging from 0 to 69 with large variation of lighting, pose, and expression. While, The MORPH database contains 1690 face images of 515 subjects ranging in age from 15 to 68 years for men and women of various ancestry groups.

5.2. Results

A Leave-One-Person-Out (LOPO) test strategy is used on the FG-NET database. MORPH is used only as a test set on the trained FG-NET database (this same strategy was used in [2]). Each image is cropped using the 75 shape feature points and resized to 59×80, which leads to a total of 6100 features per image. We use both SVM-based classifier and a support vector regressor (SVR). We build six SVR models (one for each range as shown in Table 2) and one SVM model using the experimentally selected parameters provided in Table 2. Using SVR or SVM separately cannot adequately estimate age because of the diversity of the aging process across different ages [11]. Hence, we combine SVR and SVM models by selecting which model to use over each age group, based on MSE results over the training set.

SVR	SVM
C=1,5,30,60 (ages<30), C=240,1000(ages >30)	C =1000
$\gamma = 0.000308$, €=0.1, $e = 0.002$	

Table.2 SVR and SVM model parameters.

We first measure the effect of utilizing detailed shape features by using 75 points as opposed to 68 as in [6]. Experimental results are provided in Table 5 which shows the superiority of a more detailed shape representation over FG-NET.

The second set of experiments aimed at measuring MAE and CS scores of the proposed method against state-of-the-art. The MAE results on the FG-NET are shown in Table 3 for different age ranges. The MAE of our method (EBIF) is 3.17 which is significantly smaller than the 4.77 of the state-of-the-art BIF method [6] on FG-NET. CS curves are similarly shown in Fig 4 for the EBIF and BF methods.
For the MORPH database, (which was tested in [2]), the authors used 433 images which represent only Caucasian descent, we used the same images for consistency .The MAE of EBIF on MORPH is 4.11 compared to AGES [2]

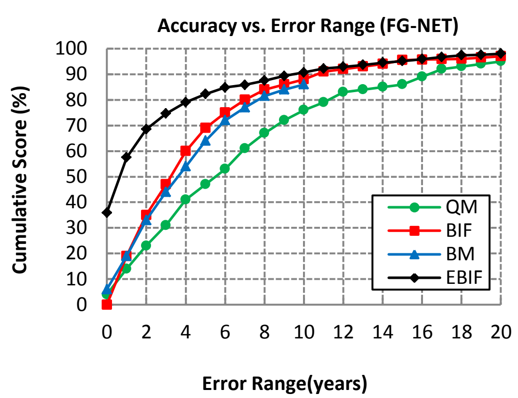
8.83. Finally, an aggregate comparison of MAE values is shown in Table 4 for both e FG-NET and MORPH databases.

Range	#img.	EBIF	BIF[6]	RUN[13]	QM[3]
0-9	372	1.67	2.99	2.51	6.26
10-19	338	1.09	3.39	3.76	5.85
20-29	144	3.31	4.30	6.38	7.10
30-39	79	7.87	8.24	12.51	11.56
40-49	46	13.02	14.98	20.09	14.80
50-59	15	18	20.49	28.07	24.27
60-69	8	26.25	31.62	42.50	37.38
Total	1002	**3.17**	4.77	5.78	7.57

Table.3 MAE (years) at different age groups on FG-NET.

6. CONCULSION AND FUTURE WORK

In this paper, we presented a human age estimator based on the bio-inspired features-based method [6]. We have combined BIF h the Active Shape Model (ASM) for initialization. Besides, we have experimented with the extraction of finer facial features as opposed to [6] and shown experimentally the superiority of the proposed contributions. Evaluated on the FG-NET and MORPH benchmark databases, our algorithm achieved high accuracy in estimating human ages compared to published methods. We also tested the proposed algorithm on MORPH database to show its generalization capabilities. As far as the future work is concerned, we aim at exploring the imaginary and magnitude parts in Gabor functions and incorporating other models like the Active Appearance Model with the biologically inspiration features.



FG-NET.

Figure 4. Cumulative scores for Quadratic Model, BIF, BM and our algorithm at absolute error levels from 0 to 20 years

FG-NET	
75 shape feature points	68 shape feature points
3.17	3.81

Table.5 MAE (years) measures for 75 and 68 shape feature points

Method	FGNET	MORPH
WAS [2]	8.06	9.32
AGES [2]	6.77	8.83
QM [3]	6.55	–
$AGES_{lda}$[2]	6.22	8.07
RUN [13]	5.78	–
BM [12]	5.33	–
LARR [11]	5.07	–
PFA [8]	4.97	–
RPK [6]	4.95	–
BIF [6]	4.77	–
EBIF(Ours)	**3.17**	**4.11**

Table.4 MAE (years) comparisons

7. REFERENCES

[1] J. Suo, S. Zhu, S. Shan and X. Chen, "A Compositional and Dynamic Model for Face Aging" IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009

[2] 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.

[3] A. Lanitis, C. Draganova, and C. Christodoulou.”Comparing different classifiers for automatic age estimation”. IEEE Trans. on SMC-B, 34(1):621–628, 2004.

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

[5]Y. Kwon and N. Lobo.“Age classification from facial images”. Computer Visionand Image Understanding, 74(1):1–21, 1999.

[6]Mu, G.W., Guo, G.D., Fu, Y., Huang, T.S. “Human age estimation using bio-inspired features”,CVPR09(112-119).

[7]T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham.“Active Shape Models | theirTraining and Application”. CVIU 61, 38{59 (1995)

[8] Y. Fu, Y. Xu, and T. S. Huang, “Estimating human ages by manifold
analysis of face pictures and regression on aging features,” in Proc. IEEE Conf. Multimedia Expo., 2007, pp. 1383–1386.

[9] A. Lanitis, C. Taylor, and T. Cootes, “Toward automatic simulation of
aging effects on face images,” IEEE Trans. Pattern Anal. Mach. Intell.,
vol. 24, no. 4, pp. 442–455, Apr. 2002.

[10] Y. Wei “Research on Facial Expression Recognition and Synthesis”, Master Thesis, 2009

[11] GG. Guo, Y. Fu, T. Huang, and C. Dyer. Locally adjusted robust regression for human age estimation. In *IEEE WACV*, 2008.

[12] S. Yan, H. Wang, T. S. Huang, and X. Tang.“Ranking with uncertain labels”.InIEEE conf. on Multimedia and Expo, pages 96–99, 2007.

[13] S. Yan, H. Wang, X. Tang, and T. Huang.”Learning auto-structured regressor from uncertain nonnegative labels”.In IEEE conf. on ICCV, 2007.

[14] K. Ricanek and T. Tesafaye, “MORPH: A longitudinal image database of normal adult age-progression,” in Proc. the 7th Int’l Conf. Automatic Face and Gesture Recognition, Southampton, UK, 2006, pp. 341–345.

[15] J. Mutch and D. Lowe. Object class recognition and localization using sparse features with limited receptive fields. In CVPR pages 11–18, 2006.

[16] V. N. Vapnik, Statistical Learning Theory, John Wiley, 1998