

A Novel Age Classification Method Using Morph-Based Models

Asuman Günay Yılmaz¹ and Vasif Nabiye²

¹Department of Computer Technologies, Karadeniz Technical University, Turkey

²Department of Computer Engineering, Karadeniz Technical University, Turkey

Abstract: Automatic facial age classification and estimation is an interesting and challenging problem, and has many real world applications. The performances of the classification methods may differ depending on the selected training samples. Also using large amount of training samples makes the classification systems more complex and time consuming. In this paper, a novel and a simple age classification method using morph-based age models is presented. The age models representing the common characteristics of age groups are produced using image morphing method. Then age related facial features are extracted with Local Binary Patterns. In the classification phase, ensemble of distance metrics is used to determine the closeness of the test sample to age groups. Then, the results of these metrics are combined with Borda Count voting method to improve the classification performance. Experimental results using the Face and Gesture Recognition Research Network (FGNET) and Park Aging Mind Laboratory (PAL) aging databases show that the proposed method achieves better age classification accuracy when compared to some of the previous methods.

Keywords: Age classification, image morphing, local binary patterns, borda count voting.

Received January 27, 2016; accepted June 13, 2016

1. Introduction

With the increasing need of automatic recognition systems, the researches on facial image processing have received considerable interest in recent decades. Face detection, face recognition, gender classification and facial expression recognition are the research topics that have been studied by many researchers in this area [17]. Facial age estimation is a relatively new topic and the interest in this topic has significantly increased because it has many real world applications. For example, an automatic age estimation system can prevent under ages from accessing cigarettes, alcohol or obscene contents on websites. In addition, age specific target advertising, face recognition systems robust to age progression and age prediction systems for finding the lost children and criminals are important age estimation applications.

Facial age estimation is a multi-class classification problem because an age label can be seen as an individual class. This makes age estimation much harder than other facial image processing problems such as gender classification, face detection, etc. Besides, real world age progression displayed on faces is varied and personalized. Aging process of a person is affected by the genetics, race, living styles, eating and drinking habits, climate, facial expressions, etc. [2]. Therefore, it is very difficult to determine the type of facial features directly represents the age. Moreover, the accuracy of age estimation systems are insufficient, even the human skills about age estimation is limited. The lack of proper large data set including the

chronological image series of individuals is another drawback in age estimation systems. Despite the restrictions mentioned above, age estimation has been demanding growing interest recently because of its crucial application areas. But it can be seen from previous classification methods that there is no consensus on the boundaries of the age groups [3, 4, 6, 7, 8, 9, 11, 12, 13, 15, 22, 23, 24]. Furthermore, the different methods tested on the same database have used different boundaries for age groups [7, 15, 23]. The classification performances of these methods are confusing, because the classification accuracy of a method can be satisfactory for one age group design, but may not be for another age group design. In addition, the theoretical explanation about the determination of the boundaries of age groups is not mentioned in previous works. Standardizing the age groups will facilitate the performance evaluation of the classification methods with each other.

In this respect, the boundaries of age groups in this study are designated according to developmental psychologist and psychoanalyst Erik Erikson's stages of psychosocial development theory [5]. According to his theory there are 8 stages through which a healthily developing human should pass from infancy to late adulthood. Each stage is characterized by a psychosocial crisis of person's two conflicting forces: biological forces and socio-cultural forces. The age ranges of these stages are 0-2 (Oral-Sensory), 2-4 (Muscular-Anal), 4-5 (Locomotor-Genital), 5-12 (Latency), 13-19 (Adolescence), 20-39 (Young Adulthood), 40-64 (Middle Adulthood) and 65 and

above (Late Adulthood). With the assumption of the facial texture does not change much in early years of human life, the age range of 0-12 years old is assigned as childhood in the proposed method. The other age ranges are used the same as the Erikson’s theory.

After setting up a theoretical basis for the boundaries of age groups, we consider it is necessary to produce age models for age groups instead of using large training datasets for classification. This will also make the recognition system simple. Based on this idea, the aim of this study is to produce age models representing the common characteristics of age groups and age classification using these age models.

The general structure of the proposed age classification method is shown in Figure 1. The original facial images are preprocessed in order to adjust the head pose, the viewpoint and the size of the images. The age groups are designated according to Erik Erikson’s stages of psychosocial development theory. Then age models representing the common characteristics of age groups are produced using image morphing. In the feature extraction module, the global and Local Binary Pattern (LBP) histograms of the age models are generated and used as feature vectors in the classification module. In the testing phase, first the preprocessing and feature extraction steps are applied to the test sample. Then the distances between the feature vectors of the test sample and the age models of age groups are calculated using ensemble of distance metrics. Each distance metric makes a preference order among these age groups according to the distances, from minimum distance to the maximum one. The results of distance metrics are combined using Borda Count (BC) voting method to improve the classification performance.

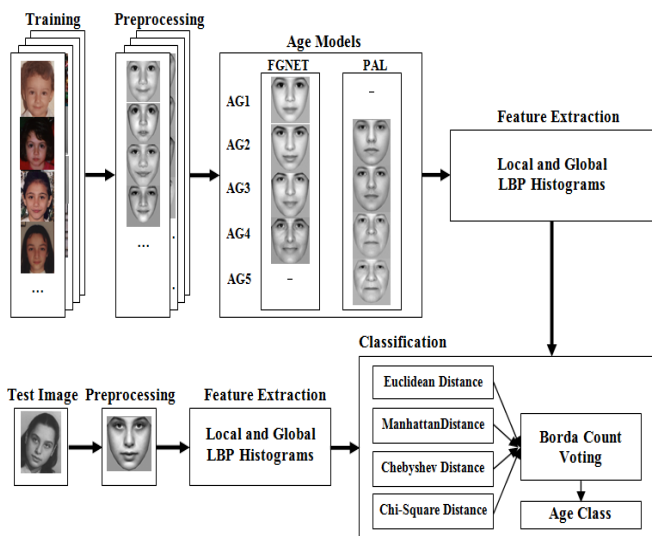


Figure 1. Proposed age classification system.

The rest of the paper is organized as follows. Some of the age classification techniques are introduced in section 2. The proposed age classification method is explained in section 3. In section 4 the experimental

results are reported and analyzed. Finally, the conclusions are outlined in section 5.

2. Related Work

The earliest paper published in the area of age classification was the work by Kwon and Lobo [12]. They computed six ratios of distances on frontal images to separate babies from adults and used wrinkle information to separate young adults from senior adults. They used a very small database containing only 47 images in their experiments and the infant identification rate is below 68%. Horng *et al.* [8] used geometric features and wrinkles to classify images into four age groups: babies (0-2), young-adult (3-39), middle-aged (40-59), and elderly (≥ 60). They extracted wrinkle information (the wrinkle density, the wrinkle depth and the average skin variance) from facial images using Sobel edge operator and classified images using a neural network. Their system performance is reported as 81.58% on a database including 230 facial images. But the age groups are not properly arranged which makes the extension of the results difficult to real applications.

Hayashi *et al.* [6] performed histogram equalization on the extracted skin regions to emphasize wrinkles and used Digital Template Hough Transform to detect wrinkles. A look up table containing the wrinkle distributions against appearance in ten-year spans is constructed and used for age classification. Their age classification accuracy is only 27% for regularly defined age groups. Their database contains 300 images greater than 15 years old. Txia and Huang [24] have used Active Appearance Models (AAMs) to locate 28 facial feature points on facial images. Then they used Sobel filter for extracting wrinkle information in forehead, eye corners, and mouth corners regions. They selected 251 (120 for train-131 for test) Caucasian descent facial images from MORPH (Craniofacial Longitudinal Morphological Face) database to classify images into four age groups (20-30, 31-40, 41-50, 51-60) and achieved 72.52% accuracy.

Dehshibi and Bastanfard [4] extracted wrinkles from the wrinkle areas (the forehead, two eye corners and two cheeks) using the Canny edge detector. Wrinkle densities of these regions and geometrical features are used for age group classification. They used a neural network to classify the faces taken from Iranian Face Database into the age groups under 15, 16-30, 31-50 and greater than 50 and achieved 86.64% accuracy. Chen *et al.* [3] have classified facial images into child, elder and adult age groups using 180 images from Face and Gesture Recognition Research Network (FGNET) and Face Recognition Technology (FERET) databases. They reported the classification accuracy as 87.8% using the edge features of facial images. But the number of images used in the experiments is limited.

The performance of the system on all images of FGNET and FERET databases is not mentioned.

Higashi *et al.* [7] have used Gabor magnitude images with Local Directional Pattern operator to classify facial images into one of the four age groups: 0-9, 10-19, 20-34 and 35-64. The classification performance of the system is reported as 60.69%. They used only 261 facial images from the FGNET database in the experiments. Tonchev *et al.* [23] classified images into two age groups: 0-16 and 17-69. Their system is based on subspace projection using principal component analysis and classification using kernel Support Vector Machines (SVM). They achieved 77.73% classification accuracy on FGNET database but only two age groups are used in this method.

Liu *et al.* [15] used AAM to extract facial features and SVM to classify images into five age groups: 0-2 (baby), 3-10 (child), 11-18 (young), 19-39 (adult) and 40-69 (middle and old age). They have chosen 25 images per each age group from FGNET database to test the classification performance and reported 79.2% accuracy. Kalansuriya and Dharmaratne [9] used wrinkle information and geometrical ratios with neural networks to classify facial images taken from FERET and FGNET databases into four age groups: 8-13, 14-25, 26-45 and 46-63. Their classification performance is 74.38%. In the experiments they used 550 images to train the neural network. In the testing phase they used 200 images to evaluate the system performance. Kohli *et al.* [11] extracted feature vectors from images using AAMs and used ensemble of classifiers trained on different dissimilarities to distinguish between child/teen-hood (0-21) and adulthood (22-69). The classification accuracy on FGNET database for two age classes is 84%.

Li *et al.* [13] used multi feature weighted decision fusion for age classification. They have used LBP histograms of Gabor images generated by Gabor filter bank and the ratios between the wrinkles and skin areas as aging features. They simulated experiments on 1600 images taken from the FGNET and self-build database. The classification accuracy of the system is reported 85.75% for four age groups: 0-17, 18-45, 46-60, >60. Similarly Sai *et al.* [22] have used LBP histograms of Gabor images for feature extraction. Then extreme learning machine is used to classify images into 4 age groups: 0-10, 11-19, 20-60 and >60. They have used 383 images taken from FGNET and Park Aging Mind Laboratory (PAL) databases with good quality and contain low degree of expression and pose, and achieved 82% accuracy. Liu *et al.* [14] presented multi-stage learning system. They extracted global features from the whole face and local features from facial components and classified faces into different age groups having different age ranges. Then these decisions are fused using different fusion schemes. Ren and Li [21] used gradient features and Gabor features for age estimation. Then feature

selection is performed to choose meaningful regions on facial images. The images are classified into the age groups of 0-15, 16-40, 41-59 and ≥ 60 .

3. Proposed Method

The proposed age classification method consists of preprocessing, age model generation, feature extraction and age classification modules. These modules are explained in detail in the following subsections.

3.1. Preprocessing

As the variations in illumination, head pose, facial expression, background, resolution etc. in facial images can adversely affect the recognition rate of the system, the preprocessing is performed for all images in the training and test sets. In order to adjust the head poses and viewpoints and to produce frontal view images, image morphing method is used in this module. Although the sizes of the images are also adjusted after image morphing, the eye center positions are still different from one image to another. For this reason the images are normalized which means that they are aligned according to eye center positions and cropped to the same size.

3.1.1. Image Morphing

In this study, image morphing method is used in the preprocessing module to produce frontal view images and in the age model generation module to produce age models for age groups. Morphing is a technique used to transform a source image into a target image smoothly which generates a sequence of images.

In order to morph one image to another, some control points of each source image must be calculated. Each control point specifies an image feature. The images are partitioned into non-overlapping regions based on these control points. Also they are used to interpolate the positions of the feature points across the morphing process. The choice and number of control points determined the accuracy of the morphing process [10].

Suppose we have two source images of which the feature points are detected and partitioned into regions. To make an intermediate image between two source images, the weightings are set to α and $(1-\alpha)$, respectively. For a feature point A in the first and the corresponding feature point B in the second image, linear interpolation can be used to generate the position of the new feature point F presented in Equation (1).

$$F = \alpha A + (1-\alpha)B \quad 0 \leq \alpha \leq 1 \quad (1)$$

The new feature point F is used to construct a point set which partitions the image in another way different from the two source images. First and second images are warped such that their feature points are moved to the same new feature points, and thus their feature

points are matched. For this purpose affine transformation is used.

In the warping process, coordinate transformations are performed for all regions respectively. After performing coordinate transformations for source images, the feature points of these images are matched. To complete the morphing process, cross-dissolving is done as the coordinate transforms are taking place. Cross-dissolving is described by the following Equation (2),

$$I_{morph}(x, y) = \alpha I_1(x, y) + (1 - \alpha) I_2(x, y), 0 \leq \alpha \leq 1 \quad (2)$$

Where I_1, I_2 are the warped source images, and I_{morph} is the morphing result. This operation is performed pixel by pixel, and each of the color components of RGB is dealt with individually.

In the preprocessing phase, image morphing technique is used to adjust the head poses and viewpoints of facial images as follows. The original facial images are first labeled with 68 landmark points as shown is Figure 2-a. In order to adjust the head poses and viewpoints of the images and produce frontal view images, original facial images are flipped in the right direction and also labeled with 68 landmark points Figure 2-e. To find the optimal triangulation for morphing process Delaunay triangulation is used Figure 2-b, and 2-d). Then intermediate (mean shape) triangulation is obtained with $a=0.5$ Figure 2-c. Finally two source images are warped to the mean shape so that their feature points are matched (Figure 2-f, 2-h). After cross-dissolving the pixel values of two warped images with $a=0.5$ the frontal view image is produced as shown in Figure 2-g.

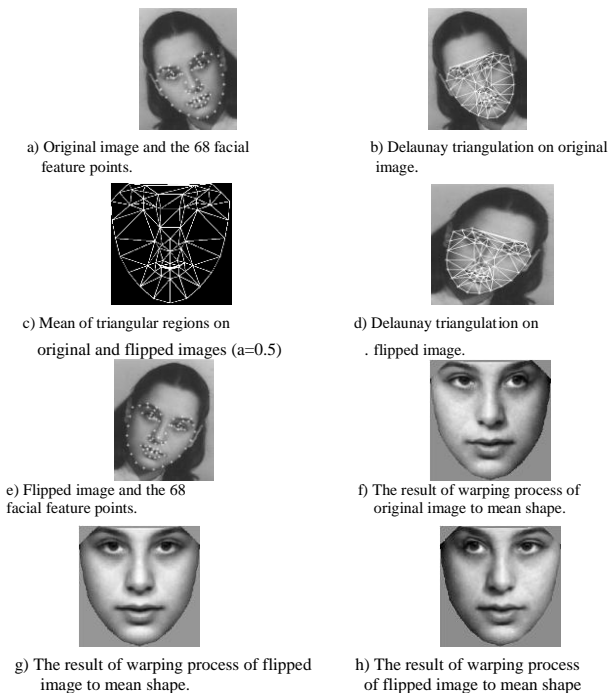


Figure 2. Image preprocessing.

3.1.2. Image Normalization

The head poses and viewpoints of the facial images are adjusted after image morphing process. But the eye center positions are still different from one image to another. For this reason the images are aligned according to eye centers, scaled and cropped to the size of 110x110. This size is selected considering the face anthropometry [20]. Furthermore, the 21 out of the 68 control points are enclosed the edges of the face in the morphing process, so the non-facial areas such as background, clothes are masked and only the facial region is used in the feature extraction step.

3.2. Age Models

The performances of the classification methods may differ depending on the selected training samples. Also using large amount of training samples to improve system performance makes the classification systems more complex and time consuming. In order to reduce the effects of different and large training datasets and make a simple classification system, producing age models representing the mean information of age groups is proposed in this study.

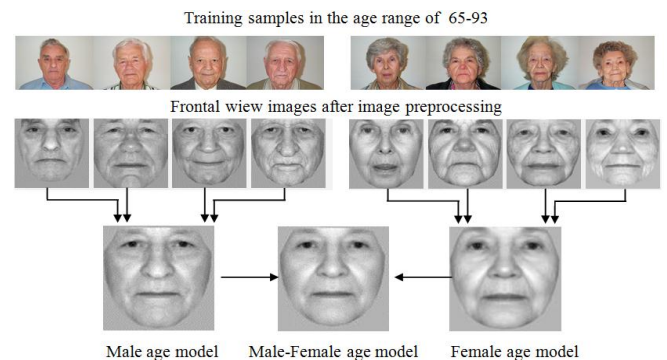


Figure 3. Age model generation process for ≥ 65 age group of PAL database.

In the study the borders of the age groups are determined as 0-12, 13-19, 20-39, 40-64 and 65-93. The border of the last age group is 65-93, because the PAL database contains faces ranging from ages 18 to 93. Age variation has the largest values in the age groups of 20-39, 40-64 and 65-93, that is there are 20, 25 and 29 different age labels in these age groups respectively. In the other age groups the age variation is less than these age groups. Therefore, the average age variation in age groups is calculated approximately 24 by taking the mean of different age labels in these age groups. Then age groups are represented with 8 images (4 male-4 female) randomly selected at 3 years intervals in age, since 3 years in age will not cause too much change in facial appearance [1]. The system performance can be affected due to the selection of the images. In the study age model generation process is performed on random images. Then one male, one female and one male-female age models are produced for each age group using image morphing technique as

shown in Figure 3. Age models produced for the age groups of FGNET and PAL databases are given in Figure 4.

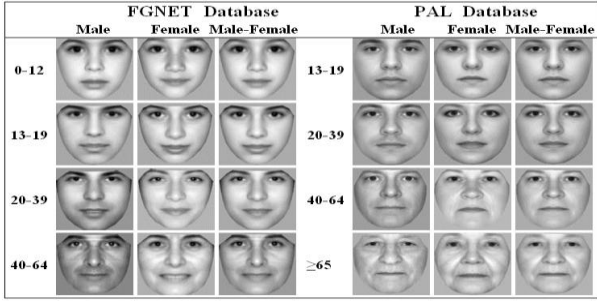


Figure 4. Male, female and male-female age models for age groups (AGs) of FGNET and PAL databases.

3.3. Feature Extraction with Local Binary Patterns

LBPs are fundamental properties of local image texture and their occurrence histogram is a powerful texture descriptor [19]. Original LBP operator assigns a label for every pixel of the image, by means of thresholding the center pixel with its 3x3 neighborhood and considering the results as a binary number as illustrated in Figure 5. Then the occurrence histogram of these labels is used as a texture feature.

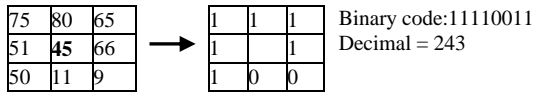


Figure 5. The basic LBP operator.

This structure was extended to use circular neighborhoods of different sizes. By changing the radius (R) and the number of neighbors (P), the multi-scale texture analysis is achieved. P controls the quantization of the angular space whereas R determines the spatial resolution of the operator. In texture description all LBP codes are not used. The uniform patterns which contains at most two bitwise transitions from 0 to 1 or vice versa when the binary pattern is considered circular, account for nearly 90% of all patterns when using (8,1) neighborhood [19].

To label any image $I(x,y)$ with LBP operator, Equation (3) is used for every pixel of $I(x,y)$.

$$LBP_{P,R}(x_c) = \sum_{p=0}^{P-1} u(x_p - x_c) 2^p \quad u(y) = \begin{cases} 1, & y \geq 0 \\ 0, & y < 0 \end{cases} \quad (3)$$

In Equation (3) x_c is the center pixel, x_p represents one of his P neighbors, R is the radius and $u()$ is the step function. After labeling the input image $I(x,y)$ with $LBP_{P,R}$ operator, histogram of the labeled image is produced with the following equation.

$$H_i = \sum_{x_c \in I(x,y)} f\{LBP_{P,R}(x_c) \in U(i)\} \quad i = 0,1,\dots,n-1 \quad (4)$$

$$f\{y\} = \begin{cases} 1, & y \text{ is true} \\ 0, & y \text{ is false} \end{cases}$$

In this Equation, n is the number of uniform patterns, $U(i)$ is the vector keeps uniform patterns produced by (P, R) neighborhood. This histogram contains information about the distribution of the local micro-patterns like edges, spots, flat areas, over the whole image. For an efficient representation and also consider shape information, local LBP histograms can be used. For this purpose image can be divided into regions and for every region (R_j) local histograms are produced as follows.

$$H_{i,j} = \sum_{x_c \in R_j} f\{LBP_{P,R}(x_c) \in U(i)\} \quad i = 0,1,\dots,n-1 \quad (5)$$

$$j = 0,1,\dots,m-1$$

In this Equation (5), $H_{i,j}$ is the i . value of the LBP histogram of j . region. These regional histograms are concatenated to build a global description of the image. According to the importance of information contained in the regions, some weights can be assigned for each region. Searching the importance of regions and assigning the weights of regions is also a difficult problem. In this study the weights are set empirically. For example, based on the observations, the regions under the eyes, cheeks and mouth corners represents the aging effects noticeably so the weights of these regions are higher than others. The local (H_L) and global (H_G) LBP histogram generation procedure is illustrated in Figure 6.

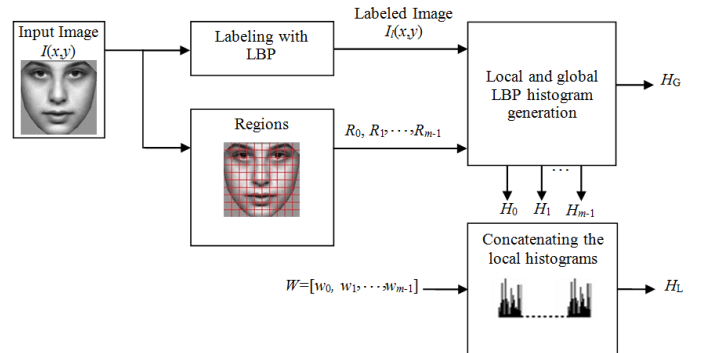


Figure 6. LBP histogram generation.

After LBP histogram generation, the classification is performed with comparing the distances between the LBP features of the test image and the age models of age groups.

3.4. Classification

Against the difficulty of estimating the parameters of sophisticated classifiers, a simple classification method is adopted in the classification module. Ensemble of distance metrics are used to calculate the distances between the feature vectors of test samples and age models. The classification performance of each metric differs from one age group to another. If the minimum distance classifier is used, the test sample is classified into the class that has minimum distance. But the distances between the first and second minimum may be very small which can make the classification result

incorrect. Basically each classifier put the possible class labels in an order according to calculated distances. The class with the minimum distance takes the first place and the next smallest class takes the second place, etc. To improve the classification performance of the system and evaluate the results of distance metrics together, BC voting method is used.

3.4.1. Distance Metrics

In this study the distances between the LBP histograms of test image and histograms of age models of age groups are computed with different distance metrics. These are Euclidean, Manhattan, Chebyshev and Chi Square statistics given below.

$$D_{Euclidean} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \tag{6}$$

$$D_{Manhattan} = \sum_{i=1}^n |p_i - q_i| \tag{7}$$

$$D_{Chebyshev} = \max_i (|p_i - q_i|) \tag{8}$$

$$\chi^2(P, Q) = \sum_i \frac{(p_i - q_i)^2}{p_i + q_i} \tag{9}$$

3.4.2. Borda Count Voting

The BC is a method of voting designed to pick a winner from a group of more than two options or candidates. The main advantage of the BC is that it takes into account voters’ preferences among the complete set of options, and therefore favors policies that have wide support. In general, if N is the number of candidates, each first-place vote is worth N points, each second-place vote is worth $N-1$ points, each N th-place (i.e., last-place) vote is worth 1 point. Whichever candidate receives the most points wins the election as shown in Figure 7.

Voter’s preferences					Candidate	BC Score
Points	Voter1	Voter2	Voter3	Voter4		
5	D	A	C	C	A	4+5+4+4=17
4	A	B	A	A	B	2+4+2+1=9
3	E	C	D	E	C	1+3+5+5=14
2	B	E	B	D	D	5+1+3+2=11
1	C	D	E	B	E	3+2+1+3=9

The winner by BC is A

Figure 7. BC score evaluation example.

In our method the voters are the distance metrics and the candidates are the age groups. If M is the number of distance metrics, N is the number of age groups and $r_{i,j}$ is the point assigned to the age group j by i th distance metric, the BC scores are calculated with the following equation.

$$BC_j = \sum_{i=1}^M r_{i,j}, \quad j=1, \dots, N \tag{10}$$

For a test sample t , the BC scores of all age groups are calculated and the test sample is classified to the age group AG_k with maximum BC score as follows.

$$t \in \{AG_k \mid \max_j (BC_j) = BC_{AG_k}\}, j=1, \dots, N \tag{11}$$

4. Experimental Results

To evaluate the performance of the proposed method, the FGNET and the PAL aging databases are used in the experiments. The FGNET aging database comprises of 1,002 images of 82 objects in the age range 0-69 years old. The images were obtained by scanning from real-life albums of different subjects. Therefore there are extreme variations in illumination, head pose, facial expression, background, resolution and noise from scanning. Besides, the subjects with respect to age are not uniformly distributed as shown in Figure 8-a. The PAL aging database was developed to be more representative of age groups across the lifespan [16]. It contains 580 individual faces ranging from ages 18 to 93. The images were captured under natural lighting conditions using a digital camera. The database includes various expressions such as smiling, sadness, anger or neutral faces. The data distribution of PAL database is given in Figure 8-b.

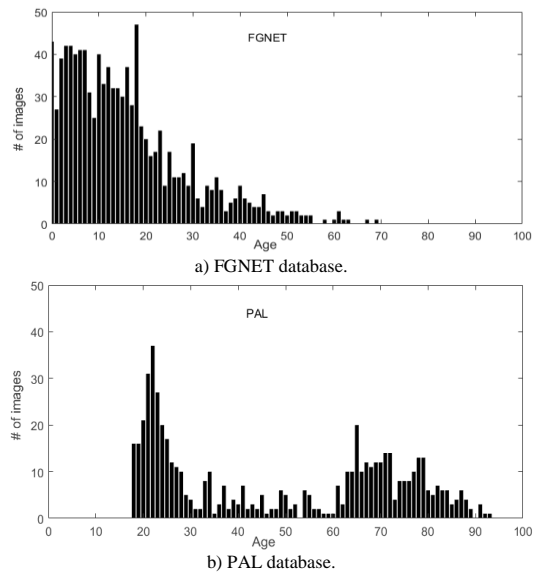


Figure 8. Data distribution according to age.

In the preprocessing and age model generation modules, control points representing the facial features of facial images are necessary for the morphing process. The FGNET database provides the positions of 68 facial feature points of all facial images. For the PAL database, 68 facial feature points are identified manually on all facial images.

In the classification phase the age ranges of age groups are designated according to Erikson’s stages of human development theory. These age groups are: 0-12 childhood, 13-19 adolescence, 20-39 young adulthood, 40-64 middle adulthood and ≥ 65 late adulthood. The age range distributions of images in FGNET and PAL databases are given in Table 1.

For all age groups, four male and four female facial images, selected randomly at 3 years intervals, are used to represent that age group. Then the age models (male, female and male-female) containing the average information of age groups are produced with image

morphing technique. The classification performances of the proposed system using two age models (male and female) for each age group and one age model (male-female) for each age group are graphically illustrated for FGNET and PAL databases in Figures 9 and 10, respectively. It can be seen from the figures that, representing the age groups with two age models gives better results for some age groups (40-64 (Female) in FGNET database and ≥ 65 (Male) in PAL database), but using a single age model for age groups generally improves the system performance. Selecting the samples properly in the age range of an age group provides the determination of common features of that age group. Thus, we can achieve classification performances similar to or better than previous works with a simpler system rather than a complex system trained with large datasets.

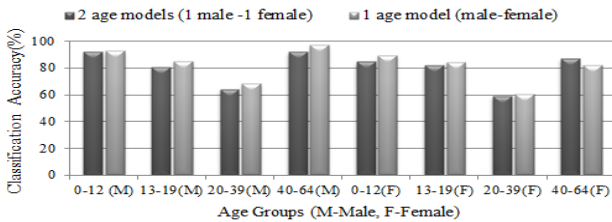


Figure 9. Classification accuracies of age models on FGNET database.

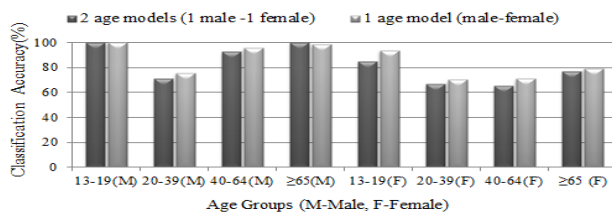


Figure 10. Classification accuracies of age models on PAL database.

After selecting the training samples and producing a single age model for each age group, the classification performance of the proposed method is evaluated using the rest of the database as testing dataset. In order to evaluate the classification performance of the method according to gender, the system is tested on male and female test samples separately and the results are given in Table 2. The classification results using all the testing samples are also given in this table. The results have shown that our system classifies the male samples more accurately than female samples. This result is supported by the studies investigating the effects of gender in facial age estimation [18]. Since the female samples include wider variations caused by cosmetics than male samples, this result is also a reflection of the reality in nature. While the classification accuracy on the male samples of FGNET database is 86.67%, the total performance decreases to 83.81%, as the classification accuracy on female subjects is 79.55%. For the PAL database, 85.48% of the male, 75.45% of

the female and 79.34% of all test samples are correctly classified with the proposed method.

The comparison between the previous age classification methods and the proposed method is given in Table 3. The columns of this table indicate the methods, the number of age groups, age group boundaries, the databases, the number of images used in the experiments, the percentage of the database used in the experiments and the classification accuracies of the methods, respectively. The classification accuracy of our system is 83.81% for FGNET and 79.34% for PAL database. It can be seen from Table 3 that classification performances of some previous works [3, 4, 11, 13] are higher than the proposed method. But, in these works small percentages of the databases are used in the experiments. Dehshibi and Bastanfard [4] achieved the classification accuracy of 86.64% by using only the 13.83% of the images in the database. Similarly, in [3] the classification performance of the method is reported 87.8% when using only 8.17% of the images in the experiments. Li *et al.* [13] used a database containing images taken from FGNET and self-build database and achieved 85.75% classification accuracy. Since we use the 100% of the images in the FGNET and PAL databases in our experiments, the results cannot be comparable with the works mentioned above. Although the percentage of the database used in experiments is 100% in [11], their 84% classification accuracy is achieved for only two age groups. Considering these conditions, the classification accuracies of the proposed outperforms the previous age classification methods.

The CPU time of the proposed method is 5.81 seconds for age model (male, female, male-female) generation for all age groups from the training set and 0.01863 milliseconds for feature extraction and classification of any test sample. These results are obtained on the system with CPU of Intel Core i5-2430M 2.40GHz, memory of 3 GB and better than the results reported in [4, 8, 22].

Table 1. Age range distribution of images in FG-NET and PAL databases.

	FGNET					PAL				
	0-12	13-19	20-39	40-64	≥ 65	0-12	13-19	20-39	40-64	≥ 65
Male	319	124	114	37	2	-	12	123	27	66
Female	162	105	109	30	-	-	20	111	64	157

Table 2. Classification accuracies (%) of the proposed method on FGNET and PAL databases (M: male test samples, F: female test samples, M+F: all test samples).

Age Range	FG-NET			PAL		
	M	F	M+F	M	F	M+F
0-12	92.69	89.24	91.54	-	-	-
13-19	85.00	84.16	84.61	100.0	93.75	95.83
20-39	68.18	60.00	64.19	75.63	70.09	73.01
40-64	96.97	80.77	90.27	96.65	71.18	78.32
≥ 65	-	-	-	98.38	78.95	84.58
Average	86.67	79.55	83.81	85.48	75.45	79.34

Table 3. Comparison between the previous works and proposed method for age-group classification (AG: Age Group, N/A: Not Available).

Method	AG #	Age Groups	Database	Image # Method	Image % (Method)	Accuracy (%)
Kwon and Lobo [12]	3	Baby, adult, elder	N/A	47	N/A	68
Horng <i>et al.</i> [8]	4	0-2, 3-39, 40-59, ≥ 60	N/A	230	N/A	81.58
Hayashi <i>et al.</i> [6]	6	10, 20, 30, 40, 50, 60 years	HOIP	300	N/A	27
Txia and Huang [24]	4	20-30, 31-40, 41-50, 51-60	MORPH	251	14.85	72.52
Dehshibi and Bastanfard [4]	4	≤15, 16-30, 31-50, ≥51	Iranian Face DB	498	13.83	86.64
Chen <i>et al.</i> [3]	3	Baby, adult, elder	FGNET +FERET	180	8.17	87.8
Higashi <i>et al.</i> [7]	4	0-9, 10-19, 20-34, 35-64	FGNET	261	26.04	60.69
Tonchev <i>et al.</i> [23]	2	0-16, 17-69	FGNET	1002	100	77.73
Liu <i>et al.</i> [15]	5	0-2, 3-10, 11-18, 19-39, 40-69	FGNET	125	12.47	79.2
Kalansuriya and Dharmaratne [9]	4	8-13, 14-25, 26-45, 46-63	FGNET +FERET	750	34.07	74.8
Kohli <i>et al.</i> [11]	2	0-21, 22-69	FGNET	1002	100	84
Li <i>et al.</i> [13]	4	0-17, 18-45, 46-60, 60-69	FGNET +self-build	1600	-	85.75
Sai <i>et al.</i> [22]	4	0-10, 11-19, 20-60, >60	FGNET +PAL	383	24	82
Proposed Method	5	0-12, 13-19, 20-39, 40-64, ≥ 65	FGNET PAL	1002 578	100 100	83.81 79.34

5. Conclusions

In this paper, we proposed a novel age classification system using morph-based age models. The common characteristics of age groups are represented by age models produced with image morphing method. The global and local LBP histograms of these age models are used as age related features in the classification module. The classification is performed using ensemble of distance metrics. As the performances of distance metrics are different from each other, their results are combined using BC voting method to improve the classification performance.

The boundaries used for age groups in previous works are considerably different from each other. Furthermore, the different methods tested on the same database have used different boundaries for age groups. Standardizing the age groups will facilitate the performance evaluation of the classification methods with each other. In this respect, the boundaries of age groups in this study are designated according to developmental psychologist and psychoanalyst Erik Erikson's stages of psychosocial development theory.

In the age model generation phase, the training samples are selected properly from age groups and this provides the determination of common properties of

that age group. Thus, our study shows us that, we can achieve classification performances similar to or better than previous works with a simpler system rather than a complex system trained with large datasets. The idea of recognition using face models can be suggested in other facial image processing research areas.

Acknowledgments

The authors would like to thank the Editor and anonymous reviewers for their valuable comments and suggestions to improve the quality of the manuscript.

References

- [1] Aeminov V., Panteleev I., and Selivanov N., *Criminalistics*, Progress Publishers, 1989.
- [2] Albert M., Ricanek K., and Patterson E., "A Review of the Literature on the Aging Adult Skull and Face: Implications for Forensic Science Research and Applications," *Forensic Science International*, vol. 172, no. 1, pp. 1-9, 2007.
- [3] Chen Y., Han M., Song K., and Ho Y., "Image-Based Age-Group Classification Design Using Facial Features," in *Proceedings of International Conference on System Science and Engineering*, Taipei, pp. 548-552, 2010.
- [4] Dehshibi M. and Bastanfard A., "A New Algorithm for Age Recognition from Facial Images," *Signal Processing*, vol. 90, no. 8, pp. 2431-2444, 2010.
- [5] Erikson E., *Identity, Youth and Crisis*, Norton Company, 1968.
- [6] Hayashi J., Yasumoto M., Ito H., Niwa Y., and Koshimizu H., "Age and Gender Estimation from Facial Image Processing," in *Proceedings of the 41st SICE Annual Conference*, Osaka, pp. 13-18, 2002.
- [7] Higashi A., Yasui T., Fukumizu Y., and Yamauchi H., "Local Gabor Directional Pattern Histogram Sequence for Age and Gender Classification," in *Proceedings of IEEE Statistical Signal Processing Workshop*, Nice, pp. 505-508, 2011.
- [8] Horng W., Lee C., and Chen C., "Classification of Age Groups Based on Facial Features," *Tamkang Journal of Science and Engineering*, vol. 4, no. 3, pp. 183-192, 2001.
- [9] Kalansuriya T. and Dharmaratne A., "Facial Image Classification Based on Age and Gender," in *Proceedings of International Conference on Advances in ICT for Emerging Regions*, Colombo, pp. 44-50, 2013.
- [10] Karungaru S., Fukumi M., and Akamatsu N., "Morphing Human Faces: Automatic Control Points Selection and Color Transition," in *Proceedings of World Academy of Science*,

- Engineering and Technology*, vol. 1, pp. 224-227, 2004.
- [11] Kohli S., Prakash S., and Gupta P., "Hierarchical Age Estimation with Dissimilarity-Based Classification," *Neurocomputing*, vol. 120, pp. 164-176, 2013.
- [12] Kwon Y. and Lobo N., "Age Classification from Facial Images," *Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 1-21, 1999.
- [13] Li W., Su H., Pan F., Gao Q. and Guo S., "Facial Age Classification Based on Weighted Decision Fusion," in *Proceedings of the 33rd Chinese Control Conference*, pp. 4870-4874, 2014.
- [14] Liu K., Yan S., and Kuo C., "Age Estimation via Grouping and Decision Fusion," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 11, pp. 2408-2423, 2015.
- [15] Liu L., Liu J., and Cheng J., "Age-Group Classification of Facial Images," in *Proceedings of 11th International Conference on Machine Learning and Applications*, Boca Raton, pp. 693-696, 2012.
- [16] Minear M. and Park D., "A Lifespan Database of Adult Stimuli," *Behavior Research Methods, Instruments and Computers*, vol. 36, no. 4, pp. 630-633, 2004.
- [17] Murtaza M., Sharif M., Raza M., and Shah J., "Analysis of Face Recognition Under Varying Facial Expressions: A Survey," *The International Arab Journal of Information Technology*, vol. 10, no. 4, pp. 378-388, 2013.
- [18] Nguyen D., Cho S., Shin K., Bang J., and Park K., "Comparative Study of Human Age Estimation with or without Preclassification of Gender and Facial Expression," *The Scientific World Journal*, 2014.
- [19] Ojala T., Pietikainen M., and Maenpaa T., "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971-987, 2002.
- [20] Parramon J., *How to Draw Heads and Portraits*, Watson-Guption Publications, 1989.
- [21] Ren H. and Li Z., "Age Estimation Based on Complexity-Aware Features," in *Proceedings of 12th Asian Conference on Computer Vision*, Singapore, pp. 115-128, 2015.
- [22] Sai P., Wang J., and Teoh E., "Facial Age Range Estimation with Extreme Learning Machines," *Neurocomputing*, vol. 149, pp. 364-372, 2015.
- [23] Tonchev K., Paliy I., Boumbarov O., and Sokolov S., "Human Age-Group Classification of Facial Images with Subspace Projection and Support Vector Machines," in *Proceedings IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Prague, pp. 439-443, 2011.
- [24] Txia J. and Huang C., "Age Estimation using AAM and Local Facial Features," in *Proceedings of IEEE 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Kyoto, pp. 885-888, 2009.



Asuman Günay Yılmaz received her BSc., MSc. and Ph.D. Degrees from Department of Computer Engineering, Karadeniz Technical University, Trabzon, Turkey in 2000, 2003 and 2015 respectively. Currently, she is a Lecturer in Trabzon Vocational School, Karadeniz Technical University, Trabzon, Turkey. Her research interests include biometric security and machine learning.



Vasif Nabiyev received BSc. and MSc. degrees in the Faculty of Computer Engineering and Automation from St. Petersburg Electro Technical University in 1985, and a Ph.D. degree in the Department of Computer Science from Moscow Technical University in 1990. Currently he is a Professor in Department of Computer Engineering, Karadeniz Technical University, Trabzon, Turkey. His research interests are artificial intelligence, biometry, security, image processing, operational research, combinatorial algorithms. He is author of textbooks titled "Artificial Intelligence: Problems, Methods and Algorithms", "Algorithms: From Theory to Applications".