

Facial Age Estimation Using Spatial Weber Local Descriptor

Asuman Günay and Vasif V. Nابیev

Abstract—This paper introduces a novel age estimation method using a new texture descriptor Weber Local Descriptor (WLD). This texture descriptor is analyzed in depth for age estimation problem. In the study, the multi-scale versions of holistic and spatial WLD (SWLD) descriptors are used to extract the age related features from normalized facial images. After finding a lower dimensional feature subspace, age estimation is performed using multiple linear regression. In addition the age estimation accuracy of each of the distinct and intersection block used in spatial texture extraction are investigated. Experiments on FG-NET, MORPH and PAL databases have shown that similar age estimation performances can be obtained by using more effective blocks in spatial histogram generation. This also provides us to reduce the number of features and computational cost.

Keywords—age estimation, regression, Weber local descriptor.

I. INTRODUCTION

Facial age estimation has been demanding growing attention recently because of its real world applications such as, internet access control, age-specific target advertising, face recognition systems robust to age progression, and age prediction systems for finding the lost children and criminals.

Age estimation is the process of associating a facial image automatically with exact age or age group. The suitable facial image representation is crucial in age estimation, as the success of the classifiers depends on the adequacy of the feature domain [1]. For this reason the selection of facial features related with aging is very important for age estimation systems design. But the variations on different individuals faces are not the same during their life span. Real world age progression is personalized as shown in Fig. 1 such that it is related with living styles, eating habits, climate, etc. Also factors such as genetics, race, frequency of facial expressions, anti-aging product usage affects the facial appearance [2]. Another challenge in age estimation systems is the availability of a good aging database. As aging is a slow process, a large aging database, especially containing the chronometrical image series of individuals is hard to collect.

Age estimation approaches generally comprises feature extraction and estimation phases. In feature extraction phase, shape-based and texture-based methods are used for efficient facial image representation. In the second phase, classification and regression methods or hierarchical classifiers are utilized to perform age estimation. Kwon and Lobo published the first

Manuscript received November 1, 2016; revised April 25, September 21 2017.

A. Günay and V. V. Nابیev are with the Department of Computer Technologies, Karadeniz Technical University, Trabzon, Turkey, e-mail: {gunaya, vasif}@ktu.edu.tr.



Fig. 1. Face aging of different individuals (Marie Curie, Charlie Chaplin www.gettyimages.com)

work for age classification [3]. They have calculated distance ratios between facial features [4] based on face anthropometry and extracted wrinkle information using snakelets, to classify facial images into three groups: babies, young and senior adults. Later on age classification methods using the geometrical features and texture features extracted with different image filters are also proposed [5], [6].

Lanitis *et al.* [7] initially proposed using Active Appearance Models (AAM) for age estimation. With the usage of AAMs the shape and appearance information of facial images are joined together and a parametric description is obtained. Then an aging function defining the relationship between the parametric description and the age of the individual is built. After their method, AAMs are frequently used in age estimation frameworks [8], [9]. In some studies AAM features are extracted as global facial features and fused with local facial features for efficient age estimation [10]. In Geng *et al.*'s approach [11] a unique aging pattern is learned from the sequence of facial images sorted in temporal order for each individual. On the contrary, a common aging pattern is learned from the images of different individuals at different ages by using manifold embedding techniques in some age estimation approaches. The mapping from the image space to the manifold space can be done either by linear or by nonlinear functions [12], [13]. Also extraction of aging related facial features using various methods such as edge detection methods, Gabor filters, Local Binary Patterns (LBP) is explored in many studies [10], [14]–[17]. Both the global and features extracted from the whole image and local features extracted from image blocks or facial regions (forehead, eyes, cheeks, etc.) are used for age estimation. Recently, deep learning methods and dynamic age groups are used in age estimation studies [18]–[21].

In this paper a novel age estimation method using textural properties of facial images is introduced. The idea of using

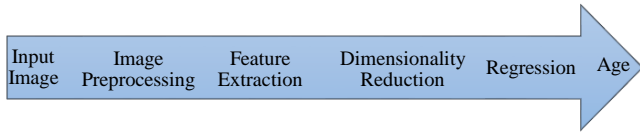


Fig. 2. Age estimation system

texture information of facial images is not new, however we use a new texture descriptor Weber Local Descriptor (WLD) [22] which has not been analyzed in depth for age estimation. The block diagram of age estimation system is shown in Fig. 2. In our method, the original facial images are normalized so that the orientation and the size of faces are adjusted and only the facial region is extracted. Then multi-scale and spatial WLD (SWLD) descriptors are used to extract the age related features from normalized facial images. After dimensionality reduction with Principal Component Analysis (PCA), multiple linear regression is used to estimate the specific age.

The main goal of the study is to analyze deeply the age estimation performance of a new texture descriptor (Weber Local Descriptor-WLD) which extracts both the edges and the directions of edges. This analysis is performed on 3 different databases with different characteristics for different scenarios. Thus, multi-scale and spatial features for different parameter values are extracted to find the optimum parameters, block size and scale for a general age estimation system. Also the age estimation accuracies of face blocks used in spatial histogram generation are analyzed separately for finding the blocks which contribute mostly to age estimation process. The aim is to obtain more accurate systems by using these blocks, also reduce the number of features and computational cost.

The rest of the paper is organized as follows. The proposed age estimation approach is explained in Section II. In Section III the experimental results on various databases are reported and analyzed. Finally, the conclusions are outlined in Section IV.

II. PROPOSED METHOD

The proposed method consist of image normalization, spatial and multi-scale WLD histogram extraction, dimensionality reduction, age estimation modules. These modules are explained in detail in the following subsections.

A. Image Processing

The orientation and the size of original images are different from each other as shown in Fig. 3(a). Besides, the original images have unnecessary features such as background, cloth and hair which are not related to the face and can affect the performance of the algorithm. Therefore, the images are rotated, scaled and cropped to extract only the facial regions and to adjust the orientation and size of the faces. (Fig. 3(b)). To adjust the orientation, the images are rotated using the angle of the line between the eye centers. Then the sizes of facial images are adjusted by scaling (up or down) and the distance between the eye centers is 50 pixels. Finally the scaled images are cropped to the size of 88×88 pixels using the eye center positions. For this reason the sizes and the eye center positions

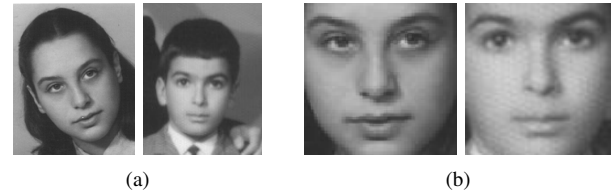


Fig. 3. Image preprocessing: (a) original images, (b) normalized images

of the normalized images are the same and only the facial region is obtained.

B. Feature Extraction with WLD

WLD [22] is based on the fact that human perception of a pattern depends not only on the amount of change in intensity of a stimulus (such as sound, lighting) but also on the original intensity of the stimulus. According to Webers law, the change of a just noticeable stimulus is a constant proportion of the original stimulus. Basically, WLD consist of two components: differential excitation (ξ) and orientation (θ). With the differential excitation component local salient micropatterns are extracted. Then statistics are built on these patterns along with the gradient orientation of the current point. Consequently WLD has the ability for detecting edges perfectly, which are very important for age estimation. Additionally, when compared to LBP [23], a powerful texture (bright/dark spots, edges, flat areas) descriptor used frequently in age estimation systems, WLD not only uses the local intensity variations to extract the edges, also uses the gradient orientations to describe the direction of edges.

1) *Differential Excitation*: The differential excitation $\xi(x_c)$ of a current pixel x_c is computed as:

$$\xi(x_c) = \arctan \left[\frac{v_s^{00}}{v_s^{01}} \right], \quad (1)$$

where $v_s^{00} = \sum_{i=0}^{p-1} x_i - x_c$, $x_i (i = 0, 1, \dots, p-1)$ denotes the i -th neighbors of x_c , p is the number of neighbors, and $v_s^{01} = x_c$.

2) *Orientation*: The orientation component of WLD is the gradient orientation, which is computed as:

$$\theta(x_c) = \arctan \left[\frac{v_s^{11}}{v_s^{10}} \right], \quad (2)$$

where v_s^{11} is the intensity difference of the two pixels on the left and right of the current pixel x_c and v_s^{10} is the intensity difference of the two pixels directly below and above of the current pixel x_c . After mapping $\theta \in [-\pi/2, \pi/2]$ to $\theta' \in [0, 2\pi]$, it is quantized into T dominant orientations as follows:

$$\Phi = \frac{2t}{T} \pi, \quad t = \text{mod} \left(\left\lfloor \frac{\theta'}{2\pi/T} + \frac{1}{2} \right\rfloor, T \right). \quad (3)$$

3) *WLD Histogram*: In this step 2D histogram $WLD(\xi_j, \Phi_t)$, ($j = 0, 1, \dots, N-1, t = 0, 1, \dots, T-1$) is computed where N is dimensionality of an image and T is the number of dominant orientations.

To obtain a more discriminative descriptor the columns of $WLD(\xi_j, \Phi_t)$ are projected into 1D histograms H_t , $t = 0, 1, \dots, T-1$. Subsequently each H_t is evenly divided into M

segments, $H_{(m,t)}$, ($m = 0, 1, \dots, M - 1$). Furthermore each sub-histogram segment $H_{(m,t)}$ is composed of S bins, i.e. $H_{(m,t)} = \{h_{m,t,s}\}$, $s = 0, 1, \dots, S - 1$. Then these sub-histogram segments ($H_{(m,t)}$) form a histogram matrix where each column corresponds to a dominant direction Φ_t . Each row of this matrix is concatenated as a histogram which is referred as WLD descriptor. The size of WLD descriptor is $T \times M \times S$. In the study, age estimation accuracy of the WLD descriptor is analyzed using different values for T , M and S parameters ($T = 4, 8$; $M = 4, 6$; $S = 3, 5, 7, 9$). In the experiments $2 \times 2 \times 4 = 16$ different parameter combinations are used for finding the optimum values.

4) *Multi-scale WLD Histogram*: Original WLD histograms are extracted from the 3×3 neighborhood [22]. Multi-scale WLD histograms are computed using a square symmetric neighbor sets of P pixels placed on a square whose sides have the length $(2R + 1)$ as shown in Fig. 4. The WLD operator is denoted as $WLD_{(P,R)}$ where P is the number of neighbors and R determines the spatial resolution of the operator. Multi-scale analysis is performed using 3 different scales ($P = 8$ $R = 1$, $P = 16$ $R = 2$, $P = 24$ $R = 3$) in the study. This means that 48 (16×3) different scenarios are used in the experiments for finding the optimum parameters and scale.

5) *Spatial WLD Histogram*:

The basic WLD descriptor represents an image as a histogram of differential excitations organized according to dominant gradient orientations. In this histogram differential excitations are collected according to their values and gradient orientations regardless of their spatial location. Consequently, two different regions in an image with similar differential excitations and gradient orientations will contribute to the same bins in the histogram, and will not be discriminated by WLD descriptor. To enhance the discriminatory power of WLD descriptor, each image is divided into number of blocks, WLD histogram for each block is computed and then these regional histograms concatenated to form a SWLD descriptor. Unfortunately, as a result of discrete evaluation of the blocks, some related information between the blocks may not be completely captured. To obtain these features, the intersection blocks obtained by combining the centers of discrete blocks are also used in spatial histogram generation. The features extracted from discrete and intersection blocks are concatenated to form a more discriminative SWLD histogram.

In the study age estimation performance of the texture descriptor is deeply analyzed by using 5 different block configurations (1×1 , 2×2 , 4×4 , 8×8 , $8 \times 8+$) for each

parameter and scale scenario. For this reason the results are obtained for 240 (48×5) different states. Also 3 databases which have different characteristics are used to determine the parameter values of the texture descriptor (T , M , S , scale and block size) to conduct a general age estimation system. In addition the age estimation accuracies of the blocks used in spatial histogram generation are analyzed separately for finding the blocks that contributes mostly to the age estimation process.

C. Dimensionality Reduction

After SWLD histogram generation, PCA is performed on these feature vectors to find a lower dimensional subspace which carries significant information for age estimation. The PCA method finds the embedding that maximizes the projected variance given below:

$$W_{opt} = \arg \max_{\|W\|=1} W^T S W. \quad (4)$$

in (4) $S = \sum_{i=1}^n (f_i - \bar{f})(f_i - \bar{f})^T$ is the scatter matrix, f_i is i -th feature vector and \bar{f} is the mean vector. Solving this problem, a set of d eigenvectors associated to the d largest eigenvalues of S is obtained. Then dimensionality reduction is performed by calculating $y_i = W^T f_i$ which gives the low dimensional representation of feature vectors.

D. Regression

In our study, we define the age estimation problem as a multiple linear regression problem as $age = f(M) \Leftrightarrow \hat{L} = \hat{f}(Y)$, where \hat{L} denotes the estimated age label, $f(\cdot)$ the unknown regression function, and $\hat{f}(\cdot)$ is the estimated regression function. The matrix formulation is given by $L = \tilde{Y}B$ where L is the age label vector and \tilde{Y} is the matrix of observed values. B is the unknown parameter vector which we need to estimate during the learning stage by ordinary least squares or robust regression. The aging function used in this study is a linear function given in $\hat{l} = \hat{\beta}_0 + \hat{\beta}_1^T y$, where \hat{l} is the estimate of age, $\hat{\beta}_0$ is the offset, $\hat{\beta}_1$ is the weight vector and y is low dimensional representation of the extracted feature vector.

III. EXPERIMENTS AND RESULTS

In this paper the FG-NET, MORPH and PAL databases are used to evaluate the performance of the proposed method. The FG-NET aging database [24] includes 1002 images of 82 individuals with uncontrolled head pose, facial expression and illumination. The MORPH database [25] consists of 1690 images taken from 515 subjects. The PAL database [26] contains 580 images with different facial expressions, captured under natural lighting conditions. The data distributions of FG-NET, MORPH and PAL databases according to age are given in Fig. 5. Also the age ranges of the databases can be seen from the graphics.

The performance of the proposed method is evaluated using 3-fold cross validation. First the samples are randomly partitioned into 3 equal sized subsamples. Then one subsample is retained as test set and two subsamples are used as training set. This procedure is repeated three times with each of the

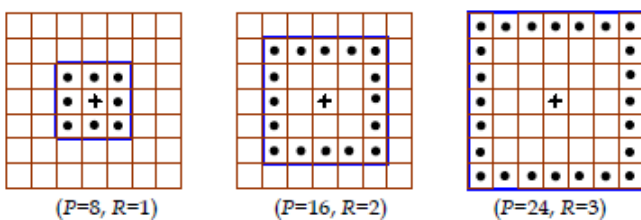


Fig. 4. Square symmetric neighborhood for different (P, R)

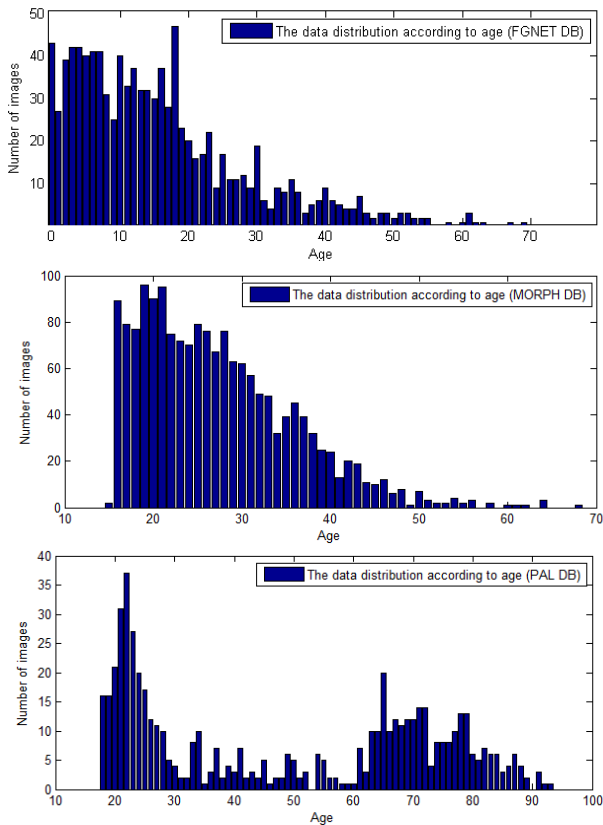


Fig. 5. The data distributions of FG-NET, MORPH and PAL databases

TABLE I
THE NUMBER OF TRAINING/TESTING SAMPLES USED IN EACH FOLD OF EVALUATION METHOD

Database	# of training images	# of testing images
FGNET	668	334
MORPH	1127	563
PAL	387	193

three subsamples used exactly once as test set. The results are then averaged to produce a single estimation. The number of testing/training images used is each fold of cross validation scheme is given in Table I.

The performance is measured using the Mean Absolute Error (MAE) metric given as:

$$MAE = \sum_{i=1}^{N_t} |\hat{l}_i - l_i| / N_t, \tag{5}$$

where \hat{l}_i is the estimated age value of i -th test sample, l_i is the real age value of i -th test sample, and N_t is the total number of test samples.

In the experiments the age estimation performances of multi-scale WLD descriptors are calculated for various combinations of T , M and S parameters. The multi-scale WLD operator, $WLD_{P,R}$ (P is the number of neighbors, R is the distance between the neighbors and the center pixel) is applied in 3 different scales: $WLD_{8,1}$, $WLD_{16,2}$ and $WLD_{24,3}$. Experiments are performed for $T = 4, 8$, $M = 4, 6$ and $S = 3, 5, 7, 9$ for finding optimal age estimation accuracies. Firstly the WLD histograms are extracted from the whole

images (1×1) and age estimation is performed using these holistic texture descriptors. Then the images are divided into various number of blocks and SWLD histograms are extracted from these blocks. As the image size is 88×88 after image normalization step, we divided them into 2×2 blocks (block size is 44×44 pixels), 4×4 blocks (block size is 22×22 pixels) and 8×8 blocks (block size is 11×11 pixels), extract WLD histograms from each block and concatenated them to form spatial texture descriptor. Besides, according to the proposed spatial histogram generation approach, the WLD histograms extracted from 8×8 blocks and 7×7 blocks obtained by combining the centers of these 8×8 blocks are concatenated to enhance the discriminative power of descriptor. The results of all experiments are tabulated in Table II. In the table $WLD_{P,R}$ is the multi-scale WLD operator used in the experiments. The result are obtained using $P = 8 R = 1$, $P = 16 R = 2$ and $P = 24 R = 3$ ($WLD_{8,1}$, $WLD_{16,2}$ and $WLD_{24,3}$). According to the results the best age estimation accuracies for FGNET, MORPH and PAL databases are obtained with $WLD_{8,1}$ operator. Also the accuracies obtained with the parameters $T = 8$ and $M = 6$ are better than the other combinations for all databases. It can be seen from the table that age estimation accuracies of $WLD_{8,1}$, histograms for MORPH database are the same for parameters $T = 8 M = 6 S = 7$ and $T = 8 M = 4 S = 9$. In order to design a general age estimation system independent from the databases and image types, image qualities, etc. common T , M and S parameter values should be selected for all databases used in the experiments. We can also see from the table that the best MAEs are obtained with $T = 8$ and $M = 6$ parameter values for FGNET and PAL databases. For this reason the same values ($T = 8$ and $M = 6$) are selected for these parameters for MORPH database. Only the value of the S parameter differs according to the data set ($S = 9$ for FGNET, $S = 7$ for MORPH and $S = 3$ for PAL database). When the results are deeply analyzed, it can be seen that the value of the S parameter can be taken $S = 5$ to make a common age estimation system for all databases with different characteristics by ignoring the small decreases in the accuracies.

The MAEs of holistic and spatial $WLD_{8,1}$, $WLD_{16,2}$,

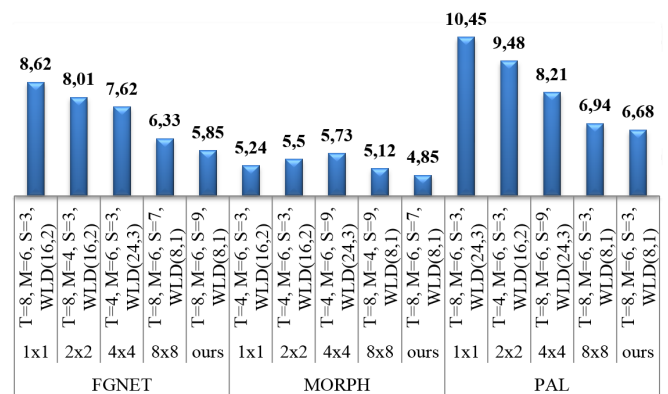


Fig. 6. The best MAE's of holistic and spatial $WLD_{8,1}$, $WLD_{16,2}$, $WLD_{24,3}$ descriptors with T , M , S parameters.

TABLE II
THE MAES OF HOLISTIC AND SPATIAL FEATURES EXTRACTED USING MULTI-SCALE WLD OPERATORS WITH DIFFERENT COMBINATIONS OF T , M AND S PARAMETERS

#of blocks	WLD $_{P,R}$	S	$T = 8, M = 6$			$T = 8, M = 4$			$T = 4, M = 6$			$T = 4, M = 4$		
			FGNET	PAL	MORPH	FGNET	PAL	MORPH	FGNET	PAL	MORPH	FGNET	PAL	MORPH
1×1	WLD $_{8,1}$	3	9.23	14.53	6.16	9.20	14.22	6.08	9.47	16.51	6.38	9.46	16.78	6.36
		5	9.16	11.42	5.87	9.11	12.34	5.99	9.35	12.22	6.00	9.40	14.41	6.24
		7	9.11	10.89	5.78	9.11	11.16	5.83	9.21	11.48	5.89	9.28	12.29	6.03
		9	9.18	10.85	5.82	9.14	10.90	5.89	9.23	11.44	5.84	9.23	11.48	5.98
	WLD $_{16,2}$	3	8.62	10.62	5.52	8.76	10.86	5.56	8.84	11.10	5.58	8.79	11.28	5.66
		5	8.73	10.91	5.51	8.88	10.63	5.51	8.73	11.37	5.54	8.91	11.07	5.54
		7	9.04	11.39	5.65	9.02	10.81	5.52	8.83	11.53	5.58	8.86	11.31	5.49
		9	9.14	11.85	5.71	9.23	11.24	5.55	8.98	11.90	5.58	8.90	11.49	5.52
	WLD $_{24,3}$	3	8.95	10.45	5.74	8.68	10.41	5.63	8.83	11.11	5.67	8.73	11.24	5.64
		5	9.13	11.09	5.88	8.80	11.05	5.78	8.90	11.31	5.24	8.79	11.44	5.67
		7	9.58	11.80	6.05	9.01	11.16	5.84	9.05	12.00	5.82	8.92	11.69	5.72
		9	10.12	12.37	6.11	9.40	11.45	5.95	9.26	12.00	5.92	8.86	11.83	5.75
2×2	WLD $_{8,1}$	3	8.34	11.92	5.69	8.27	11.96	5.67	8.62	12.50	6.00	8.71	12.91	6.00
		5	8.52	10.78	5.72	8.54	10.93	5.70	8.51	11.08	5.83	8.69	11.43	5.87
		7	8.58	10.37	5.90	8.61	10.32	5.80	8.68	10.91	5.77	8.76	10.67	5.84
		9	9.01	10.90	6.01	9.13	10.70	5.95	8.70	10.89	5.82	8.93	10.80	5.88
	WLD $_{16,2}$	3	8.29	9.96	5.73	8.01	9.50	5.60	8.22	9.48	5.50	8.30	9.56	5.64
		5	8.62	9.74	5.88	8.25	10.14	5.77	8.68	10.07	5.62	8.49	9.88	5.66
		7	9.29	10.41	6.33	8.88	10.06	6.01	9.06	10.25	5.97	8.71	10.07	5.76
		9	9.37	11.22	6.50	9.35	10.97	6.15	9.53	11.04	5.92	8.83	10.30	5.92
	WLD $_{24,3}$	3	8.74	10.10	6.13	8.62	9.64	5.85	8.39	10.07	5.84	8.13	9.67	5.75
		5	9.18	10.06	6.58	9.05	10.42	6.18	8.99	10.98	6.15	8.54	10.01	5.94
		7	9.29	10.06	7.21	9.75	9.99	6.69	9.61	10.77	6.37	8.77	10.90	6.08
		9	9.41	10.57	7.26	9.89	10.09	6.81	9.90	11.38	6.70	9.16	10.99	6.47
4×4	WLD $_{8,1}$	3	8.63	10.35	5.97	8.60	9.95	5.94	8.03	9.79	5.73	8.03	9.84	5.75
		5	8.47	9.97	6.40	8.44	9.96	6.30	8.53	10.45	6.03	8.59	10.11	6.08
		7	8.63	9.42	6.44	8.60	9.27	6.52	9.21	10.77	6.27	9.14	10.73	6.33
		9	8.65	9.69	6.43	8.69	9.51	6.40	9.24	10.49	6.44	9.15	10.20	6.47
	WLD $_{16,2}$	3	8.17	9.47	6.43	8.36	9.91	6.43	7.62	9.69	5.97	7.88	9.84	5.93
		5	7.85	9.41	6.53	8.15	9.59	6.59	8.25	9.85	6.35	8.19	9.96	6.34
		7	7.68	9.21	6.51	8.07	9.26	6.66	8.49	9.52	6.57	8.45	10.03	6.76
		9	7.75	8.86	6.34	8.03	9.38	6.53	8.58	9.34	7.03	8.59	9.66	6.89
	WLD $_{24,3}$	3	8.04	8.80	6.32	8.54	9.48	6.60	8.31	9.31	6.07	8.17	9.75	6.12
		5	7.75	8.56	6.37	8.17	8.97	6.52	8.64	8.97	6.47	8.63	9.88	5.56
		7	7.54	8.37	6.23	7.89	8.81	6.51	8.54	9.05	6.64	8.78	9.84	6.70
		9	7.47	8.21	6.09	7.69	8.55	6.30	7.97	8.85	6.35	8.36	9.59	6.99
8×8	WLD $_{8,1}$	3	6.99	6.94	5.82	7.00	7.01	5.78	7.81	8.42	6.38	7.89	8.58	6.46
		5	6.60	7.15	5.40	6.58	7.17	5.38	7.51	8.24	6.28	7.57	8.38	6.31
		7	6.33	7.25	5.24	6.34	7.27	5.21	7.16	8.37	6.25	7.23	8.43	6.22
		9	6.34	7.54	5.14	6.34	7.57	5.12	7.17	8.21	6.00	7.18	8.23	5.99
	WLD $_{16,2}$	3	6.65	8.36	5.68	6.81	8.32	5.84	7.67	8.66	6.36	8.10	8.47	6.41
		5	6.57	8.52	5.53	6.66	8.44	5.59	7.37	8.61	6.14	7.75	8.25	6.29
		7	6.57	8.58	5.41	6.87	8.66	5.42	7.34	8.74	5.99	7.60	8.58	6.02
		9	6.40	8.65	5.42	6.57	8.83	5.45	7.18	8.61	5.87	7.46	8.69	5.92
	WLD $_{24,3}$	3	6.68	8.72	5.60	7.02	9.12	5.69	7.35	9.31	5.53	7.51	9.81	5.84
		5	6.64	8.69	5.55	6.92	9.16	5.59	7.12	9.46	5.49	7.41	9.78	5.80
		7	6.71	8.85	5.43	6.92	9.19	5.48	6.91	9.51	5.50	7.26	9.86	5.74
		9	6.61	8.95	5.43	6.89	9.34	5.40	6.87	9.55	5.49	7.17	9.88	5.68
ours	WLD $_{8,1}$	3	6.22	6.68	5.18	6.21	6.72	5.19	7.03	7.61	5.91	7.08	7.65	5.98
		5	6.13	6.97	4.94	6.12	7.02	4.94	6.65	7.79	5.60	6.68	7.80	5.64
		7	6.03	7.02	4.85	6.02	7.06	4.86	6.44	7.78	5.52	6.48	7.85	5.54
		9	5.85	7.26	4.85	5.99	7.29	4.85	6.47	7.56	5.40	6.48	7.61	5.42
	WLD $_{16,2}$	3	6.75	8.69	5.33	6.73	9.12	5.41	7.37	8.12	5.87	7.42	8.16	5.85
		5	6.70	8.85	5.11	6.53	9.25	5.21	7.03	8.24	5.57	7.10	8.20	5.62
		7	6.68	9.03	5.10	6.58	9.43	5.19	6.96	8.50	5.51	6.92	8.59	5.44
		9	6.78	9.13	5.17	6.71	9.60	5.17	6.91	8.57	5.47	6.83	8.71	5.39
	WLD $_{24,3}$	3	6.83	8.56	5.30	7.02	9.05	5.38	7.33	8.52	5.32	7.51	9.29	5.67
		5	6.73	8.77	5.22	6.97	9.18	5.36	7.16	8.85	5.21	7.36	9.49	5.56
		7	6.76	8.93	5.18	7.05	9.31	5.32	7.01	9.07	5.20	7.37	9.55	5.51
		9	6.69	9.11	5.20	7.02	9.51	5.32	6.95	9.15	5.21	7.34	9.69	5.48

TABLE III
EXECUTION TIMES (IN SECONDS) FOR TRAINING AND TEXTING SAMPLES

Database	Training time	Testing Time per 1 image
FG-NET	196.4033	0.2818
MORPH	338.8221	0.2848
PAL	108.5187	0.2781

MORPH and PAL databases are given in Fig. 6 for better understanding the effects of WLD parameters and the spatial features for age estimation. These results are the best ones for multi-scale WLD descriptors using different combinations of T , M and S parameters. We can see from the figure that the results obtained using holistic WLD is not promising and increasing the number of blocks improves the estimation performance for all databases used in the experiments. Also it

WLD $_{24,3}$ descriptors with T , M , S parameters on FG-NET,

can be seen from Table II that the age estimation accuracy is increased according the number of blocks regardless of the scale and parameters of the WLD descriptor. Because increasing the number of blocks increases the discriminative power of the descriptor by ensuring that the different regions in an image with similar properties will not contribute to the same bin in the WLD histogram. In addition better estimation accuracies are obtained by including the features extracted from the intersection blocks for all databases. The minimum MAEs of 5.85 years for FG-NET database, 4.85 years for MORPH database and 6.68 years for PAL database are obtained using proposed approach with $WLD_{8,1}$ descriptor with parameters of $T = 8$ and $M = 6$. The values of S parameter for FGNET, MORPH and PAL databases are 9, 7 and 3 respectively.

In the experiments a PC with an Inter core i5 2.40 GHz CPU and 3GB memory is used. The training and testing times of the age estimation system are tabulated in Table III. In the table training time includes the time needed for spatial feature extraction, dimensionality reduction and age estimation phases. As 3-fold cross validation evaluation scheme is used, the number of training images for FG-NET, MORPH and PAL databases are approximately 668, 1127 and 387 respectively. For this reason differences between the training times of databases are observed.

When the testing time per a facial image is analyzed, times needed to estimate the age of a test sample are similar regardless of the databases.

Chen *et al.* [22] demonstrated that WLD outperforms LBP [23] which is a powerful texture descriptor in texture recognition and is frequently used in state-of-art age estimation methods. For this reason we compare the age estimation performance of WLD with LBP in our study. In our experiments $WLD_{8,1}$ descriptor with $T = 8$, $M = 6$ and $S \in \{3, 5, 7, 9\}$ gives better results than $WLD_{16,2}$ and $WLD_{24,3}$ descriptors and other parameter values. For this reason we compare the estimation accuracies of $WLD_{8,1}$ descriptor with $T = 8$, $M = 6$, $S \in \{3, 5, 7, 9\}$ and $LBP_{8,1}$ descriptor, and the results are shown in Fig. 7. It can be seen from the figure that WLD descriptor generally outperforms LBP descriptor for FG-NET, MORPH and PAL databases. Besides, using the proposed approach improves the estimation accuracy regardless of the texture descriptors and databases. Furthermore, the results of texture analysis in [23] show that much of the discriminative texture information is contained in high spatial frequencies such as edges. In this context our results also indicate that the WLD descriptor extracts more powerful discriminating features than the LBP descriptor for age estimation.

In the study the age estimation accuracies of the blocks used in spatial feature extraction are calculated separately. In this process the scales and the parameters of the WLD operator for the databases are selected according to the best results obtained and mentioned previously. Then age estimation is performed by using only the WLD features extracted from one block of the image. The age estimation performances of 64 blocks for FGNET, PAL and MORPH databases are shown in Fig. 8, Fig. 9 and Fig. 10 respectively. It can be seen from the Fig. 8 that the estimation performances of the blocks around the eyes

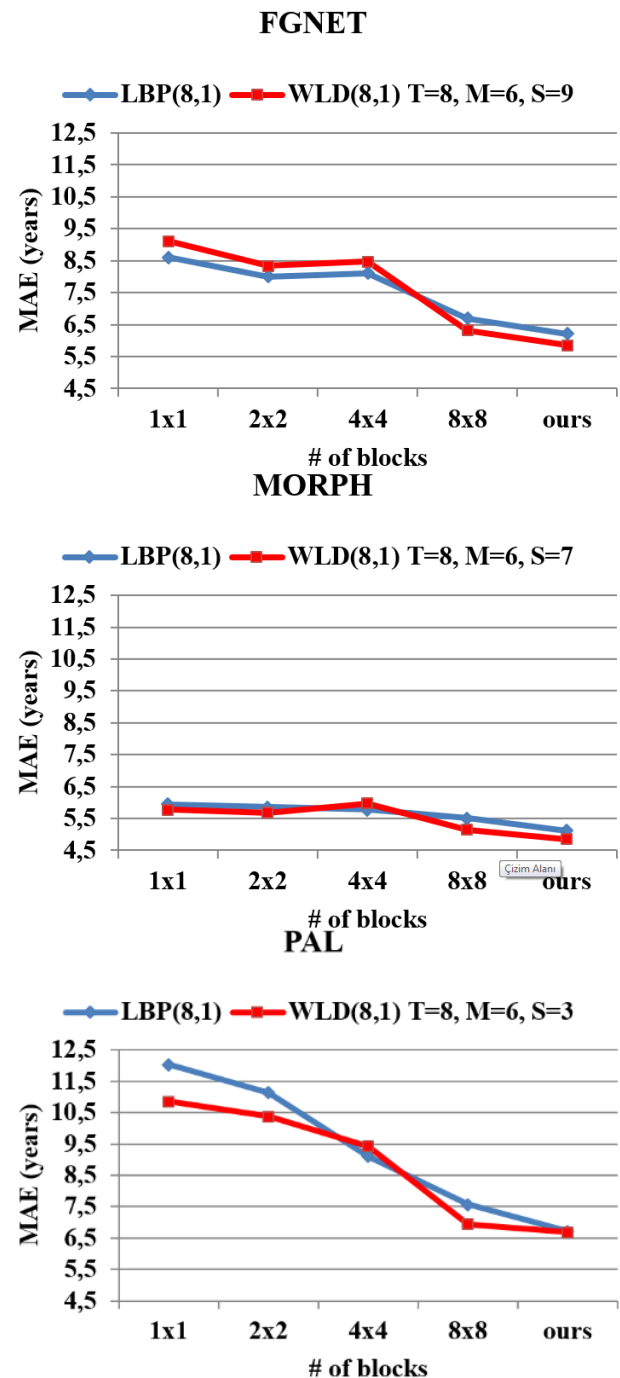
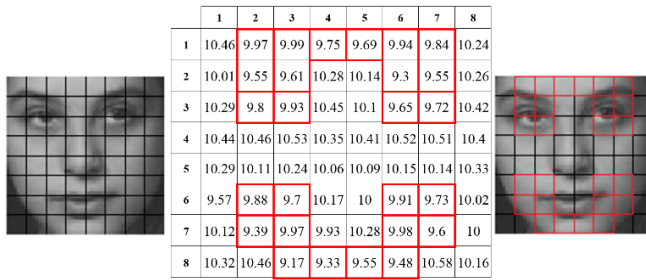


Fig. 7. The comparison of MAE's of holistic and spatial $WLD_{8,1}$ ($T = 8$, $M = 6$, $S \in \{3, 5, 7, 9\}$) and $LBP_{8,1}$ descriptors for a) FGNET b) MORPH and c) PAL databases.

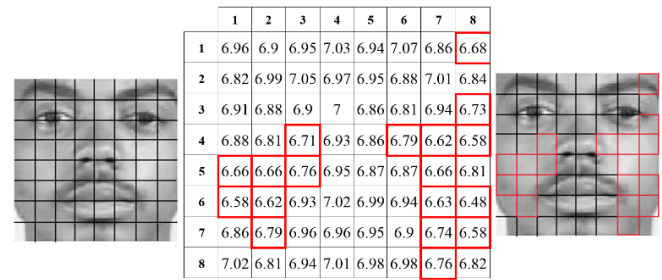
and the mouth are better than the other blocks for FGNET database. Also when we consider the intersection sections similar results are obtained. As a result 26 blocks indicated with red lines in Fig. 8 and the 23 intersection blocks with higher estimation accuracies are selected and age estimation is performed by concatenating the WLD histograms extracted only from these 49 blocks for FGNET database.

When Fig. 9 is examined, it can be seen that the age estimation accuracies of the wrinkle blocks such as near, between and under the eyes, sides of nose, and sides of mouth



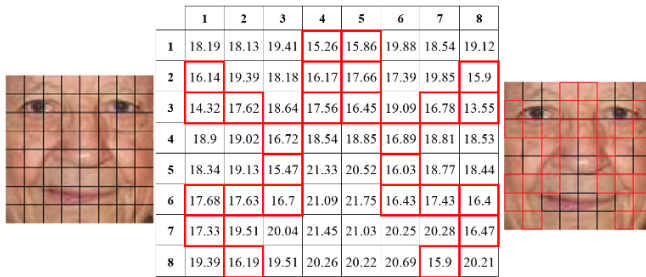
	1	2	3	4	5	6	7	8
1	10.46	9.97	9.99	9.75	9.69	9.94	9.84	10.24
2	10.01	9.55	9.61	10.28	10.14	9.3	9.55	10.26
3	10.29	9.8	9.93	10.45	10.1	9.65	9.72	10.42
4	10.44	10.46	10.53	10.35	10.41	10.52	10.51	10.4
5	10.29	10.11	10.24	10.06	10.09	10.15	10.14	10.33
6	9.57	9.88	9.7	10.17	10	9.91	9.73	10.02
7	10.12	9.39	9.97	9.93	10.28	9.98	9.6	10
8	10.32	10.46	9.17	9.33	9.55	9.48	10.58	10.16

Fig. 8. Age estimation accuracies of image blocks for FGNET database.



	1	2	3	4	5	6	7	8
1	6.96	6.9	6.95	7.03	6.94	7.07	6.86	6.68
2	6.82	6.99	7.05	6.97	6.95	6.88	7.01	6.84
3	6.91	6.88	6.9	7	6.86	6.81	6.94	6.73
4	6.88	6.81	6.71	6.93	6.86	6.79	6.62	6.58
5	6.66	6.66	6.76	6.95	6.87	6.87	6.66	6.81
6	6.58	6.62	6.93	7.02	6.99	6.94	6.63	6.48
7	6.86	6.79	6.96	6.96	6.95	6.9	6.74	6.58
8	7.02	6.81	6.94	7.01	6.98	6.98	6.76	6.82

Fig. 10. Age estimation accuracies of image blocks for MORPH database.



	1	2	3	4	5	6	7	8
1	18.19	18.13	19.41	15.26	15.86	19.88	18.54	19.12
2	16.14	19.39	18.18	16.17	17.66	17.39	19.85	15.9
3	14.32	17.62	18.64	17.56	16.45	19.09	16.78	13.55
4	18.9	19.02	16.72	18.54	18.85	16.89	18.81	18.53
5	18.34	19.13	15.47	21.33	20.52	16.03	18.77	18.44
6	17.68	17.63	16.7	21.09	21.75	16.43	17.43	16.4
7	17.33	19.51	20.04	21.45	21.03	20.25	20.28	16.47
8	19.39	16.19	19.51	20.26	20.22	20.69	15.9	20.21

Fig. 9. Age estimation accuracies of image blocks for PAL database.

which are indicated with red lines are better than other blocks. The same situation is valid when we consider the intersection blocks. For this reason the WLD histograms extracted from 26 distinct and 14 intersection blocks (totally 40 blocks) are concatenated to form spatial WLD histogram for the images in PAL database.

According to the age estimation accuracies of the 64 blocks of MORPH database, unlike other databases, a clear and regular pattern cannot be obtained as shown in Fig. 10. Only the blocks around the mouth gives better estimation accuracies than other blocks. Therefore the 18 distinct blocks indicated with red lines in Fig. 10 and the 15 intersection blocks (totally 33 blocks) with better estimation accuracies are used in spatial WLD histogram generation for MORPH database.

The results of this experiment are tabulated in Table IV. In Table IV the number of blocks before and after the block selection and the MAEs for the FGNET, MORPH and PAL databases are given. When we compared the results, it can be seen that similar age estimation performances can be obtained by using more effective blocks in spatial histogram generation. This also provides us to reduce the number of features and computational cost.

IV. CONCLUSION

In this paper a novel age estimation method using a new texture descriptor WLD is introduced. The age estimation performances of multi-scale versions of holistic and spatial WLD descriptors with various combinations of parameters are compared for finding the optimal parameters. Furthermore the block and inter-block based information is evaluated together to improve the age estimation performance. Besides, the advantage and accuracy of the proposed approach is confirmed with the experiments using LBP texture descriptor. As the dimensionality of the SWLD histograms increases, we

TABLE IV
THE MAE'S OF ALL REGIONS AND SELECTED REGIONS FOR FGNET, MORPH AND PAL DATABASES

Database	# of blocks	# of features	MAE
FGNET	113	48816	5.85
	49	21168	6.05
MORPH	113	37968	4.85
	33	11088	5.25
PAL	113	16272	6.68
	40	5760	7.22

perform dimensionality reduction using PCA to find a lower dimensional subspace. Besides age estimation accuracies of the blocks used in spatial feature extraction are calculated separately. By using the blocks with higher estimation accuracies, the dimensions of the feature vectors and the computational cost is decreased without losing much accuracy. Determining the weights of the dominant blocks so increasing the age estimation accuracy by using these weighted features is our future work.

REFERENCES

- [1] A. Hadid, M. Pietikainen and T. Ahonen, "A discriminative feature space for detecting and recognizing faces", *Proc. of Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 797–804, 2004.
- [2] M. Gonzalez-Ulloa and E. S. Flores, "Senility of the face-Basic study to understand its causes and effects", *Plastics & Reconstructive Surgery*, vol. 36, no. 2, pp. 239–246, 1965.
- [3] Y. H. Kwon and N. V. Lobo, "Age Classification from Facial Images", *Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 1–21, 1999.
- [4] M. Oravec, B. Kristof, M. Kolarik, and J. Pavlavcova, "Extraction of Facial Features from Color Images", *Radioengineering*, vol. 17, no.3, pp. 115–120, 2008.
- [5] W.-B. Horng, C.-P. Lee, and C.-W. Chen, "Classification of Age Groups Based on Facial Features", *Tamkang Journal of Science and Engineering*, vol. 4, no. 3, pp. 183–192, 2001.
- [6] M. M. Dehshibi and A. Bastanfard, "A new algorithm for age recognition from facial images", *Signal Processing*, vol. 90, no. 8, pp. 2431–2444, 2010.
- [7] A. Lanitis, C. Taylor, and T. Cootes, "Toward Automatic Simulation of Aging Effects on Face Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 442–455, 2002.
- [8] S. Kohli, S. Prakash, and P. Gupta, "Hierarchical age estimation with dissimilarity-based classification", *Neurocomputing*, vol. 120, pp. 164–176, 2013.
- [9] W.-L. Chao, J.-Z. Liu, and J.-J. Ding, "Facial age estimation based on label-sebsitive learning and age oriented regression", *Pattern Recognition*, vol. 43, pp. 628–641, 2013.
- [10] S. E. Choi, Y. J. Lee, S. J. Lee, and K. R. Park, "Age estimation using a hierarchical classifier based on global and local facial features", *Pattern Recognition*, vol. 44, no. 6, pp. 1262–1281, 2011.
- [11] X. Geng, Z. H. Zhou, and K. S. Miles, "Automatic Age Estimation Based on Facial Aging Patterns", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2234–2240, 2007.

- [12] Y. Fu and T. S. Huang, "Human Age Estimation with Regression on Discriminative Aging Manifold", *IEEE Trans. on Multimedia*, vol. 10, no. 4, pp. 578–584, 2008.
- [13] J. Lu and Y.-P. Tan, "Ordinary Preserving Manifold Analysis for Human Age and Head Pose Estimation", *IEEE Trans. on Human-Machine Systems*, vol. 43, no. 2, pp. 249–258, 2013.
- [14] A. Günay and V.V. Nابیev, "Automatic age classification with LBP", *IEEE 23rd International Symposium on Computer and Information Sciences*, pp.1–4, 2008.
- [15] G. Guo, G. Mu, Y. Fu, and T. S. Huang, "Human Age Estimation Using Bio-Inspired Features", *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 112–119, 2009.
- [16] H. Han, C. Otto, X. Liu, and A. K. Jain, "Demographic Estimation from Face Images: Human vs. Machine Performance", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 37, no. 6, pp. 1148–1161, 2015.
- [17] Y. Ma, J. Liu, Y. Yang, and N. Zheng, "Double layer multiple task learning for age estimation with insufficient training samples", *Neuro-computing*, vol. 147, pp. 380–386, 2015.
- [18] J. Xing, K. Li, W. Hu, C. Yuan, and H. Ling, "Diagnosing deep learning models for high accuracy age estimation from a single image", *Pattern Recognition*, vol. 66, pp. 106–116, 2017.
- [19] K. Li, J. Xing, W. Hu, and S. J. Maybank, "D2C: Deep cumulatively and comparatively learning for human age estimation", *Pattern Recognition*, vol. 66, pp. 95–105, 2017.
- [20] J. Huang, B. Li, and J. Zhu, "Age classification with deep learning face representation", *Multimedia Tools and Applications*, vol. 76, no. 19, pp. 20231–20247, 2017.
- [21] D. Lai, Y. Chen, and X. Luo, "Age estimation with dynamic age range", *Multimedia Tools and Applications*, vol.76, no. 5, pp. 6551–6573, 2017.
- [22] J. Chen, S. Shan, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "WLD: A Robust Local Image Descriptor", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1705–1720, 2010.
- [23] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [24] FG-Net aging database. Available: <http://sting.cycollge.ac.cy/~alanitis/fgnetaging>.
- [25] K. Ricanek Jr., and T. Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression", *IEEE 7th Int. Conf. on Automatic Face and Gesture Recognition*, pp. 341-345, 2006.
- [26] M. Minear, and D. C. Park, "A lifespan database of adult stimuli", *Behavior Research Methods, Instruments and Computers*, vol. 36, no. 4, pp. 630–633, 2004.

Asuman Günay 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 image processing, biometric security and machine learning.

Vasif V. Nابیev 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" and "Algorithms. From Theory to Applications".