

Investigating the Effects of Facial Regions to Age Estimation

Asuman Günay ^{*1}, Vasif V. NabiyeV ²

Accepted 3rd September 2016

Abstract: Aging process causes evident alterations on human facial appearance. Real world age progression on human face is personalized and related with many factors such as, genetics, living style, eating habits, facial expressions, climate etc. The wide degree of variations on facial appearance of different individuals affects the age estimation performance. In accordance with these facts discovering the aging information contained in facial regions is an important issue in automatic age estimation. Thus the facial regions emphasizing the aging information can be used for more accurate age estimation. In this context, age estimation performances of facial regions (eye, nose, mouth and chin, cheeks and sides of mouth) are investigated in this paper. For this purpose, an age estimation method is designed to produce an estimate of the age of a subject by using the texture features extracted from facial regions. In this method the facial images are warped into the mean shape thus variations of head pose and scale are eliminated and the texture information of facial images are aligned. Then the holistic and spatial texture features are extracted from facial regions using Local Phase Quantization (LPQ) texture descriptor, robust to blur, illumination and expression variations. After the low dimensional representation of these features, a linear aging function is learned using multiple linear regression. In the experiments FGNET and PAL databases are used to evaluate the age estimation accuracies of facial regions i.e. eye, nose, mouth and chin, cheek and sides of mouth, separately. The results have shown that the eye region carries the most significant information for age estimation. Also the mouth and chin, cheek regions are effective in the prediction of age. The results also have shown that, using the spatial texture features enhances the discriminative power of the texture descriptor and thus increases the estimation accuracy.

Keywords: Age estimation, Local Phase Quantization, Facial Regions, Regression.

1. Introduction

Age estimation is the process of estimating the age or age group of an individual according to his/her facial information. During the aging process, evident alterations occur on human facial appearance. These variations are personalized and affected by factors such as race, genetics, living conditions, eating habits, the frequency of facial expressions, etc. This makes the age estimation problem much harder than other facial image processing problems. Therefore the accuracy of the age estimation systems are insufficient, even the human skills about age estimation are limited. In this context, discovering the amount of aging information contained in facial regions and thus improving the accuracy of age estimation systems by using the facial regions emphasizing the aging information is an important issue in this field.

In this paper we investigate the age estimation performances of facial regions: eye, nose, mouth and chin, cheek, sides of mouth. For this purpose we designed an age estimation system using the texture features of these facial regions to produce an estimate of the age of a subject. The block diagram of the system is shown in Fig.1. In our method the input images are normalized so that the shape variations such as scale, head pose are eliminated and the facial texture is aligned. Then Local Phase Quantization (LPQ) is

used to extract holistic and spatial texture features from facial regions. After feature extraction, dimensionality reduction is performed with Principal Component Analysis (PCA). Finally an aging function is learned using Multiple Linear Regression (MLR) for age estimation. The feature extraction, dimensionality reduction and aging function learning steps are performed separately for each of the facial regions to discover the amount of aging information contained in that region.

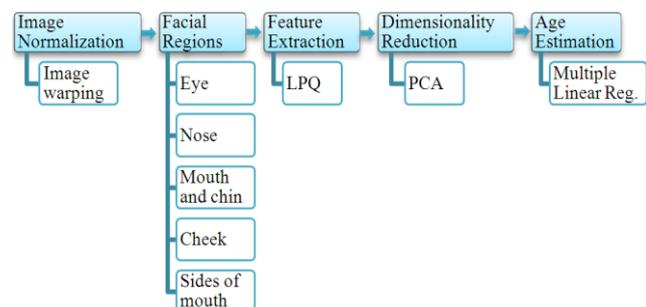


Figure 1. System overview.

The rest of the paper is organized as follows. A survey on the age estimation methods are given in Section 2. The proposed age estimation approach is explained in Section 3. In Section 4 the experimental results on various databases are reported and analyzed. Finally, the conclusions are outlined in Section 5.

2. Related Work

Age estimation systems generally consist of age image representation and age estimation modules. The aim of age image representation module is to extract the shape or texture based

¹ Department of Computer Technologies, Trabzon Vocational School, Karadeniz Technical University, Trabzon/Turkey

² Department of Computer Engineering, Karadeniz Technical University, Trabzon/Turkey

* Corresponding Author: Email: gumaya@ktu.edu.tr

Note: This paper has been presented at the 3rd International Conference on Advanced Technology & Sciences (ICAT'16) held in Konya (Turkey), September 01-03, 2016.

features from facial images. Then classification techniques are used to classify the facial images into multiple age groups or regression techniques are applied to estimate the age of the subjects.

The age image representation techniques can be grouped under five topics. The anthropometric models rely on the facial geometry. In these models the distances and the ratios of these distances are calculated using the fiducial points on the facial images. As these geometrical features can only deal with young ages, wrinkle features are used with geometric features to strengthen the classification performance for older ages [1-3]. Active Appearance Model (AAM) based age estimation methods incorporate shape and appearance information together. For this reason AAM's are frequently used in age estimation methods [4-6]. In some studies AAM features are extracted as global facial features and fused with local facial features for efficient age estimation [7]. In Aging Pattern Subspace method, the sequence of an individual's aging face images are used to model the aging process [8]. But age manifold methods use the images of different individuals at different ages to learn the common aging pattern. They utilize manifold embedding techniques to discover the aging trend in a low dimensional space [9, 10]. Appearance models are mainly focused on aging-related facial feature extraction using various texture descriptors such as local binary patterns, Gabor filters, histograms of gradients etc. [7, 11-13].

All the studies mentioned above generally use the whole face in feature extraction phases. Unfortunately the number of works examining the age estimation performances of facial regions/parts is insufficient in the literature. Lanitis [14] investigated the significance of facial parts in age estimation. In the experiments, the age estimation performances of whole face (including the hairline), internal face, upper part of the face and lower part of the face are calculated. AAMs are used to represent the shape and appearance of facial parts with model parameters. The results showed that the upper part of the face gives the minimum age estimation error than other parts. El-Dib and Onsi [15] used bio-inspired features to analyze the different facial parts: eye wrinkles (covering the eyes and the area under the eyes), internal face and whole face. They built six support vector regression and one support vector machine model to estimate the age of a subject. Their results showed that the eye region covering the eyes and the area under the eyes contains the most important aging features when compared with others.

3. Proposed Method

3.1. Image Normalization

In order to eliminate the shape variations such as head pose, scale, size, etc. and to align the texture information, image normalization is applied on facial images. Image normalization is performed by warping the facial images into the mean shape obtained from training set. The facial images in the training set are labelled with 68 landmark points as shown in Fig.2-a. Mean shape is obtained by taking the mean of all coordinates of these points. Then all the images are warped into the mean shape using Delaunay triangulation (Fig. 2-b) and affine transformation, so the landmark points of all images are matched with the mean shape and also the texture information is aligned. The result of warping process is given in Fig. 2-c. As the head poses of facial images varies, the warped images can be inclined, therefore rotation is the last step in image normalization (Fig.2-d).

3.2. Facial Regions

Investigating the aging information contained in facial regions is important for age estimation system's design. In this way, the accuracies of age estimation systems can be improved by using the facial regions emphasizing the aging information. For this purpose the facial images are divided into regions and the age estimation algorithm is applied to these regions separately. In the study the age estimation performances of eye (190×55), nose (55×70), mouth and chin (85×70), cheek (65×60) and side of mouth (40×45) regions are determined. The facial regions used in the experiments are shown in Fig. 3.

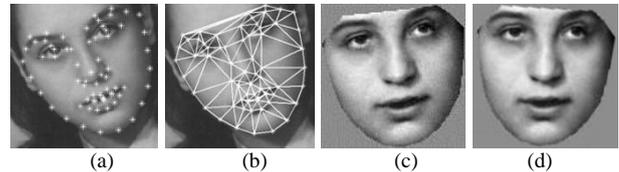


Figure 2. Image normalization a) Facial image labeled with 68 landmark points b) Delaunay triangulation c) The image after warping into the mean shape d) Rotated image.

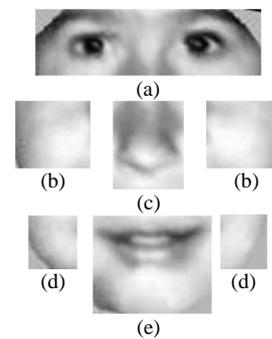


Figure 3. Facial regions a) eye b) cheeks (left and right) c) nose d) side of mouth (left and right) e) mouth.

3.3. Feature Extraction with LPQ

LPQ is a blur insensitive texture descriptor based on the blur invariance property of the Fourier phase spectrum [16]. In this method LPQ codes are computed in local image windows using discrete Fourier transform (DFT) and the results are presented as a histogram.

The spatially invariant blurring of an image can be expressed by a convolution between the image and the point spread function (PSF). In the frequency domain this is equal to $G(u) = F(u) \cdot H(u)$, where $G(u)$, $F(u)$ and $H(u)$ are the DFTs of the blurred image, the original image and the PSF, respectively. Considering the phase of the spectrum we have $\angle G(u) = \angle F(u) + \angle H(u)$. If we assume that the blur PSF $h(x)$ is centrally symmetric, $h(x) = h(-x)$, its Fourier transform is always real valued, 0 for $\angle H(u) \geq 0$ and π for $\angle H(u) < 0$. This means that $\angle G(u) = \angle F(u)$ for all $H(u) \geq 0$. In other words the phase of the observed image $\angle G(u)$ is invariant to centrally symmetric blur, at the frequencies where $H(u)$ is positive.

If the $N \times N$ neighborhood around a pixel x is denoted as \mathcal{N}_x , the two dimensional (2-D) DFT of \mathcal{N}_x is defined by,

$$F(u, x) = \sum_{y \in \mathcal{N}_x} f(x - y) e^{-j2\pi u^T y} = w_u^T f_x \quad (1)$$

where w_u is the basis vector of the 2-D DFT at frequency u , and f_x is the vector containing all N^2 pixels in \mathcal{N}_x . Only the complex coefficients of $u_1=[a, 0]^T$, $u_2=[0, a]^T$, $u_3=[a, a]^T$, $u_4=[a, -a]^T$ are considered in LPQ. a is a scalar frequency below the first zero crossing of $H(u)$ that satisfies $H(u) \geq 0$. For each pixel position this results in a vector given by,

$$F_x^c = [F(u_1, x), F(u_2, x), F(u_3, x), F(u_4, x)] \quad (2)$$

$$F_x = [\text{Re}\{F_x^c\}, \text{Im}\{F_x^c\}]^T, \quad (3)$$

where $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ return real and imaginary parts of a complex number, respectively. Then G_x is computed for all image positions, i.e., $x \in \{x_1, x_2, \dots, x_N\}$, and the resulting vectors are quantized using a simple scalar quantizer,

$$q_j = \begin{cases} 1, & \text{if } g_j \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where g_j is the j^{th} component of G_x . The quantized coefficients are represented as integer values between 0-255 using binary coding using (5), and the histogram of these integer values is used as a feature vector.

$$b = \sum_{j=1}^8 q_j 2^{j-1} \quad (5)$$

The LPQ texture descriptor represents the input image as a histogram of 256 bins. In this holistic representation, the histogram is produced without taking into account the spatial information of the pixels. The discriminative power of the texture descriptor can be enhanced by using the spatial histograms. The spatial histograms are produced by concatenating the local histograms extracted from small image blocks.

3.4. Dimensionality Reduction

In order to find a lower dimensional subspace of the extracted features and to obtain the significant features for age estimation, dimensionality reduction is performed using PCA. PCA method finds the embedding that maximizes the projected variance given by $W_{opt} = \arg \max_{\|W\|=1} W^T S W$ where $S = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T$ is the scatter matrix, x_i is i^{th} feature vector with $x_i \in R^D$ and \bar{x} is the mean of the feature vectors. By solving this problem, a set of $d \leq D$ eigenvectors associated to the d largest eigenvalues of S is obtained. Then dimensionality reduction is performed by projecting all samples on the projection subspace using $y_i = W^T x_i$ with $y_i \in R^d$.

3.5. Age Estimation

After finding the low dimensional representation of features, the age estimation problem is recast as a multiple linear regression as $L = YB + e$ where Y is the data matrix, B is the unknown parameter vector, L is the age label vector and e is the error vector. In the learning phase the unknown parameters are estimated by means of least squares, or robust regression. The regression function used in this study is a linear function given by $\hat{l} = \hat{\beta}_0 + \hat{\beta}_1^T y$ where \hat{l} is the estimate of age, $\hat{\beta}_0$ is the offset, $\hat{\beta}_1$ is the weight vector containing the coefficients for each value in the feature vector and y is the low dimensional representation of the extracted feature vector.

4. Experiments and Results

In this paper the FGNET and PAL databases are used to evaluate the age estimation performances of facial regions. FG-NET database [17] is composed of 1,002 images that were retrieved from real-life albums of 82 subjects, thus includes variations of head pose, occlusion, illumination, facial expressions, etc. The age range in this database is 0-69 years, but the images are not uniformly distributed according to the ages. This can be a disadvantage for the estimation accuracy.

The PAL aging database [18] contains 580 images of different persons taken under natural lighting conditions using a digital camera. The images include various expressions such as neutral

faces, anger, sadness or smiling. The age distribution of the images in this database is between 18 and 93 years old and also not uniform.

The age estimation performance of the system is evaluated using n -fold cross validation method. In this method all the samples are randomly partitioned into n equal sized subsamples. Then one subsample is used as test set and $n-1$ subsamples are used as training set. This procedure is repeated n times until each of the subsamples are used once as a test set. Then the average of all n estimations is considered as the system performance. In our experiments we set $n=3$.

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

$$\text{MAE} = \sum_{i=1}^{N_t} |\hat{l}_i - l_i| / N_t \quad (6)$$

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 facial regions are calculated using the features extracted with LPQ texture descriptor and regression. LPQ is performed on 5×5 local image windows. The LPQ histograms are first extracted from the whole facial region resulting the holistic description of that region. Then the region is divided into $n \times n$ blocks, LPQ histograms are extracted from each block and concatenated to obtain the spatial description of the region. The spatial features are used to enhance the discriminative power of the texture descriptor. The spatial LPQ histograms are extracted for $n=2, 3, 4, 5, 6$ and age estimation is performed using these features. The effect of using the spatial texture features of facial regions for age estimation on FGNET and PAL databases are given in Fig. 4.

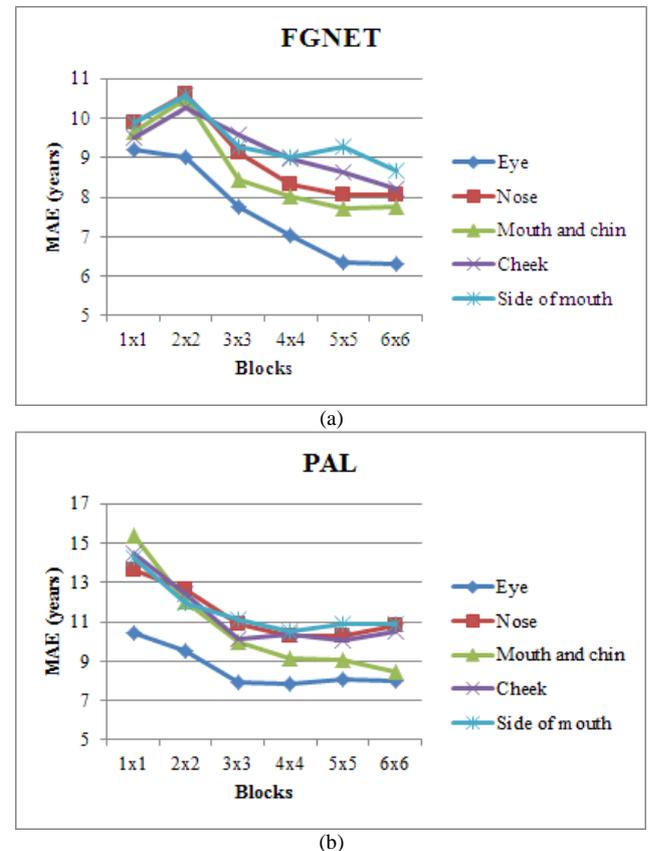


Figure 4. The effect of using spatial texture features of facial regions for age estimation a) FGNET database b) PAL database.

We can see from the figure that the holistic feature representation is not encouraging and increasing the number of blocks in spatial feature representation generally increases the estimation accuracy of the facial regions.

The age estimation accuracies of the whole face and facial regions are tabulated in Table 1. We can see from the results that the age estimation accuracy of the whole face and the eye region are close to each other. These results indicate that the eye region carries the most significant information for age estimation. Also the mouth and chin, cheek regions are more effective than other regions in the prediction of age.

Table 1. MAE's of facial regions for FGNET and PAL databases

Regions	FGNET	PAL
Whole face	5.83	6.94
Eye	6.30	7.83
Nose	8.04	10.24
Mouth and chin	7.71	8.46
Cheek	8.21	10.04
Side of mouth	8.66	10.52

5. Conclusion

In this paper the age estimation performances of the facial regions are investigated. For this purpose the facial image is divided into five regions: eye, nose, mouth and chin, cheek, side of mouth. Then LPQ is used to extract features from these regions. In the feature extraction phase holistic representation of the region is obtained by applying the texture descriptor on the whole region. Also the spatial representation is obtained by dividing the region into number of blocks, extracting features from these blocks and concatenating them. The age estimation accuracies of the facial regions are evaluated separately and the experiments on FGNET and PAL databases have shown that the most of the aging information is contained in the eye region. Moreover the age estimation accuracy of the eye region is close to the whole face. As the estimation of the whole face is better than the regions alone, determining the weights of these regions in age estimation and thus achieving better estimation accuracy is our future work.

References

[1] Kwon Y. H. and Lobo N. V. Age Classification from Facial Images, *Computer Vision and Image Understanding*, Vol. 74, No. 1, 1999, pp. 1-21.

[2] Horng W. B., Lee C. P. and Chen C. W. Classification of Age Groups Based on Facial Features, *Tamkang Journal of Science and Engineering*, Vol. 4, No. 3, 2001, pp. 183-192.

[3] Dehshibi M. M. and Bastanfard A. A new algorithm for age recognition from facial images, *Signal Processing*, Vol. 90, No. 8, 2010, pp. 2431-2444.

[4] Lanitis A., Taylor C. and Cootes T. Toward Automatic Simulation of Aging Effects on Face Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 4, 2002, pp. 442-455.

[5] Kohli S., Prakash S. and Gupta P. Hierarchical age estimation with dissimilarity-based classification, *Neurocomputing*, Vol. 120, 2013, pp. 164-176.

[6] Chao W. L., Liu J. Z. and Ding J. J. Facial age estimation based on label-sensitive learning and age oriented regression, *Pattern Recognition*, Vol. 43, 2013, pp. 628-641.

[7] Choi S. E., Le Y. J., Lee S. J., Park K. R. and Kim J. Age estimation using a hierarchical classifier based on global and local facial features, *Pattern Recognition*, Vol. 44, 2011, pp. 1262-1281.

[8] Geng X., Zhou Z. H. and Miles K. S. Automatic Age Estimation Based on Facial Aging Patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 12, 2007, pp. 2234-2240.

[9] Fu Y. and Huang T. S. Human Age Estimation with Regression on Discriminative Aging Manifold, *IEEE Transactions on Multimedia*, Vol. 10, No. 4, 2008, pp. 578-584.

[10] Guo G., Fu Y., Dyer C. R. and Huang T. S. Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression, *IEEE Transactions on Image Processing*, Vol. 17, No. 7, 2008, pp. 1178-1188.

[11] Chen C., Yang W., Wang Y., Ricanek K. and Luu K. Facial Feature Fusion and Model Selection for Age Estimation, *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG'11)*, 2011, pp. 200-205.

[12] Guo G., Mu G., Fu Y. and Huang T. S. Human Age Estimation Using Bio-Inspired Features, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 112-119.

[13] J. Liu, Y. Ma, L. Duan, F. Wang and Y. Liu, "Hybrid constraint SVR for facial age estimation", *Signal Processing*, vol. 94, pp. 576-582, 2014.

[14] Lanitis A. On the Significance of Different Facial Parts for Automatic Age Estimation, *14th International Conference on Digital Signal Processing*, Vol. 2, 2002, pp. 1027-1030.

[15] El Dib M. Y. and Onsi H. M. Human age estimation framework using different facial parts, *Egyptian Informatics Journal*, Vol. 12, No. 1, 2011, pp. 53-59.

[16] Ojansivu V. and Heikkila J. Blur Insensitive Texture Classification Using Local Phase Quantization, *Image and Signal Processing*, Vol. 5099, 2008, pp. 236-243.

[17] FG-Net aging database. Available: <http://sting.cyccollege.ac.cy/~alanitis/fgnetaging>. May 2006.

[18] Minear M. and Park D. C. A lifespan database of adult stimuli, *Behavior Research Methods, Instruments and Computers*, Vol.36, No.4, 2004, pp.630-633.