Facial Age Estimation Using Hybrid Haar Wavelet and Color Features with Support Vector Regression

Maral Arvanaghi Jadid *
Faculty of Computer and Information Technology Engineering, Qazvin branch, Islamic Azad University,
Qazvin, Iran.
M_arvanaghi@qiau.ac.ir

Mahdi Rezaei
Faculty of Computer and Information Technology Engineering, Qazvin branch, Islamic Azad University,
Qazvin, Iran.
Rezaei@qiau.ac.ir

Abstract— **Face appearance is one of the most important visual features of human which varies significantly over the aging. Therefore, automatic age estimation is a demanding research topic in the field of facial feature analysis. In the task of age estimation, feature extraction is the first influential step which highly effects on a learning method and its obtained results. The second important step of an age estimation system is training of pattern recognition method based on the extracted feature vector. Considering the importance of the feature extraction and training steps, this paper utilizes the combination of Haar wavelet transform and color moment approaches to extract full-informative and influencing feature elements of face image. To improve the training step, the paper trains a Support Vector Regression (SVR) model, based on the extracted feature vector for age estimation. Experimental results of the proposed method are performed on FG-NET and MORPH datasets and prove the superiority of the method compared with the state-of-the-art methods.**

Keywords— age estimationt; Haar wavelet; Support Vector Regression.

I. INTRODUCTION

Human face is one of the major properties that characterize the individuals in terms of gender, emotion, or age [1-4]. Among the above features, age has a direct relationship to the face due to the appearing of several variations, such as face size and skin, on human face over the aging process. By analyzing such gradual change and variations over the aging, some patterns can be extracted to estimate the age of individual for his/her own specific face image that can be considered as the task of automatic age estimation [5-7]. In other words, automatic age estimation is a process of assigning an age label to a given individual based on his/her own face image. The face aging process in males may differ from the females where the face image of a female is sound younger than the a male due to the makeups. In this regard, extraction of discriminative features for age estimation in females is more difficult than males. This is considered as an open problem [10].

A typical age estimation approach consists of two fundamental steps: Feature extraction and age estimation [11-13]. In the first step, an efficient approach is used to extract full-informative and influencing feature elements of input face image [14]. In this regard, several feature extraction methods have been proposed such as age manifold [8-9], Active Appearance Model (AAM) [15], age manifold [8-9], and AGing pattErn (AGES) Subspace [16-17]. The basic contribution of AAM is comparing the estimated appearance model of face image to the target image in order in to find the matched components of the models. Manifold learning technique is based on the modeling of the manifold data samples in a sufficient low-dimensional embedding space where a linear regression method is trained on the samples for the task of age estimation. In AGES, an appearance model [19] is utilized to extract the feature vectors of p individual face images in different ages. After that, a representative subspace is generated based on the extracted feature vector to model the aging pattern for each individual.

The second step of a typical age estimation method is training a learning approach on the feature vector. In [20], a regression method with quadratic function is proposed to improve the performance of age estimation method. Also, this work is improved by proposing three hierarchical architectures to the previous system [21]. It is important to note that the efficiency of a typical age estimation method highly depends on the above mentioned steps, where training of an appropriate learning method on an inefficient feature vector or training a weak method on an influencing feature vector will cause poor performance. To solve such problems, we propose a novel age estimation method to improve the performance of both feature extraction and learning steps. In the proposed method, a combination of Haar-wavelet transform and color moment is used to extract full-informative and influencing feature elements. Since the variation of facial features appears in the image texture, utilization of Haar wavelet features in an age estimation system certainly improves the quality of prediction. Furthermore, the variation of facial features over the aging process in a colorful image is more differentiable than gray-level image and so, extraction of the color features of input face image can be more appropriate.

First, the Haar wavelet transform in applied to the input face images, then the feature elements are extracted by using

F-norm theory. Using color moments, the color features of input face image are extracted and finally two types of extracted features will be combined as a single feature vector. Furthermore, the learning step of the proposed method is performed using Support Vector Regression (SVR) which is trained on the combined feature vector of Haar wavelet transform and color moment. To obtain the minimum generalization error of SVR, a regression function is employed to train in a high-dimensional feature space. The rest of the paper is organized as follows: The Haar wavelet and color features are introduced in Section 2. The proposed method is described in Section 3. Evaluations and experimental results of the proposed method are presented in Section 4; and the paper is concluded in Section 5.

II. BACKGROUND

In this section, two feature extraction methods that are widely used in machine vision and pattern recognition are described as follows:

A. Wavelet transform

Wavelet transform is a multi-resolution approach for image texture analysis that consists of two steps: 1) recursive filtering and 2) sub-sampling for a 2D signal (image or voice) [22]. At each level, wavelet transform function decomposes the input signal into four frequency sub-bands, HH, HL, LL, and LH as shown in Fig. 1, where H and L mean high and low frequency, respectively.

LL	LH
HL	HH

Fig. 1. First level of wavelet transform.

B. Haar wavelet transform

Haar wavelet is development of wavelet transform to simplify the computation of wavelet transform. Haar wavelet transform function consists of several coefficients which are computed as follows:

$$wc_i = \frac{p_i - p_{i+1}}{2}, \qquad i \in [1, N] \tag{1}$$

where wc is wavelet coefficient and p refers to element of one-dimensional dataset where the dataset consists of N elements. Since p_i and p_{i+1} are two adjacent elements of data, there will be $\frac{N}{2}$ wavelet coefficients for a given dataset with N elements. Furthermore, the vector of wavelet coefficients can be completed using the average of two adjacent elements of dataset as follows:

$$A_i = \frac{p_i + p_{i+1}}{2}, \qquad i \in [1, N] \tag{2}$$

where A_i shows the average of two adjacent elements p_i and p_{i+1} and thus, there will be $\frac{N}{2}$ averages for a dataset with N elements. After computation of the wavelet coefficients and the averages, they will be integrated as a single vector where the averages are stored in the first half of the vector, and the wavelet coefficients are stored in the second half [22]. This process is repeated on the average part of vector until a single average and coefficient are obtained. If the dataset is a two-dimensional image, the above one-dimensional Haar wavelet decomposition process is applied on each row and column of input image, separately.

C. Color feature

In color image, the RGB space provides a useful starting point to find color features of images, but it is not perceptually uniform [22]. One of the main approaches to extract influencing color feature elements is color moment, which is based on this assumption that any colour distribution in an image can be analyzed as a probability distribution; hence, the main parameters of a typical probability distribution, such as mean and variance, should be defined for color distribution. We provide more details for color moment in the next section.

III. PROPOSED METHOD

As mentioned before, two main steps of feature extraction and learning have a high impact on the performance of an age estimation system. In this paper we propose a novel age estimation system in which the two mentioned steps are considerably improved. The paper, proposes a combination of Haar wavelet transform and color moment as the improved feature extraction step. Using SVR model, we also propose a new learning step with further improvements in the final results. Next sections provide further details.

A. Improved feature extraction

As mentioned earlier, the proposed feature extraction method consists of two distinct approaches, Haar wavelet transform and color moment, in which their extracted feature vector will be combined as a single feature vector. In the proposed method, first, the Haar-wavelet decomposition function is applied on the input RGB face image for four times (levels) [22]. After that, the F-norm theory is utilized to reduce the dimension of the generated feature-vector. Representing the input face image by *I*, we define the sub-image of original image *I* in *i*-th level of Haar wavelet decomposition process as follows [22]:

$$I = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix} \tag{3}$$

$$I_i = \begin{bmatrix} p_{1i} & \cdots & p_{1i} \\ \vdots & \ddots & \vdots \\ p_{ni} & \cdots & p_{ni} \end{bmatrix}, \quad i = 1 \; to \; n \tag{4}$$

Also, the F-norm of the sub-image I_i is given as follow [22]:

$$\|I_i\|_F = \sqrt{\sum_{j=1}^{i} \sum_{k=1}^{i} p_{kl}^2} \tag{5}$$

Where p_{kl} is the pixel of input image in row *k* and column *l*. If we define $\Delta I_i = \|I_i\|_F - \|I_{i-1}\|_F$ and $\|I_0\|_F = 0$, then the Haar wavelet-based feature vector of matrix A can be represented as:

$$FV_{HW} = [\Delta I_1, \Delta I_2, \dots, \Delta I_n] \tag{6}$$

After the extraction of Haar wavelet-based feature vector, the color feature vector of input face image will be obtained by applying color moment approach on input image. The definition of color-moment feature comes from statistical distribution of colour values within the input image based on some specific moments such as mean, and variance. The color moment feature of input face image is defined as a triple (m_i, σ_i, s_i) where m_i represents the average of probability distribution of color channel *i*, σ_i shows the standard division of probability distribution of color matrix *i*, and s_i is the skewness of probability distribution of color matrix *i*. To compute the average color distribution of color channel *i*, the following equation is utilized [22]:

$$m_i = \frac{1}{N}\sum_{j=1}^{N} p_{ij} \tag{7}$$

where N is the total number of the pixels and p_{ij} is the j^{th} pixel of image in the i^{th} color channel. The standard deviation of color distribution of input face image for the color channel *i* is given as follow:

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(p_{ij} - m_i\right)^2} \tag{8}$$

Also, the skewness factor of color probability of the color channel *i* is computed as follow [22]:

$$s_i = \sqrt[3]{\frac{1}{N}\sum_{j=1}^{N}\left(P_{ij} - E_i\right)^3} \tag{9}$$

Since the RGB model is used in this paper for colorful image, there is a color moment feature vector with length of 9 as follow:

$$FV_{CM} = [m_i, \sigma_i, s_i] \quad , \quad i \in \{\text{red}, \text{green}, \text{blue}\} \tag{10}$$

At the final step of feature extraction, the feature vectors of the Haar wavelet transform and the color moment should be combined as a simple form as follows:

$$\begin{aligned} FV_{HWCM} = {} & FV_{HW} + FV_{CM} \\ & = [\Delta A_1, \Delta A_2, \dots, \Delta A_n, M_i, \sigma_i, S_i], \\ & i \in \{\text{red}, \text{green}, \text{blue}\} \end{aligned} \tag{11}$$

The final feature vector consists of texture information and color probability of the input face image, which are efficient and impressive for the task of age estimation. Since the variation of facial features appears in the image texture, the application of Haar wavelet features in an age estimation system certainly improves the accuracy of our final age estimation procedure. Furthermore, the variation of facial features over the aging process in a colorful image is more differentiable than gray-level image; therefore, extraction of the color features of input face image can be more appropriate. After the feature extraction, a learning method should be employed to train on the extracted feature vector for the task of age estimation. For this purpose, we utilize the SVR model for age estimation that is presented in details in the next sub-section.

B. Improved learning step

SVR is the development of Support Vector Machine (SVM) classifier that aims to find the regression function $f(x)$ with the maximum ε deviation from the test data in the problem space [23]. To introduce the concept of SVR, suppose the dataset $\{(x_1, y_1) \dots (x_l, y_l)\}, x \in R^n, y \in R$ where x_i refers to sample *i* with desired label y_i. The goal of SVR is prediction of the function $y = f(x)$, which can be defined as follows:

$$f(x) = w\phi(x) + b \tag{12}$$

where $\phi(x)$ is the feature vector of input face image, which is mapped from space x to a higher-dimensional space; parameter b is a bias and w is a vector for the regression coefficient [23]. Since these two parameters should be estimated, the SVR model uses regularized risk function as follows [23]:

$$\begin{aligned} R_{reg}[f] &= R_{emp}[f] + \lambda\|w\|^2 \\ &= \frac{1}{l}\sum_{i=1}^{l} L(f(x_i) - y_i) + \lambda\|w\|^2 \end{aligned} \tag{13}$$

where l is the total number of samples in dataset, and $L(f(x_i) - y_i)$ is ε-insensitive loss function that is given as follow [23]:

$$L(f(x_i) - y)_{\square} = \begin{cases} |(f(x) - y)| - \square, & \text{if } |(f(x) - y)| \geq \square \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

where ε defines a threshold for regression function as the maximum radial distance from data points in problem space. In SVR model, there are two slack variables ξ^+ and ξ^1 for the data pints located outside the radial distance [23]. These variables are embedded in the following object function to be minimized as follows [23]:

$$\begin{gathered} Min\colon\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left(\xi^+ + \xi^-\right) \\ y_i - w_i\phi(x_i) + b \leq \square + \xi^+ \\ w_i\phi(x_i) + b - y_i \leq \square + \xi^- \\ \xi^+ + \xi^- \geq 0 \end{gathered} \tag{15}$$

where C is a reutilized constant to make a balance between empirical error and flatness of function. Also, w is defined as

$w = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\phi(x_i)$. So, the main function of SVR model is given as follows [23]:

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\,\phi(x_i)\phi(x) + b \\ = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\,K(x_i, x) + b \tag{16}$$

where $K(x_i, x)$ is a kernel function which is set to RBF type in this paper. Furthermore, α_i and α_i^* are computed using convex optimization tool as follow [23]:

$$Max: -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\,K(x_i, x_j) - \\ \square\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) + \sum_{i=1}^{l} y_i\,(\alpha_i - \alpha_i^*) \tag{17}$$

$$\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*) = 0, \alpha_i \geq 0, \alpha_i^* \leq C \tag{18}$$

By definition of α_i and α_i^*, the bias is computed as follow [23]:

$$b = y_i - w_i\emptyset(x) - \square \quad \text{for} \quad \alpha_i \in (0, C) \tag{19}$$

$$b = y_i - w_i\emptyset(x) + \square \quad \text{for} \quad \alpha_i^* \in (0, C)$$

In this paper, the SVR model is used for the task of age estimation. To this end, the SVR model is trained on the feature vector extracted by Haar wavelet transform and color moment. In next section we show that the results of the proposed method on different datasets prove the ability of SVR model to efficiently train on the combined feature vectors.

IV. EXPERIMENTAL RESULTS

This section provides the evaluation results of the proposed method on the facial age estimation problem. The proposed method is evaluated on two common aging image datasets of FG-NET [24] and MORPH [25]. FG-NET dataset includes 640 color and grayscale face images from 80 individuals in range of newborns up to 50 years old. The MORPH dataset contains 760 color face images from 93 subjects within the range of 16 up to 77 years old. Some samples of both datasets are demonstrated in Fig. 2.

The used datasets are divided into two parts of training and testing with the size of 70% and 30%, respectively. The evaluation result of the proposed method is based on 10-fold cross validation in which the training set is shuffled and after that divided into 10 parts, and the SVR model is trained on nine parts and then validated on 10th part. Note that in each fold, the validation part is rotated through the 10 parts.

The evaluation measures utilized in the experiments are Mean Absolute Error (MAE) and Cumulative Score (CS) where the MAE computes the error between the estimated ages and the actual ones as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|A_i - P_i| \tag{20}$$

where N is the number of test samples, A_i and P_i are the actual and estimated ages of individual's face image i, respectively. Also, the CS measure is defined as follows:

$$CS(t) = \frac{N_{e\leq t}}{N} \times 100\% \tag{21}$$

where $N_{e\leq t}$ is number of test samples with an absolute error value, e, equal or lower than t.

V. RESULTS

In the first step of evaluation, the estimation results of SVR with different feature extraction methods such as SIFT [26], AAM [27] and the proposed Haar Wavelet and Color Moment (HWCM) on all datasets are analyzed in Table 1 in terms of MAE measure. According to Table 1, the best results are obtained using HWCM feature extractor. Since HWCM benefits from both texture and color features information of the input face image, it is expectable to see the capability of extracting an effective feature vector of face image by the propose method. Since, both SIFT and AAM depend on the face geometry and appearance model of the face, aging process decreases their estimation results. Furthermore, the shape of individual's changes over the aging, extraction of feature vector using the SIFT and AAM models is not effective due to the dependency of them to the face geometry, while HWCM is independent of face geometry and its variations.

TABLE 1. EVALUATION OF SVR MODEL WITH DIFFERENT FEATURE EXTRACTION METHODS IN TERMS OF MAE MEASURE

Dataset	**Feature extraction method**		
	SIFT	***AAM***	***HWCM***
FG-NET	5.35	4.37	**1.97**
MORPH	6.14	4.88	**2.43**

In the second step of evaluation, the result of the proposed method is compared with the results of the other learning methods such as C4.5[32], SVM[33], AAS[30], AGES[16], IIS-LLD[29], WAS [31], CPNN [28] and MLP on both FG-NET and MORPH datasets in terms of MAE measure (as shown in Table 2). According to this table, the proposed method also takes the best rank among other methods, and the second best rank is obtained by the IIS-LLD method. The reason of the success of proposed method refers to the hybrid feature vector that consists of texture and color features. Since the variations of face image over the aging are more obvious in image texture, extraction of Haar wavelet-based feature vector to achieve the influential best feature elements is more appropriate.

In addition, the generalization ability of SVR model to deal with the age estimation problem is proved in this section. One of the main drawbacks of IIS-LLD is utilization of the maximum entropy model to fit to real age distribution which is inappropriate for age estimation. In this comparison, the MLP model is ranked as the last method.

Fig. 2. Sample images from FG-NET and MORPH datasets. The first row shows a set of image samples of different individuls from FG-NET and the second row presents the aging process of an individual in FG-NET. The third row shows a set of image samples of different individuls from MORPH and the last row presents the aging process of an individual in MORPH.

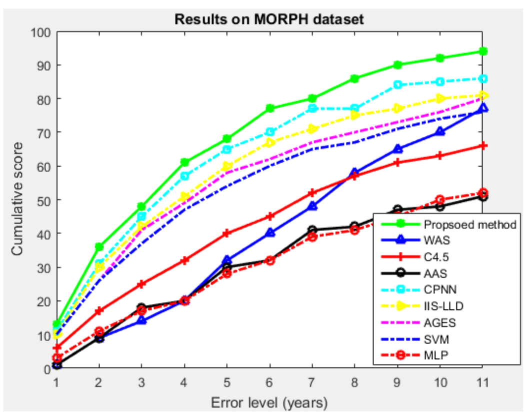

Fig. 3. Comparison of all age estimation method on FG-NET dataset in terms of CS.

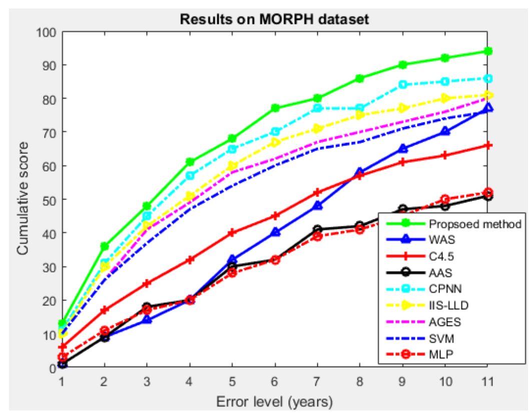

Fig. 4. Comparison of all age estimation method on MORPH dataset in terms of CS.

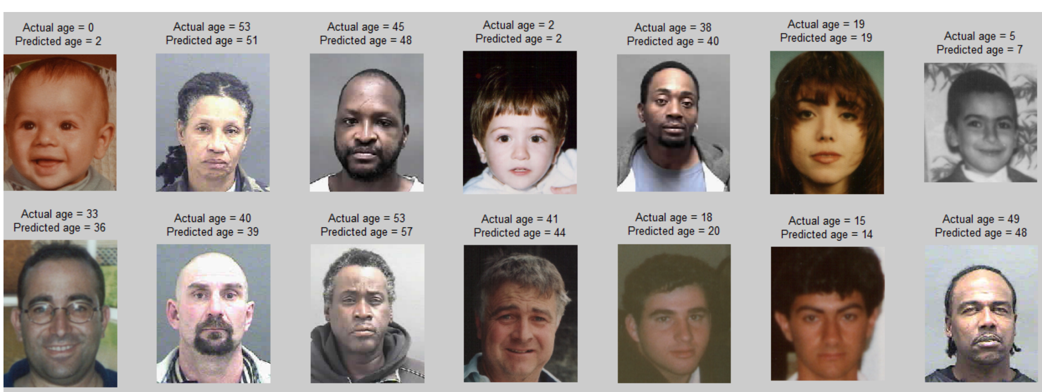

Fig. 5. Visual age estimation results of the proposed method on different images of FG-NET and MORPH datasets

TABLE 2. COMPARISON OF SEVERAL AGE ESTIMATION METHODS ON FG-NET AND MORPH DATASETS IN TERMS OF MAE

Dataset	**Age estimation methods**								
	AAS	***MLP***	***C4.5***	***SVM***	***AGES***	***WAS***	***IIS-LLD***	***CPNN***	***Proposed method***
FG-NET	11.85	14.83	9.34	7.25	6.77	8.06	5.77	4.76	**1.97**
MORPH	12.03	15.17	9.89	7.30	7.00	8.10	5.92	5.33	**2.43**

Since the cases of rotation, scaling and shifting of face inside the image are accrued frequently, a robust feature extraction method is required to efficiently extract full informative feature vector of input image. Because of that, utilization of HWCM as the feature extractor looks to be a proper solution to extract the best feature vector of face image in the aforementioned cases. According to the presented results, it is shown that combination of HWCM and SVR can perform as a complete and innovative age estimation method to obtain very promising results.

Fig.3 illustrates the results of all age estimation methods for FG-NET dataset in terms of CS measure. The horizontal axis shows the error level in unit of year and the vertical axis shows the CS value. According to this figure, the best CS values for all thresholds are obtained using the proposed method. This proves that the proposed method is able to predict the maximum number of individuals' age with minimum error compared to eight other methods. As can be seen in Fig. 4, the second best method is CPNN and IIS-LLD is ranked as the third best method. Similarly, Figure 5 shows the outcome of the proposed method on MORPH datasets in terms of CS measure. Visual results of age estimation is shown in Fig. 5 where the actual and predicted ages of each face image are mentioned for each face image. The little difference between the actual age and the estimated age proves the ability of the proposed method for real-world age estimation tasks.

Based on the presented results, the propose method takes the best estimation results on both FG-NET and MORPH datasets in terms of MAE and CS measures. The first reason of this success refers to the feature vector which contains texture and color features of face image. The second reason is SVR model which is trained on efficient and influencing feature vector and finally obtains high generalization ability in the problem of age estimation. As a results, the proposed method can be considered as an efficient and robust age estimation method that is able to predict the age of individuals in almost all situations.

VI. CONCLUSION

This paper introduced a robust technique for facial age estimation using HWCM as a feature extractor and SVR model as the learning method. As an important advantage of the proposed method, HWCM encompasses both texture and color feature of face image which are more appropriate for the task of age estimation. Analogously, the main advantage of SVR is its generalization ability to deal with age estimation problem. The proposed method was evaluated on FG-NET and MORPH datasets in terms of "mean average error" and "cumulative score" measures and the experimental result proved that the proposed method outperforms the state of the art, including eight new methods compared in this paper. As a result, we can conclude that the proposed method can be considered as an efficient and robust age estimation method which is able to predict the age of individuals in real-world applications.

References

[1] R. Singh, "Recognizing Altered Facial Appearances Due to Aging and Disguise," Signal and Image Processing for Biometrics. Springer Berlin Heidelberg, pp. 77-106, 2014.

[2] M. Valstar, "Automatic Facial Expression Analysis," Understanding Facial Expressions in Communication, Springer India, pp. 143-172, 2015.

[3] C. Shan, "Learning local binary patterns for gender classification on real-world face images," Pattern Recognition Letters, Vol. 33, No. 4, pp. 431-437, 2012.

[4] W. Gu, C. Xiang, Y.V. Venkatesh, D. Huang and H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis," Pattern Recognition, Vol. 45, No. 1, pp. 80-91, 2012.

[5] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation," IEEE Trans. on SMC-B, Vol. 34, No. 1, pp.621–628, 2004.

[6] H. Ren and Z.-N. Li, "Age estimation based on complexity-aware features," In IEEE conf. on FGR, 2006.

[7] G. Guo, Y. Fu, C. Dyer, and T. Huang, "Image-based human age estimation by manifold learning and locally adjusted robust regression," Computer Vision—ACCV, Springer International Publishing, pp. 115-128, 2015.

[8] 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.

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

[10] G. Guo and G. Mu, "Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 657-664, 2011.

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

[12] C. Fernández, I. Huerta and A. Prati, "A comparative evaluation of regression learning algorithms for facial age estimation," In Face and Facial Expression Recognition from Real World Videos, pp. 133-144. Springer International Publishing, 2015.

[13] G. Guo, Y. Fu, T. S. Huang, and C. Dyer, "Locally adjusted robust regression for human age estimation," In IEEE Workshop on Applications of Computer Vision, 2008.

[14] G. Guo, Y. Fu, T. S. Huang, and C. Dyer, "A probabilistic fusion approach to human age prediction." In IEEE CVPR-SLAM workshop, 2008.

[15] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models," In European Conference on Computer Vision, pp. 484–498, 1998.

[16] G. Panis, A. Lanitis, N. Tsapatsoulis and T.F. Cootes , "Overview of research on facial ageing using the FG-NET ageing database," IET Biometrics, 2015.

[17] C. Li, Q. Liu, J. Liu, and H. Lu, "Learning ordinal discriminative features for age estimation," In IEEE Conference on Computer Vision and Pattern Recognition (CVPR*),* pp. 2570-2577, 2012.

[18] H. Lakdawala, R. Mastan, J. Patel and E.N. Pendhari, "Analysis of Human Age Estimation Process", Analysis, Vol. 4, No. 3, 2015.

[19] H. Dibeklioğlu, A.A. Salah and T. Gevers, "A statistical method for 2-d facial landmarking", IEEE Transactions on Image Processing, Vol. 21, No. 2, pp. 844-858, 2012.

[20] Z. Li, U. Park and A.K. Jain, "A discriminative model for age invariant face recognition," IEEE Transactions on Information Forensics and Security, Vol. 6, No. 3, pp. 1028-1037, 2011.

[21] H. Dibeklioglu, F. Alnajar, A.Ali Salah and T. Gevers, "Combining Facial Dynamics With Appearance for Age Estimation", IEEE Transactions on Image Processing, Vol. 24, No. 6, pp. 1928-1943, 2015.

[22] I.H. Sarker, and S. Iqbal, "Content-based image retrieval using Haar wavelet transform and color moment" SmartCR, Vol. 3, No. 3, pp. 155-165, 2013.

[23] V. Vapnik, Statistical Learning Theory, Wiley NY, 1998.

[24] FG-NET Aging Database, http://www.fgnet.rsunit.com

[25] K. Ricanek, T. Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression", 7th In-ternational Conference on Automatic Face and Gesture Recognition, Southampton, UK, pp. 341-345, 2006.

[26] D.G. Lowe, "Object recognition from local scale-invariant features," Proceedings of the International Conference on Computer Vision, Vol. 2, pp. 1150–1157, 1999.

[27] G. J. Edwards, C. J. Taylor, and T. F. Cootes, "Interpreting face images using active appearance models". Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. p. 300, 1998.

[28] X. Geng, K. A. Smith-Miles, Z.-H. Zhou, "Facial age estimation by learning from label distributions," Proceedings of the 24th AAAI Conference on Artificial Intelligence, Atlanta, GA, pp. 451–456, 2010.

[29] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," IEEE Transactions on PAMI, Vol. 29, No. 12, pp. 2234-2240, 2007.

[30] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 34, No. 1, 621–628, 2004.

[31] A. Lanitis, C.J. Taylor, and T.F. 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.

[32] J. R. Quinlan, "*C4.5: Programs for Machine Learning"*, Morgan Kaufmann Publishers, 1993.

[33] C. Cortes, and V. Vapnik, "Support-vector networks", Machine Learning, Vol. 20, No. 3, p. 273, 1995.