

An Efficient Age Estimation System based on Multi Linear Principal Component Analysis

¹V. Tamil Selvi and ²K. Vani

¹Department of Electronics and Instrumentation Engineering,
RMK Engineering College,

²Department of Information Science and Technology, CEG,
Anna University, Chennai, India

Abstract: Problem statement: Human age estimation is an active research topic in computer vision. A vital characteristic in establishing identity of the person is the age. Age estimation from face images continues to be an extremely challenging task compared to other cognition problems. Age is a crucial factor in ascertaining the identity of a person. Age is estimated from the human face images available in the database by existing systems. They cannot estimate the age of an unknown person. **Approach:** A new age estimation system was developed to estimate the age of humans from their facial images. The proposed system consists of three processes. Face detection was the first process that normalizes the facial images. The next process extracts shape feature, frequency feature, texture feature and color feature for age estimation. Then the age is estimated using Multilinear Principal Component Analysis (MPCA). **Results:** The efficiency of the proposed system was compared with single, two and three features extracted from the facial image using FGNET and Indian Database. It is compared with SVR. The proposed method for Indian database gives the highest performance with MAE reduced by two years and with error tolerance of 10 ages's estimation rate increased by 10% when compared with support vector regression method using three features extraction. **Conclusion:** The efficiency of the system is found to be increased for three features with Indian database from the comparative analysis. The implementation results illustrates that this age estimation process effectively estimates the age from the known and unknown person's facial image with remarkable mean absolute error and cumulative scores.

Key words: Multi Layer Perceptron (MLP), Local Binary Pattern (LBP), quasi-invariants, Empirical Mode Decomposition (EMD), delaunay triangulation, SVR method, particular neuron, cumulative scores, mean absolute error

INTRODUCTION

Biometric features based age determination for a person is known as age estimation. Age estimation from face images is still very challenging compared to other cognition problems. This is primarily because many external factors influence the aging progress. The aging process can be accelerated or slowed down by physical condition, living style. The most important parts for the process of automatic age estimation are the internal parts of the face and in particular the area around the eyes.

In our work, we proposed an efficient age estimation system based on MPCA. The age estimation process was based on facial features extracted from the input image. The proposed system will effectively estimate the age from the known as well as unknown person's image. The age estimation system with proposed face detection process is based on Gabor

feature extraction method and ANN, Feature Extraction and age estimation process. Experimental results and analysis of the proposed methodology are discussed. Finally, concluding remarks are provided.

Related works: A brief review of some important contributions from the existing literature is presented as follows.

Guo *et al.* (2009) have described that estimating the human age without human intervention by employing facial image analysis has plenty of potential in real-world applications, in their study; they have introduced the age manifold learning scheme for extracting face aging features and have designed a locally adjusted robust regressor for learning and prediction of human ages.

Long (2009) have described a general distance metric learning scheme for regression problems. He conducted experiments on the public FG-NET

Corresponding Author: V. Tamil Selvi, Department of Electronics and Instrumentation Engineering,
RMK Engineering College, Anna University, Chennai, India

database by Gaussian process regression in the learned metric to validate the framework, which shows that its performance is improved over traditional regression methods.

Rhodes (2009) have described that precise age estimation is vital in a variety of settings. Their present study has been to evaluate data on the estimation of age, focusing on changes that take place to the ageing face, methodology and findings. The collected data has demonstrated that the age estimation of strange faces can be pretty precise.

Rehman and Mandic, (2010) have proposed an algorithm that extracts rotating components embedded within the signal and performs accurate time-frequency analysis, via the Hilbert-Huang transform. Simulations on synthetic trivariate point processes and real-world three-dimensional signals support the analysis.

Ruiz-Hernandez *et al.* (2010) describe an optimal tensorial configuration and compare the use of two different configurations with Multi linear Principle component analysis to reduce tensor order and Relevance vector machines as regressor. Rizon (2010) reports the human emotion assessment using Electroencephalogram (EEG).

MATERIALS AND METHODS

In our method, the age can be estimated by employing facial image. A new technique based on neural networks can be employed to estimate the age of a person. Three processes are constructed in the proposed system. Normalization of the input image is the primary process. By Gabor feature extraction method and Neural Networks the face can be detected. The next process is the extraction of facial features such as (1) Shape feature (2) Frequency feature (3) Texture feature (4) Color Feature. By employing the Multi linear Principal Component Analysis (MPCA), the age can be estimated after the features have been extracted. The proposed system is detailed as follows

Normalization: Let I be the image of size $M \times N$ in the database D which comprises lot of images such as $\{I_1, I_2, \dots, I_N\}$ and N be the total number of images in the database. It is necessary to normalize the face images for age estimation because the original images have unnecessary features including background, clothes and hair. The face image is normalized using Gabor feature extraction method and neural networks.

Gabor feature extraction method: Gabor filters, modeling the responses of the simple cells in the primary visual cortex, are simply plane waves restricted by a Gaussian envelope function. This approach not only reduces computational complexity, but also improves the performance in the presence of occlusions.

Feature Extraction algorithm for the proposed face detection method comprises of 2 major steps.

Feature point localization: In this, a 9×9 window is used to search feature points on Gabor filter responses. A feature map is constructed for the face to each of 40 Gabor filters.

Feature vector computation: Feature vectors are generated at the feature points as a composition of Gabor wavelet transform coefficients. In this, the algorithm will check all potential face contained windows and the windows around them using neural network. The result will be the output of the neural network for checked regions.

Training methodology: The objective of this study is to implement a classifier based on neural networks (Multi-layer Perceptron) for face detection. The ANN is used to classify face patterns. The MLP neural network has feed forward architecture with in the input layer, a hidden layer and an output layer. Let us consider a MLP with 3 layers, the input layer is a vector constituted by n^2 units of neurons ($n \times n$ pixel input images). The hidden layer has n neurons and the output layer is a single neuron which is active to 1 if the face is presented. The activity of a particular neuron j in the hidden layer is written by:

$$S_j = \sum_{i \in \text{input}} w_{ji} x_i, \quad x_i = f(S_j) \quad (1)$$

Equation 1 is a sigmoid function, where w_{ji} the set of weights of neuron is i , b is the threshold and x_i is an input of the neuron. Similarly the output layer activity: (Eq. 2)

$$S_j = \sum_{i \in \text{input}} w_{ji} x_i \quad (2)$$

In our system, the dimension of the retina is 27×18 pixels represent human faces and non-face, the input vector is constituted by p neurons, the hidden layer has q neurons. We are designing a feed forward neural network with one hundred neurons in the hidden layer and one neuron in the output layer. We will adjust the histogram of the image for better contrast. Then the image will be convolved with Gabor filters by multiplying the image by Gabor filters in frequency domain.

Multi-Layer Perceptron (MLP) with a feed forward learning algorithms was chosen for the proposed system because of its simplicity and its capability in supervised pattern matching. By using Gabor feature extraction and Artificial Neural Network, the face can be detected for all the images in the database. Thus we form a new face detection database FD.

Facial feature extraction: Feature extraction is a unique form of dimensionality reduction. Feature Extraction is a technique of capturing visual content of images for indexing. The tracing of those pixels in an image that have some specific features is defined as Feature extraction. The four features can be computed from an individual image in a database FD. The features are Shape Feature, Texture Feature, Color Feature and Frequency feature. Before extracting the facial features, the images are reshaped to the size $K \times L$ to reduce the timing complexity. Then, make a copy of database images and convert the images of one database from RGB to grayscale for further computation. The RGB images are only used for extracting the color feature.

Shape feature: The centroid context based shape descriptor that was introduced by the proposed system can characterize the interior of the shape. Assume that P is a posture extracted in binary form by image subtraction. Before applying triangulation, the image must be segmented in to number of regions based on the facial feature points such as eye, nose and mouth. Then perform Delaunay triangulation on each region for efficient age estimation.

Assume that P is a binary posture and by using the above technique P is decomposed into set of triangular meshes. The centroid contexts are extracted from the eyes and center of the face posture. It represents the centroid histogram of the gravity center of the posture P .

Texture feature: To extract the texture feature from the image region, we build a color texture histogram by using a Local Binary Pattern (LBP). It is defined as a gray scale invariant texture measure, derived from a common definition of texture in a local neighborhood. From the standard LBP operator both rotation and gray scale invariance can be achieved. It is stored as the LBP texture feature vector of the corresponding image.

Color feature: By employing quasi invariants method the color feature can be extracted from the image region. The derivative of an image $F_x = (R_x, G_x, B_x)^T$ is projected on three directions. These projections are termed as variants. The complementary set of derivatives known as Quasi-Invariants is constructed by removing the difference from the imitative parts of the image. The projection of the derivative on the shadow-shading direction is called the shadow-shading variant. The quasi invariant has stronger meaning than the variant.

As a result, the shadow shading specular quasi-invariant is stored as the color feature of the image.

Frequency feature: The frequency feature of the normalized image can be primarily employed for age estimation. By employing Empirical Mode Decomposition (EMD) the frequency feature can be extracted. A recently introduced technique that decomposes a signal into a sum of components, each with slowly varying amplitude and phase is used. Using the Hilbert Transform the properties of each component can be analyzed. Every component of the EMD is called an Intrinsic Mode Function (IMF).

More particularly, a real valued input signal is represented by $X(k)$ and then the application of EMD generates a set of M IMFs. The IMFs are successively obtained from the element with high frequency. Hence, the residual signal $res(k)$ has the lowest frequency. Then, for evaluating the imaginary part Hilbert Transform is employed. The Hilbert Transform is used to obtain $Y(t)$ from $X(t)$ as follows: (Eq. 3)

$$Y(t) = \frac{1}{\pi} p \int_{-\infty}^{\infty} \left(\frac{X(\tau)}{t - \tau} \right) d\tau \quad (3)$$

where, P indicates the Cauchy principal values.

A Spectrum that is calculated using the IMF is called a Huang *et al.* (1998) spectrum and it is defined by the following equation: (Eq. 4)

$$H(t, \omega) = \begin{cases} \alpha(t) & \omega(t) = \omega \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Here we can form a Hilbert spectrum contour plot $H(t, \omega)$ which displays the amplitude (z - axis) as the function of time and frequency. The spectrum H is qualitative, but certain integral quantities provide statistics.

Age estimation using MPCA: MPCA is a multi linear subspace learning method that extracts features directly from multi-dimensional objects. MPCA receives the set of face image samples of the same dimensions as input for feature extraction. We convert the feature vectors of the face images into the feature samples. This feature samples are the input for MPCA. Hence, the age is estimated from the facial image which can be performed by the MPCA which is not only for dimensionality reduction but also for age estimation process. The resultant output of the MPCA is the dimensionally reduced feature projection matrix of face images. The pseudo-code implementation of the MPCA algorithm is as follows.

Input: A set of face images.

Output: Dimensional reduced set of face images of the input face samples with maximum variation captured.

Algorithm:

- Step 1: (Pre-processing): Center the input face samples.
- Step 2: (Initialization): Calculate the Eigen-decomposition
- Step 3: (Local optimization): Calculate the mean
- Step 4: (Projection): The dimension reduced feature projection matrix after projection is obtained as:

$$\left\{ \begin{array}{l} Y_i = X_i \times_1 \tilde{U}^{(1)T} \times_2 \tilde{U}^{(2)T} \dots \times_N \tilde{U}^{(N)T}, \\ i = 1, \dots, m \end{array} \right\}$$

Where:

- Y_i = Reduced projection vector
- $\tilde{U}^{(n)}$ = Matrix to consist of the P_n eigenvectors
- \tilde{X}_i = Input face sample

After completing the training using MPCA, we obtain the reduced projection vector Y_i and then we estimate the age based on these training. When we select the input image feature from the feature vector is compared with the trained feature vector using MPCA. Then using the normalization method the age can be calculated based on the weight age assigned. Suppose, the feature vector of the test image i is b_i , then the mean feature vector of the age i over the training image set is \bar{b}_i , then the aging vector is composed by: (Eq. 5 and 6)

$$x_c(\omega) = [b_1^T, b_2^T, \dots, b_n^T] \quad (5)$$

Where

$$b_i = w_i b_i + (1 - w_i) \bar{b}_i, i = 1, \dots, n \quad (6)$$

Where $0 \leq w_i \leq 1$

In MPCA, the principal Eigen vectors are calculated as the linear combination of the training vector. An indicator vector ω^* can be obtained through minimizing the objective function $L(w)$ with respect to ω subject to the condition that $0 \leq w_i \leq 1$. The maximum element in ω^* will then indicate the suitable age for the test image.

RESULTS

Our approach is implemented in MATLAB (7.10).

FGNET database: We have used the large set of FG-NET Aging database to estimate the age. 250 images are used for training. The FG-NET Aging Database is an image database containing face images showing a number of subjects at different ages. The FG-NET Aging Database is widely used for research of age related facial image analysis. The database contains 1002 high resolution color or grey scale face images with large variations of pose, illumination and expression. There are 82 subjects in total with the age ranges from 0-69 years. FG-NET is very useful for age progression study such as estimation and simulation. The gradual results attained by the proposed age estimation system are given Fig. 1.

The face detection process using Gabor feature extraction method and artificial neural network is shown in Fig. 1.

Figure 2 represents the shape feature extracted from the facial image.

The Color feature histogram, EMD frequency histogram and LBP histogram for the facial image are represented in Fig. 3.

Indian database: The database contains 100 high resolution color or grey scale face images. There are 9 subjects in total with the age ranges from 0 to 78 years. Consequently, there is an average of 12 images per subject in the database.

Figure 4 represents the shape feature extracted from the facial image. The Color feature histogram, EMD frequency histogram and LBP histogram for the facial image in Indian database are represented in Fig. 5.

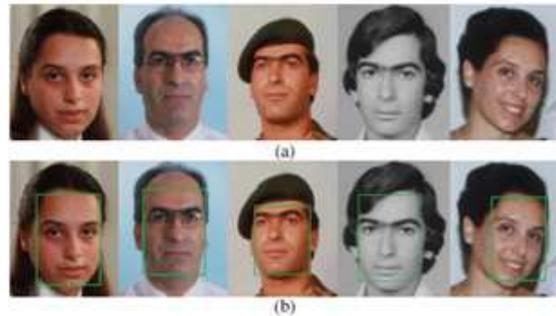


Fig. 1: Sample output obtained from face detection process (a) Original images (b) face detected images of the corresponding original images

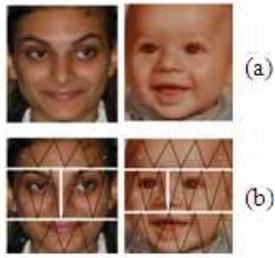


Fig. 2: Sample output obtained from the shape feature e (a) original images (b) triangular mesh and centroid context of the original images

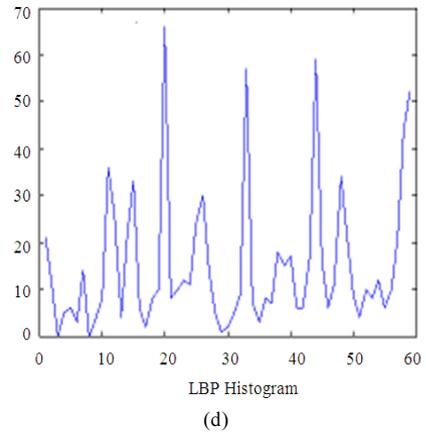


Fig. 3: Sample output obtained from the feature extraction process (a) Original image (b) Color histogram (c) EMD frequency histogram and (d) LBP texture histogram

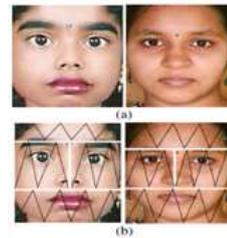
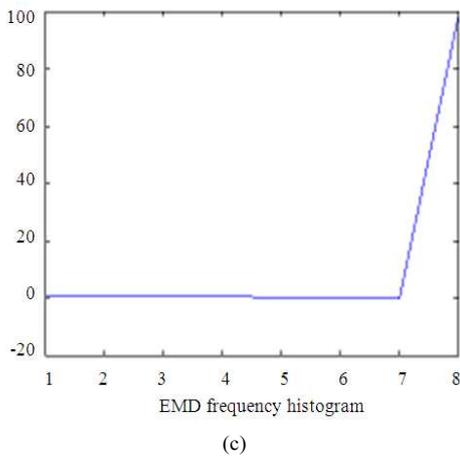
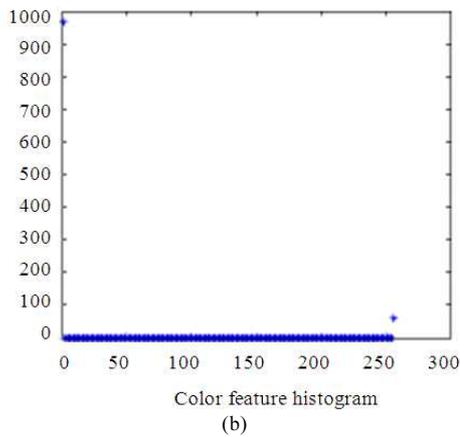
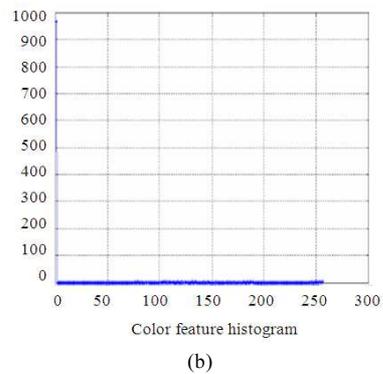


Fig. 4: Sample output obtained from the shape feature extraction in Indian Database (a) Original image (b) triangular mesh and centroid context



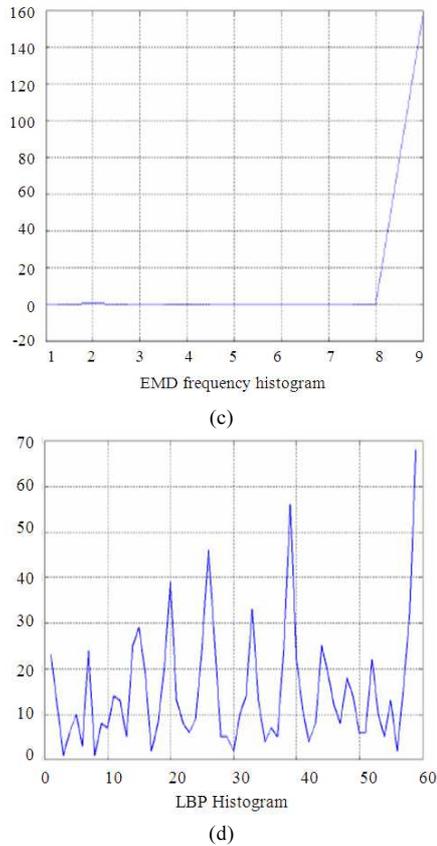


Fig 5: Sample output obtained from the feature extraction process (a) Original image (b) Color histogram (c) EMD frequency histogram and (d) LBP texture histogram

Table 1: MAE for single feature extraction

Facial Features	Mean Absolute Error (MAE) in years			
	FGNET database		Indian database	
	SVR method	Proposed method	SVR method	proposed method
Color	8.71	7.86	6.790	5.870
Texture	8.23	7.35	6.412	5.479
Shape	8.05	7.12	6.158	5.201
Frequency	8.54	7.69	6.610	5.723

Table 2: MAE for two features extraction

Facial Features	Mean Absolute Error (MAE) in years			
	FGNET database		Indian database	
	SVR method	Proposed method	SVR method	Proposed method
Color+shape	7.92	6.21	6.31	4.35
Color+texture	8.25	6.51	6.43	4.50
Color+frequency	8.57	6.97	6.49	4.59
Shape+texture	7.12	5.42	6.02	4.12
Shape+frequency	7.41	5.73	6.19	4.21
Texture+frequency	7.67	6.05	6.25	4.28

Cumulative scores for three features extraction

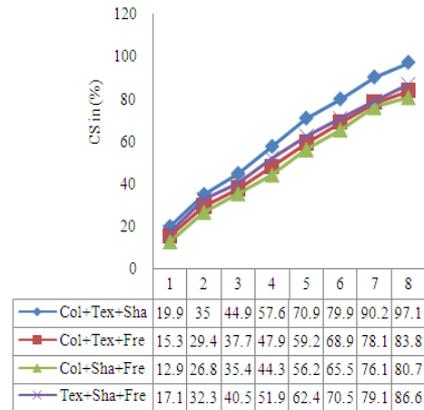


Fig. 6: Cumulative Score using MPCA in Indian database

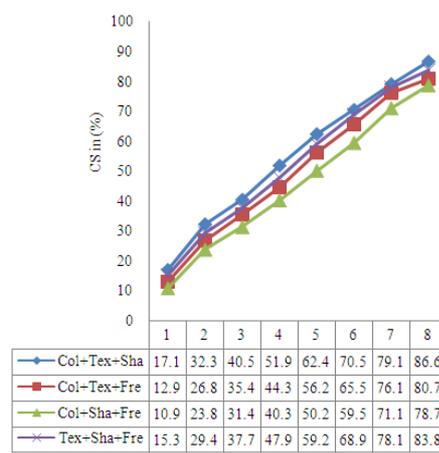


Fig. 7: Cumulative Score using SVR in database

Table 3: MAE for three features extraction

Facial Features	Mean Absolute Error (MAE) in years			
	FGNET database		Indian database	
	SVR method	Proposed method	SVR method	Proposed method
Color+texture+shape	7.54	5.780	5.41	3.950
Color+texture+frequency	8.43	6.515	6.3	4.605
Color+shape+frequency	8.18	6.510	6.02	4.370
Texture+shape+frequency	7.71	6.270	5.62	4.125

The Mean Absolute Error (MAE) is defined as the mean of absolute error between the estimated ages and the ground age: (Eq. 7)

$$MAE = \frac{\sum |I_k - I_k^*|}{N} \quad (7)$$

Where:

- I_k^* = The estimated age for the sample
- I_k = The ground truth age of the sample
- N = The total number of testing images

Age estimation results: Table 1 represents the MAE for both proposed and SVR method for single feature extraction using SVR method and proposed method in FGNET database and Indian database.

In this Table 1, the MAE for shape feature is less when compared to the other features.

In this Table 2, the MAE for shape and texture features is less when compared to the other features.

From this table, we observe that the MAE for color, shape and texture features is less when compared to the other features.

Table 1-3 represents the MAE for both proposed and SVR method. From this, table we observe that our proposed method is efficient than the SVR method.

The cumulative score is defined as: (Eq. 8)

$$CS(j) = N_{e \leq j} / N * 100\% \quad (8)$$

where, $N_{e \leq j}$ is the number of test images on which the age estimation makes the absolute error no higher than j years. We propose the estimation error over than 10 years is unacceptable and plots CS curves with $j \leq 10$.

Cumulative score for Indian database using MPCA:

Figure 6 represents the cumulative score for different combination of three features. Human estimation has an error about 4-5 years, which can be taken as a baseline of automatic age estimation. In our proposed method, error tolerance of 8 ages estimation rate attains 97.1% in three features extraction using MPCA for Indian database. It is a remarkable CS % compared to SVR method.

Cumulative score for Indian database using SVR:

From the Fig. 7, we observe that our method will efficiently estimate the age of a human from its face image compared to existing SVR methodology.

DISCUSSION

Color, Shape and texture features are having an important role in age estimation. Basically the age can be estimated by using the shape feature. The shape of the face is different for each person such as child, youth and old person. Likewise, the texture feature is also

different for each person, because the old person has more wrinkles in his face. Color feature will eliminate the imitative parts of an image. The age estimation using color, shape and texture has less error rate than the frequency feature.

Also we observe that the MAE for Indian database is less than the FG-NET database. Because, in Indian database the images have different texture, color and shape when compared to FGNET database.

Cumulative scores for Color, shape and frequency features is less for the age ranges from 1 to 8 years compared to other combinations. Also CS % for Indian data base is less than FGNET database.

CONCLUSION

In this study, we have implemented a frame work for face image based automatic age estimation. In this study, we have proposed the age estimation system based on MPCA method. Initially the face can be detected by using the Gabor feature extraction method and Artificial Neural Networks. This method will effectively detect the face and will reduce the timing complexities. The estimation process is based on the facial feature extracted from the normalized images. Finally, the age estimation process is done by employing Multi linear Principal Component Analysis (MPCA) method. Hence, the age is estimated from the facial image which can be performed by the MPCA which is not only for dimensionality reduction but also for age estimation process. The implementation results illustrates that this age estimation process effectively estimates the ages for the facial images from the database using single, two and three features. This could be visualized from the comparative analysis. Thus we conclude that an efficient age estimation based on MPCA gives better age estimation compared to SVR method in FGNET and Indian Databases.

REFERENCES

Guo, G., G. Mu, Y. Fu, C. Dyer and T. Huang, 2009. A study on automatic age estimation using a large database. Proceeding of the 12th IEEE Conference on Computer Vision, Sept. 29-Oct. 2, IEEE Xplore Press, Kyoto, pp: 1986-1991. DOI: 10.1109/ICCV.2009.5459438

Long, Y., 2009. Human age estimation by metric learning for regression problems. Proceeding of the 6th International Conference on Computer Graphics, Imaging and Visualization, Aug. 11-14, IEEE Xplore Press, Tianjin, pp: 343-348. DOI: 10.1109/CGIV.2009.91

- Rizon, M., 2010. Discrete wavelet transform based classification of human emotions using electroencephalogram signals. *Am. J. Applied Sci.*, 7: 878-885. DOI: 10.3844/ajassp.2010.878.885
- Rhodes, M.G., 2009. Age Estimation of faces: A review. *Applied Cognitive Psychol.*, 23: 1-12. DOI: 10.1002/acp.1442
- Ruiz-Hernandez, J.A., J.L. Crowley and A. Lux, 2010. "How old are you": Age estimation with tensors of binary gaussian receptive maps. *Brit. Mach. Vision Assoc.*,
- Rehman, N.U. and D.P. Mandic, 2010. Empirical mode decomposition for trivariate signals. *IEEE Trans. Sign. Proc.*, 58: 1059-1068. DOI: 10.1109/TSP.2009.2033730