ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 9(4), DOI: 10.17485/ijst/2016/v9i4/72315, January 2016* ISSN (Online) : 0974-5645

Challenging Aspects for Facial Feature Extraction and Age Estimation

A. Deepa* and T. Sasipraba

Computer Science and Engineering, Sathyabama University, Chennai – 600119, Tamil Nadu, India; deepa.sathyabama@gmail.com, sasipraba@yahoo.com

Abstract

In this paper we have discussed the steps for facial age estimation and a comparative study of various methodologies in each step has been briefed. The face identification, tools for extraction, feature normalization, features to be extracted is all explained. In spite of various challenges, feature extraction is a vital step based on which the classification and estimation of age are done. From the comparative study DCT method has provided better results across all age groups.

Keywords: DCT (Discrete Cosine Transform), HMM (Hidden Markov Models), LBP (Local Binary Pattern), MAE (Mean Absolute Error), PCA (Principle Component Analysis)

1. Introduction

Face recognition technologies provide improved performance which paves way for its deployment in various applications from web cam based system to high end biometric accesses. The various challenges in facial age estimation are manual locations of landmark points on each face image is needed, capturing of images under controlled conditions (e.g., frontal pose, normal illumination, neutral expression as in Figure 1), limited inter-age group variation, diversity of aging variation, availability of data. There are large shape and texture variations over a long period. Perceived face age depends on global non facial factors and is difficult to collect face images of same person over long time period. Age variations exist due to external factors and also lacks in quantitative measurements

2. Literature Survey

Existing methods[1-4] for facial age estimation typically consist of 2 main steps: Image representation and age prediction. The general models used for representing images are Active Shape Model (ASM), Active Appearance Model (AAM), Craniofacial Growth Model, Aging Pattern Subspace, Manifold Learning whereas for age estimation multiclass classification problem or regressing problem. In ASM model feature extraction is done by characterizing changes in process of facial appearance by face contours. The AAM approach uses the geometric ratio of local face features combined with wrinkle analysis. Using these, classification is made into 3 groups as young, adults and elders. In[5] an aging function that used parametric model for human faces is constructed. The processes of automatic age progression, age estimation, face recognition across age was compared with neural network approach. The feature extraction was done using Principle Component Analysis (PCA) and Active Appearance Model (AAM). In[6] a hybrid approach for estimating age using Active Shape Model (ASM) and Radon, DCT Transformation was constructed. Support Vector Regression (SVR) was used for learning based classification and regression.

** Author for correspondence*

Figure 1. Challenges: varying poses, illumination and expressions.

3. Methodologies used

The major steps in age estimation are identifying the face in an image, extracting the features and classifying the age of the face[7-9]. The Local features and global features are combined together to obtain a more accurate model. The global features considered are AAM, distance, face angle and ratio. The local features considered are wrinkles and texture.

3.1 Face Identification

The most commonly used 2D methods for Face identification[10,11] are principle component analysis, HMM(Hidden Markov Models),template matching model.

- Eigen face method is one of the mostly used methods in face identification. Eigen face method is also called Principle Component Analysis (PCA). Eigen features such as eigen eyes, eigen nose, and eigen mouth are used to compensate the negative effects of changing facial expressions and appearance. The result are better only in normal condition and does not suits in case different lighting conditions and different facial expressions. Therefore, eigen face method is not applicable for a successful face identification system.
- Hidden Markov Models (HMM) are used in facial identification methods. HMM was tested with 200 neutral frontal faces with high resolution. Eyes, mouth, nose, forehead regions are represented in this algorithm. 180 of 200 tested faces are found to be recognized successfully in this study. HMM utilizes one dimensional sequences which can be temporal or spatial. So technique which converts images to one dimensional sequence is used as helper. Band sampling technique is used to convert images to one dimensional observation sequence. However the classification time was very expensive (10 seconds for a face) and therefore it is not an applicable method for video based facial identification systems.
- Template matching is another used technique for face identification. Template matching is based on comparing faces with a suitable metric. Features used are eyes, nose, mouth, and the whole face. The template of an example face with the faces in the training set requires much more time than geometrical feature methods. Therefore, template matching approach is not suitable for the face identification.
- A widely used technique which uses geometrical features of the face is geometrical feature matching. In this approach, the main facial landmarks such as; eyes, eyebrows, mouth, nose, and the overall shape of the face are used for identification. These landmarks are represented as vectors. The geometric features seems to provide good performance results. So in this paper we consider geometric features also for facial feature extraction.

3.2 Determination of Tool for Feature Extraction

The various methodologies used for facial feature extraction are Local binary patterns (LBP), Histograms of Oriented Gradients (HOG) and Discrete Cosine Transform (DCT).

3.2.1 Local Binary Pattern(LBP)

The LBP is an illumination invariant descriptor method. Here the face image is divided into 8x8 pixels blocks and the LBP feature is calculated. The pixel in each block is compared with the 8 neighbors. This comparison phase results in an 8-digit binary number which is then converted to decimal format. A 256 bin histogram is computed and values of the bins of histogram are concatenated to construct the local binary pattern feature vector of the face. The feature vector can then be processed using random forest (Random Trees), to obtain the classifier.

3.2.2 Histogram of Oriented Gradients (HOG)

HOG is used for human detection and pedestrian detection. In this method 8x8 pixels blocks are used. The face image is divided into blocks, horizontal and vertical gradients are extracted. The magnitudes of the gradients are scaled and the weights are calculated by these magnitudes. HOG method constructs nine bins for -∞ to +∞. 2x2 pixels block are accumulated to yield a block of 12x12 pixels, and the weights are multiplied within the block so that weights are scaled. The descriptor of a block is obtained by appending four scaled histograms within that block. Appending obtained feature descriptors yields to the overall feature vector. This vector is used as the input of classifiers.

3.2.3 Discrete Cosine Transform(DCT)

DCT[12] is an invertible linear transform which expresses a finite sequence of data points as a sum of cosine functions. Transformation of original signal to frequency domain and vice versa are possible by DCT and inverse DCT. 2D-DCT overcomes the challenges such as illumination angles, face occlusions, colors and pose. DCT is widely used as a feature extraction[12] and compression method in various applications due to its properties such as de-correlation, energy compaction, separability and orthogonality. DCT method represents local regions of an image. It retrieves facial features from various frequency bands i.e minimum, medium and maximum frequency bands. It is data independent model and preparation of representative set of training data set is not required. DCT provides frequency information which can be utilized in handling charges in facial appearance. In this method the image is divided into (NxN) blocks.

DCT is performed on overlapping 8X8 pixels block and 15 DCT coefficients are extracted diagonally scanning the upper left part of the DCT coefficient block. As in Figure 2. DCT is applied on the entire face image. Low frequency DCT coefficients are discarded from feature representation to proved robustness against illumination variations. Feature extraction is done by projecting the whole face and facial features to eigen spaces. The classification is done by selecting different DCT coefficients with sliding window of size M and performing classification within each frequency band. The band that provides maximum separation between the closest two candidates is chosen as most reliable band. The algorithm then normalizes the features. Most of the humans have same facial features configuration so the position of facial features can be roughly estimated. After facial alignment eye center positions are to be determined. It has two points first to decide the frequency band for classification and next block based DCT is used to extract the features. The subset of the detained DCT coefficients are vector quantized and the classification is done using nearest neighbor classifier. When the block size is increased, the performance is improved. 16X16 and 32X32 when considered. DCT coefficients of image block are used as observation vectors of an embedded HMM(Hidden Markov Model).

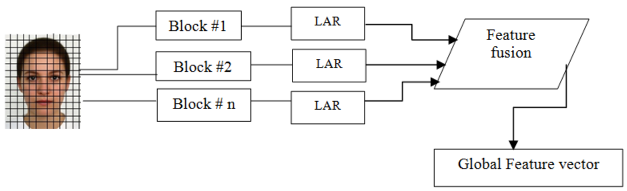

Figure 2. DCT applied on input image.

3.3 Feature Normalization

Normalization is a process that changes the range of pixel intensity values. Since normalization enhances the contrast of the image, it is also known as contrast stretching or histogram stretching. Normalization is done so as to transform the image such that parameters are mapped onto normalized values (certain fixed interval). In normalizing a facial image, the input face image is divided into sixteen local regions and the edge levels are identified. The desired regions with low complexity and large gray scale values are used as reference points in estimating the direction of illumination and error function. With these reference values the surrounding regions are optimized and the finding the range of gray levels, range stretching is done to obtain the normalized face image. The Retinex algorithm is used for normalization of face image based on illuminant direction estimation by[6]. Using Histogram Equalization (HE), the illumination adverse are rectified. Various equalization algorithms are used by various researchers. Among block equalization, orientation equalization, local equalization and adaptive equalization, local equalization eliminates illumination problems along with shadow effects.

3.4 To be Determination of Features Extracted

Feature extraction is the key for face segmentation and recognition. With respect to the conclusions of various researches, facial feature extraction is very vital and geometric feature extraction provides more accurate results. The fiducial points that are to be extracted are as shown in Figure 3 and Table 1.

Figure 3. Facial feature points.

Table 1. Features to be extracted

1	Top of the head
2	Left eyebrow outer corner
3	Left eye brow inner corner
4	Right eye brow outer corner
5	Right eye brow inner corner
6	Right eye outer corner
7	Right eye inner corner
8	Left eye outer corner
9	Left eye inner corner
10	Left most point of face
11	Right most point of face
12	Left nose corner
13	Right nose corner
14	Top of the nose
15	Left corner of mouth
16	Middle of mouth
17	Right corner of mouth
18	Tip of chin
19	Midpoint of eyebrow inner corners
20	Pupil of right eye
21	Pupil of left eye

The horizontal and vertical distances are computed. The horizontal distances that are to be considered are shown in Table 2.

The vertical distances that are to be considered are shown in Table 3.

Table 2. Horizontal distances to be considered

1	Width of left eye brow
2	Width of right eye brow
3	Width of left eye
4	Width of right eye
5	Eye outer corner distance
6	Eye inner corner distance
7	Eyebrow inner corner distance
8	Nose corner distance
9	Nose corner and middle point distance
10	Width of mouse
11	Width of face
12	Width of nose corner to right most point of face

Table 3. Vertical distances to be considered

1	Top of head and chin points distance
2	Top of head and nose middle points distance
3	Nose and mouth middle point distance
4	Mouth middle and chin point distance
5	Nose middle and chin point distance
6	Midpoint of eyebrow inner corners

With the 21 points and the relationship between the feature points the estimation can be done.

3.5 Age Group Classifiers

All the age estimation models provide a classifier with accuracy in certain age groups. The widths of age group found were by 10 to 20 years. The classifiers that are used are neural network classifier, nearest neighbor classification, KNN, hierarchical classifier, AGES, Support Vector Machine. The Figure 4 depicts the various classifiers used and their MAE.

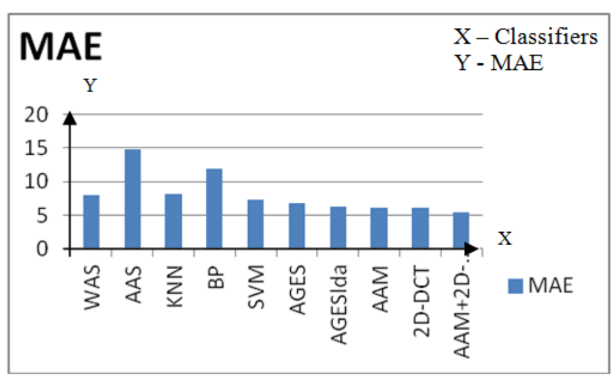

Figure 4. Classifiers and their mean absolute error rate.

Based on the results DCT method provides better accuracy. In DCT, the ratios between facial landmarks and the wrinkles on the face are determined. The information about the relations between facial feature points and the wrinkles are extracted with the DCT features of the face. Our random forest classifier uses these DCT features as the input. Random forests are an ensemble learning method for classification, regression and other tasks. Multiple decision trees are constructed at the training stage. The output is the mode of the classes (classification) or mean prediction (regression) of the individual trees. During the training phase, DCT feature vectors of all training images are extracted and these feature vectors are stored a file. After this phase, random forest classifier is trained with these features. LBP features of some selected regions of the face are computed. Wrinkle features around some regions on a face can give some clues about the age of a person. Wrinkle structures around the eyes, cheeks and foreheads have different characteristics which can differ by age. The calculation of LBP features yields to a 256 bin histogram. The vector with the values of the bins can be used as a feature vector. During the training phase, LBP feature vectors of all training images are extracted and these feature vectors are written to a file. After this phase, the random forest classifier is trained with these features.

3.6 Databases

3.6.1 MORPH2

MORPH2 is a database of mugshot images, with associated metadata giving the age, ethnicity, and gender of each subject in the database. This mugshot dataset was acquired from the Pinellas County Sherrif's Office (PCSO). This dataset includes information of the date of birth also. MORPH consists of face images of adults with varying ages. The data base has two albums MORPH album 1 and MORPH album 2. The MORPH album 1 consists of 1690 images of 515 individuals. The MORPH album 2 consists of 15204 images of 4000 individuals. The metadata of the images such as age, sex, ethnicity, height, weight are all available. The MORPH-2 database consists of 55608 color images of 13673 subjects of the age between 16 and 99 years, where 47057 images correspond to male persons and 8551 to female persons. 42897 of these images depict black faces, 10736 white, 1753 Hispanic, 160 Asian, 57 Indian and 5 faces are of other ethnicities. The images have varying resolutions of either 200x240 or 400x480 pixels. While this dataset is highly imbalanced towards black male persons and missing images of persons below the age of 16, this adds an additional challenge, which could also point out the generalizability of machine learning approaches.

3.6.2 FERET

FERET is a comprehensive database that addresses multiple problems related to face recognition such as illumination variations, pose variations, facial expressions. The images are also available with separation of age and having expressions of the subjects. The FERET dataset has gallery set containing 1196 images and duplicate probe set containing 234 images.

3.6.3 LFW

The Labeled Faces in the Wild (LFW) dataset contains faces of 5749 individuals (4263 male, 1486 female) collected from the web using a Viola-Jones face detector. Of these there are 1680 people for which more than one image is available. As a result there are 10256 male images and 2977 female images. These color images have a resolution of 250x250.

3.6.4 FDDB

Face Detection Data Set and Benchmark (FDDB), a data set of face regions designed for studying the problem of unconstrained face detection. This data set contains the annotations for 5171 faces in a set of 2845 images taken from the faces in the wild data set.

3.6.5 CMU Multi-PIE

The CMU Multi-PIE face database contains more than 750,000 images of 337 person collected in between five months period in various illuminations and expressions. Various facial points to be considered in frontal view are as shown in Figure 5.

3.6.6 FG-NET

FG-NET consists of 1,002 images of 82 individuals.6-18 images are available for per subject in the age range of 0 to 69. The database provides 68 landmark features to be identified manually. There is a mixture of 1002 color and grayscale images, which were taken in totally uncontrolled environments the metadata information such as image size, age, gender, spectacles, hat, mustache, beard, horizontal pose, vertical pose are all available. The experiments on FG-NET used the entire dataset, using a Leave-One-Person-Out (LOPO) protocol. FG-NET markup scheme of the face database is used within the FG-NET. The images from the bioid face database are selected with several additional feature points, which are very useful for facial analysis and gesture recognition. Additional feature points, very useful for facial analysis and gesture recognition are also selected. The dataset contains more than 1500 gray level images of 23 different test persons. For comparison reasons the set also contains manually set eye positions

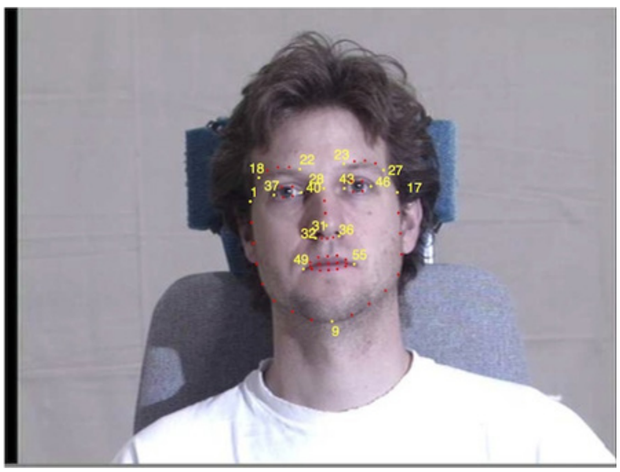

Figure 5. Facial points considered in frontal view.

4. Conclusion

Facial age estimation is one of the tedious research which paves attention to various challenges and grouping them to a single algorithm. Based on the effects of various researchers, methodologies providing better results in each stage are analyzed. From these studies, the DCT method provides better extraction of the facial features and the global features. The global feature extraction helps in identifying the age group classification. The classification method using DCT and Random Forest Classifier. In face identification module, feature extraction methods like LBP, DCT Mod2 and HOG are used and compared against each other. Then the feature extraction and classification methods are explained. The DCT feature extraction method gives the better accuracy results for a limited class count and varying training sample count. The main challenge lies in identifying the best combination of the features (distance, ratio and angle). Future work involves in selection of best combination of features which will help in liberalizing the values and will automatically improve the efficiency of the age classifier.

5. References

1. Moghadamfard H, Khanmohammadi S, Ghaemi S, Samadi F. Human age-group estimation based on ANFIS using the HOG and LBP features. International Journal on Electrical and Electronics Engineering (ELELIJ). 2013 Feb; 2(1):21–9.
2. Liang Y, Wang X, Zhang L, Wang Z. A hierarchical framework for facial age estimation. Mathematical Problems in Engineering. Hindawi; 2014; 2014:8.
3. Liu J, Maa Y, Duan L, Wang F F, Liu Y. Hybrid constraint SVR for facial age estimation. Hybrid constraint SVR for facial age estimation. 2014 Jan; 94:576–82.
4. Smeets D, Claes P, Hermans J, Vandermeulen D, Suetens P. A comparative study of 3-d face recognition under expression variations. IEEE transactions on Systems, Man and Cybernetics—Part C: Applications and Reviews. 2012 Sep; 42(5):621–8.
5. Lanitis A, Draganova C, Christodoulou C. Comparing different classifiers for automatic age estimation. IEEE Transactions on Systems, Man and Cybernetics—Part B: Cybernetics. 2004 Feb; 34(1):624–8.
6. Suo J, Zhu SC, Shan S, Chen X. A compositional and dynamic model for face aging. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010 Mar; 32(3):385–401.
7. Li Z, Park U, Jain AK. A discriminative model for age invariant face recognition. IEEE Transactions on Information Forensics and Security. 2011 Sep; 6(3):1028–37.
8. Ramanathan N, Chellappa R. Face Verification across Age Progression. IEEE Transactions on Image Processing. 2006 Nov; 15(11):3349–61.
9. Fu Y, Zheng N. M-Face: An appearance-based photo realistic model for multiple facial attributes rendering. IEEE Transactions on Circuits and Systems for Video Technology. 2006 July; 7:830–42.
10. Lanitis A, Taylor CJ, Cootes TF. Towards automatic simulation of aging effects on face images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002 Apr; 24(4):442–55.
11. Ling H, Soatto S, Ramanathan N, Jacobs DW. Face verification across age progression using discriminative methods. IEEE Transactions on Information Forensics and Security. 2010 Mar; 5(1): 82–91.
12. Dharavath K, Talukdar FA, Laskar RH. Improving face recognition rate with image preprocessing. Indian Journal of Science and Technology. 2014 Aug; 7(8):1170–5.