International Journal of Research
Available at https://edupediapublications.org/journals

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018

Integrating Global and Local Features of Facial Image in FG-Net Database

J. Suneetha
Associate Professor, Department of Information Technology,
Malla Reddy College of Engineering and Technology, Maisammaguda,
msuneetha25@gmail.com

Abstract— *This paper proposes a novel age estimation method - Global and Local feature based Age estimation (GLAAM) - relying on global and local features of facial images. Global features are obtained with Active Appearance Representations (AAR). Local features are extracted with regional 2D-DCT (2-dimensional Discrete Cosine Transform) of normalized facial images. GLAAM consists of the following modules: face normalization, global feature extraction with AAR, local feature extraction with 2D-DCT, dimensionality reduction by means of Principal Component Analysis (PCA) and age estimation with multiple linear regressions. Experiments have shown that GLAAM outperforms many methods previously applied to the FG-NET database. To integrate the global and local features, a feature level fusion approach is used.*

Keywords—2D-DCT; AAR; Age Estimation; PCA; Regression

1. Introduction

The wide-ranging topic of facial image (FI) processing has been receiving considerable interest lately because of its real world applications such as forensic art, electronic consumer relationship management, security control and surveillance, entertainment and biometrics. In the FI context, age recognition (or estimation) has been demanding growing attention. Age synthesis, also called age progression is defined as re-rendering FIs with natural and rejuvenating effects. Age estimation (AE) can be defined as the process of associating a FI automatically with an exact age or age group. In order to facilitate AE, suitable facial representations are necessary. Otherwise, even the most robust classifiers will fail due to the inadequacy of the domain where the feature recognition is done. Hence, the design of face recognition systems requires careful selection of the Face Feature Recognition (FFR) domain.

Some issues that should be contemplated are: (i) good discrimination of different people with tolerance to discrepancies inside a class; (ii) FFR must be effortlessly performed from raw face images to speedup processing; and (iii) the FFR must lie in a low dimensional space, in order to facilitate the implementation of the classifiers.

The FI characteristics make the FFR problem very difficult to solve. The most important hindrances are: (1) AE is not a standard classification problem; (2) a large aging database, especially a chronometrical image series of an individual is often hard to collect; and (3) real world age progression displayed on faces is uncontrollable and personalized.

Several techniques have been suggested to represent FIs for recognition purposes, but there is still no consensus on the best when it comes to age recognition/classification. Appearance-based techniques consider an FI as a 2D array of pixels and focus on deriving descriptors for face appearance without precise geometrical representations. Holistic (nonparametric) methods such as the Principal Component Analysis (PCA) and the Linear Discriminate Analysis (LDA) along with more recent approaches like 2D-PCA and 2D-LDA have been broadly studied.

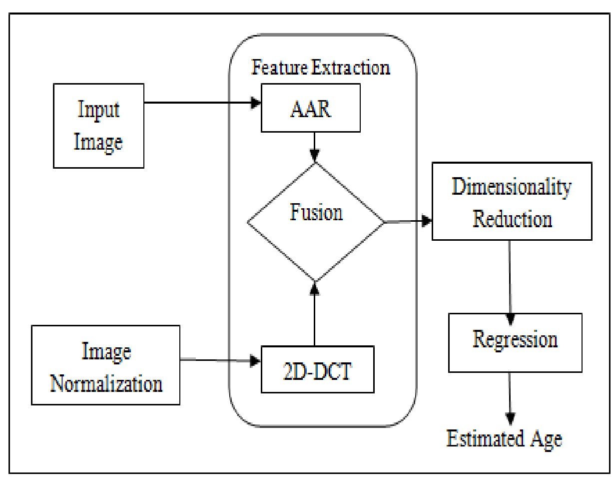


Fig:- System Architecture

In this paper, we propose a novel Global and Local feature based Age estimation (GLAAM) method. The

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018

input images are normalized and the local features are extracted using regional 2D-DCT. Global features are obtained with AAR. After feature extraction, dimensionality reduction is performed with PCA. Then, AE is cast as a regression problem. Our method uses global and local considerations and does not rely on a complex Bayesian framework; besides that, it is simple and relatively fast when compared to other ones.

2. Age Estimation Method

AAR based approaches consider both shape and texture rather than just the facial geometry as in the anthropometric model based methods. AAR uses a statistical model of object shape and appearance to synthesize a new image throughout a training stage which provides to the training supervisor a set of images and coordinates of landmarks existing in all of the images. AARs represent a familiar group of algorithms for fitting shape models to images. Training a model requires labeling a database of images where a set of locations called landmarks typify the object group in question.

The formulation chooses a linear and generative model, i.e. an explicit model of the input data has to be provided. This leads to an iterative type procedure, where the error between the current image features and those synthesized using the current location of the model in the image are used to derive additive updates to the shape model parameters. Nonetheless, the computational load is heavy, since an explicit image feature model must be stated and evaluated at each algorithm iteration.

Extended AARs for aging faces by proposing an aging function, age=f(b) which explains the variation in age. But they have to deal with each aging face image separately. Extracted feature vectors from images using AARs and used ensemble of classifiers trained on different dissimilarities to distinguish between child/teen-hood and adulthood. By using the different aging functions, accurate age of the classified image is estimated. An age estimation method using AAR features. Their approach is based on label sensitive learning and age-oriented regression.

3. The Global And Local Feature Based Age Estimation

This paper introduces an innovative AE method – known as GLAAM – relying on local and global facial features of images. Local features are extracted using regional 2D-DCT of normalized FIs and the global features are produced by AARs.

This method consists of the following modules: (I) face normalization, (II) global feature extraction with AAR and local feature extraction with 2D-DCT, (III) dimensionality reduction with PCA, and (IV) AE with multiple linear regression.

I). Face Normalization:-

Since shape and local variations of images during aging suffer an evident influence from rotation, scaling and translation, all the images have to be compatible with a common shape model produced by means of a training set of samples. In order to train the shape model, each image is represented by the coordinates of 68 landmark points. Then, the statistical shape model is trained and all images are warped to the mean shape, so that shape variations within the training set are eliminated. The warping process employs affine transformation and Delaunay triangulation.

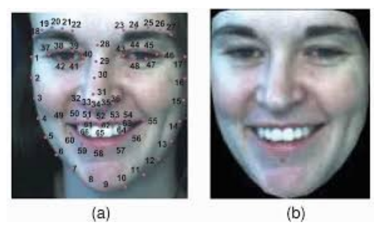


Fig: Example of face image labeled with 68 landmark points

So we rotate these images and cropped the main face part and scaled to the size 88×88 . Thus we almost eliminate unreasonable regions for feature extraction. In the local feature extraction phase the images are divided into blocks 8×8 block size. So we scale images to the size of each having 88×88 which is large enough for age related feature extraction .This image size is also efficient in terms of computational costs.

II. Feature Extraction:

The feature extraction module consists of two phases: global feature extraction with AAR and local feature extraction with 2D-DCT computation. These steps will be explained in detail in the following sections.

a) Global Feature Extraction with AAR:

AAR is a statistical shape and appearance model of FIs. These models are generated by combining a model of shape variations with a model of the appearance variations in a shape-normalized frame. A statistical shape model can be generated with a training set of face images labeled with landmark points. Let us represent all the landmark points of training images by $X = [x_i: x_i \varepsilon R^D]$. The mean shape is produced with taking the mean of the landmark points in the training set. Then, PCA is applied to the data to extract the main principal components along which the training set varies from the mean shape. The projection is chosen to maximize the determinant of the total scatter matrix of the projected samples. The set of eigenvectors of S corresponding to the d largest eigenvalues. Then a linear transformation maps the D-dimensional data space into a dimensional parameter space.

To build a statistical appearance model, each image has to be normalized, so that its control points match the mean shape using a Delaunay triangulation. Then, the gray-level intensities within a pre-specified image region are stacked to form vector g are used for training an intensity model. By applying PCA to the gray level intensities, a linear model is obtained as follows:

$$g = g + P_g b_g$$

where g is the mean gray-level vector, P_g is a set of orthogonal modes of variation and is a set of gray-level parameters. The shape and appearance of any image can be summarized by the vector b_g . Since there may be correlations between the shape and gray-level variations, a further PCA is applied to them and, finally, the integrated shape and appearance parameters are obtained. For the intensity model, approximately 7000 gray-level intensities in the facial region of the corresponding shape-free image are used to represent the training samples. The resulting integrated shape and intensity model requires 277 model parameters to explain 95 percent of the variance in the training set. These model parameters are used as a global descriptor of FI's.

b) Local Feature Extracting with 2D-DCT:

DCT is an invertible linear transform that can express a finite sequence of data points in terms of a sum of cosine functions. The original signal is converted to the frequency domain by applying the direct DCT transform and it is possible to convert back the transformed signal to the original domain by applying the inverse DCT transform (IDCT). After the original signal has been transformed, its DCT coefficients reflect the importance of the frequencies that are present in it. The 2D-DCT is commonly used as a pre-processing step in face recognition, because it attenuates the problems created by changes due to illumination angles, face occlusions, colors and pose. Using the face images directly for recognition purposes resulted in inefficiencies because of the high information redundancy and correlation in such images. Therefore DCT is widely used as a feature extraction and compression method in various applications due to its properties such as de-correlation, energy compaction, reparability and orthogonality. All these properties lead us to use 2D-DCT in AE field.

De-correlation: The principle advantage of image transformation is the removal of redundancy between neighboring pixels. This leads to un-correlated transform coefficients which can be encoded independently without compromising coding efficiency.

Energy-compaction:

Efficacy of a transformation scheme can be directly gauged by its ability to pack input data into as few coefficients as possible. This allows the quantize to discard coefficients with relatively small amplitudes without introducing visual distortion in the reconstructed image. DCT exhibits large variance distribution in a small number of coefficients for highly correlated images such as face images. In other words DCT packs energy in the low frequency regions. Therefore some of the high frequency content can be discarded without significant quality degradation.

Orthogonality:

In pattern recognition techniques to make the model computationally efficient, transform orthogonality is as important as the class separation in applications like face recognition. Unlike Gabor elementary functions, which are a set of overlapping functions and not mutually orthogonal, the DCT basis functions are orthogonal. In addition to its de-correlation characteristics, this property renders some reduction in the pre-computation complexity. In the proposed method, the normalized 88×88 images are divided into 11×11 blocks each having dimension of 8×divided into 11 2D-DCT is applied to them. This block size was adopted by the JPEG compression standard.

In the developmental phases the processing of larger blocks was seen as being prohibitively slow for the computer to execute. Also the experts observed that the use of larger blocks did not result in appreciably greater compression and quantization artifacts become more visible as the block size increases. In practice for a wide range of images and viewing 8×8 has been found to be the optimum DCT block conditions, 8×8 size and is specified in most current coding standards. After applying 2D-DCT we have 64 coefficients for each block.

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018

To eliminate the high frequency coefficients, quantization is performed. Coefficients are arranged in a vector following a zigzag fashion and the first 21 coefficients to represent that image block. Hence, the dimension of a local feature vector is 21×11×11=2541.

After the global and local features are extracted, they are integrated in a single vector in order to perform dimensionality reduction with PCA. To integrate the global and local features, a feature level fusion approach is used. For this purpose, the feature vectors are normalized by the z-score normalization as:

$$f_{i,j} = f_{i,j}-\mu_j / \sigma_j \text{ with } j=1,2 \text{ and } i=1,\ldots,n$$

where n is the number of images, $f_{i,j}$ is the j-th feature vector of i-th image and μ_j , σ_j are the mean and standard deviation of feature vector $f_{i,j}$ respectively. Then, the fused feature vector is created by concatenating the normalized global and local feature vectors

III. Dimensionality Reduction:

After the feature extraction module, PCA is performed in order to find a lower dimensional subspace which carries significant information for AE. Then, high-dimensional feature vectors are projected onto a low-dimensional subspace in order to improve the efficiency. Using this technique the p dimensional feature vector f is transformed into a d-dimensional vector y with .

The solution of this problem is given by the set of d eigenvectors associated to the d largest eigenvalues of the scatter matrix. Once the projection subspace is determined, training and testing images were projected on it. The low dimensional representation of feature vectors is calculated with allowing thus dimensionality reduction.

IV. Regression:

After finding a lower dimensional representation of facial images, we recast the AE problem as a multiple linear regression as follows:

$$Age = F(M) \leftrightarrow L = F'(Y)$$

Where L denotes the estimated age label , F(·) the unknown regression function , and F( ) is the estimated regression function.

The corresponding matrix formulation is $\tilde{L} = \tilde{Y} B + e$, $Var(e) = \sigma^2 I$

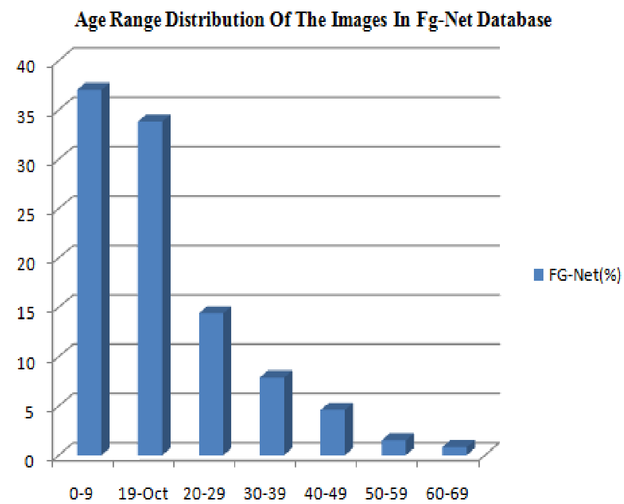
Where L is the age label vector. $\tilde{Y}$ is a known matrix including a column of 1s for the intercept and observed values. The vector B is the unknown parameter vector which we need to estimate during the learning stage. The error vector e consists of unobservable random variables, and assumed to have zero mean and uncorrelated with common variance.

4. Experiments and Results

In this paper, the FG-NET Aging Database is used to train and test the proposed method. This database contains 1,002 face images from 82 subjects with approximately 10 images per subject. The ages in the database are distributed in a wide range from 0 to 69. The age distribution of the FG-NET database. One can see from the table that the images are not distributed uniformly. A typical aging sequence from the FG-NET database. Besides the aging variation, most aging sequences display variations in pose, illumination, facial expression, occlusion, etc.

Age Range	FG-Net(%)
0-9	37.03
10-19	33.83
20-29	14.37
30-39	7.88
40-49	4.59
50-59	1.50
60-69	0.80

Table:- Age Range Distribution Of The Images In Fg-Net Database



Although these variations may increase computational complexity, all the images have been used in the experiments to avoid restrictions. The normalization phase determines the mean shape from the 68 landmarks of training samples. Next, all images are warped to the mean shape using affine transformation and Delaunay triangulation and scaled to the size of 88x88. Furthermore, each image is represented with 277 AAR model parameters that are used as global face features. In the local feature extraction step, the normalized 88x88 images are divided into 11x11 blocks each having 8x8 size and 2D-DCT is applied to them. After 2D-DCT computing, we have 64 coefficients for each block.

To eliminate the high frequency coefficients, quantization is performed. Coefficients are arranged in a vector according to a zigzag fashion. In this phase the determination of the number of DCT coefficients is done experimentally.

Then global and local feature vectors are normalized according to their mean and standard deviation and concatenated into a single vector. After that, a low dimensional age manifold is learned with PCA. AE is performed with multiple linear regressions in the low dimensional space.

As FG-NET contains face images from 82 subjects, after 82 folds, each subject has been used as test set once, and the final results are calculated based on all estimations. This scenery is very close to real life applications and, hence, it is very adequate for testing. The Mean Absolute Error (MAE) has been chosen as a metric for performance comparison.

Number of DCT Co-efficients	MAE
1	7,25
3	6,85
6	6,65
10	6,43
15	6,26
21	6,18
28	6,20
36	6,19

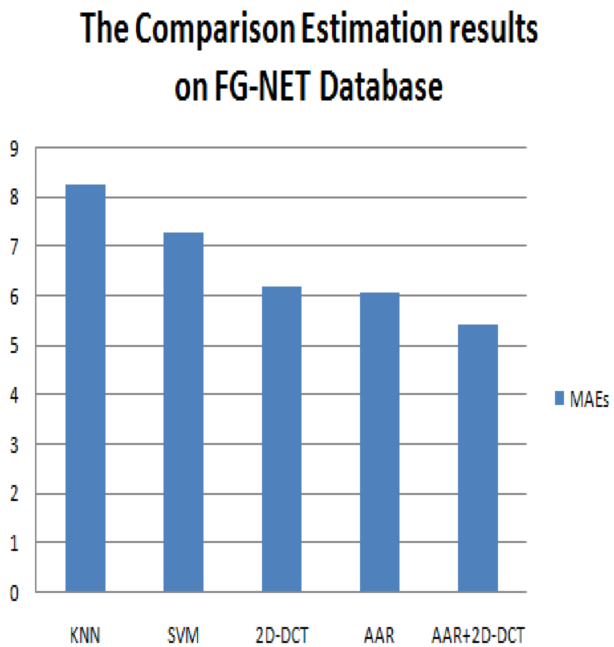
Table:- The Estimation Results Of Different Number Of DCT Co-efficients

GLAAM achieves better results than earlier methods like KNN, SVM, 2D-DCT, AAR and AAR+2D-DCT on the FG-NET database. DCT encodes facial texture and edge information in the frequency domain. Moreover local appearance information is captured using the block based DCT, but the global ones are ignored. So we use AAR, because it encodes the geometrical and global facial texture information in spatial domain. These feature sets capture differential complementary information. The combination of these feature vectors outperforms the AE accuracy of each one of the feature vector alone. We also investigate the AE performance of our method in various age ranges. we can observe that, GLAAM outperforms the AE accuracy of global features and local features alone, almost in all age ranges.

Methods	MAEs
KNN	8.24
SVM	7.25
2D-DCT	6.18
AAR	6.07
AAR+2D-DCT	5.39

Table:- The Comparison Estimation results on FG-NET Database

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018



5. Conclusion

In this paper, an AE method relying on an AAR model named GLAAM has been introduced. Its main contribution is a set of parameters accounting for both global texture features as well as local features of FI's. Locality is preserved by regional DCT coefficients and this is the main advantage/contribution of GLAAM over its competitors because DCT captures more accurately local features in FIs. Moreover, the proposed method is simple and relatively fast when compared to other ones used as benchmark, because 2D-DCT is recast and computed by means of 1D-DCT operations. Shape variations are eliminated via normalization with respect to mean shape obtained from a training set consisting of FIs. Moreover, the proposed method is simple and relatively fast when compared to other ones used as benchmark, because 2D-DCT is recast and computed by means of 1D-DCT operations. Shape variations are eliminated via normalization with respect to mean shape obtained from a training set consisting of FIs. Furthermore, these local features are integrated with global features of images. Experimental results using the FG-NET aging database show that GLAAM is better than earlier methods.

6. References

[1] A. Hadid, M. Pietikäinen and T. Ahonen, "A discriminative feature space for detecting and recognizing faces", Proc. of Computer Vision and Pattern Recognition, pp.797-804, 27 June-2 July 2004.

[2] W. Chen, M. J. Er and S. Wu, "PCA and LDA in DCT domain", Pattern Recognition Letters, vol. 26, no. 15, pp. 2474-2482, 2005.

[3] M. Li and B. Yuan, "2D-LDA: A statistical linear discriminant analysis for image matrix", Pattern Recognition Letters, vol. 26, no. 5, pp. 527-532, 2005.

[4] H. Yu and M. Bennamoun, "1D-PCA, 2D-PCA to nD-PCA", 18th International Conference on Pattern Recognition (ICPR'06), pp. 181- 184, 2006.

[5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[6] J.-M. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison", SIAM Journal on Imaging Sciences, vol. 2, no. 2, pp. 438-469, 2009.

[7] M. Hamouz, J. Kittler, J.-K. Kamarainen, P. Paalanen, H. Kalviainen and J. Matas, "Feature-based affine-invariant localization of faces", IEEE Transactions on Pattern Recognition and Machine Intelligence, vol. 27, no. 9, pp. 1490-1495, 2005.

[8] M. Bicego, A. Lagorio, E. Grosso and M. Tistarelli, "On the use of SIFT features for face authentication", Proc. of the Computer Vision and Pattern Recognition Workshop (CVPRW'06), pp. 17-22, June 2006.

[9] P. Paalanen, J-K. Kamarainen, J. Ilonen and H. Kälviäinen, "Feature representation and discrimination based on Gaussian mixture model probability densities-practices and algorithms", Pattern Recognition, vol. 39, no. 7, pp. 1346-1358, 2006.

[10] S. E. Choi, Y. J. Lee, S. J. Lee and K. R. Park, "Age estimation using a hierarchical classifier based on global and local facial features", Pattern Recognition, vol. 44, no. 6, pp. 1262-1281, June 2011.

[11] B. Ni, Z. Song and S. Yan, "Web image and video mining towards universal and robust age estimator", IEEE Transactions on Multimedia, vol. 13, no. 6, pp. 1217-1229, December 2011.

[12] Y. H. Kwon and N. V. Lobo, "Age classification from facial images", Computer Vision and Image Understanding, vol. 74, no. 1, pp. 1-21, April 1999.

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue-01
January 2018

[13] T. R. Alley, Social and Applied Aspects of Perceiving Faces, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.
[14] W.-B. Horng, C.-P. Lee and C.-W. Chen, "Classification of Age Groups Based on Facial Features", Tamkang Journal of Science and Engineering vol. 4, no.3, pp. 183-192, 2001.
[15] M. M. Dehshibi and A. Bastanfard, "A new algorithm for age recognition from facial images", Signal Processing, vol. 90, no.8, pp. 2431-2444, 2010.
[16] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework", International Journal of Computer Vision, vol. 56 , no. 3, pp. 221-255, 2004.
[17] T. Cootes, G. Edwards and C. Taylor, "Active appearance models", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 681-685, Jun 2001.
[18] A. Lanitis, C. Taylor and T. Cootes, "Toward automatic simulation of aging effects on face images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 442-455, April 2002.
[19] S. Kohli, S. Prakash and P. Gupta, "Hierarchical age estimation with dissimilarity-based classification", Neurocomputing, vol. 120 pp. 164- 176, 2013.