Automatic human age estimation system for face images Using Local Binary Pattern

Merga Tafa and Yusheng Zhu

DePaul University

Abstract – This paper examines the use of local binary pattern and for the purpose of age estimation based on face images. After preprocessing, local face regions were selected to extract skin texture features using LBP. These features were then fit in lasso regression models. Results showed good performance when compared with other methods.

Keywords -- Age estimation; Local binary pattern; regression

I INTRODUCTION

Human age estimation from facial image is the automated method of labeling the exact age of the individuals from their facial images using algorithms. Age is one of the base for communication and understanding as it determines habits, preferences and level of understanding during communication. It also the best way to identify certain biological changes such as skin wrinkles, hairs, etc. Age estimation provides information in fighting crimes as it save time and reduces costs.

II RELATED WORK

Human age estimation from facial images can be done in many methods. Those methods require model designing and the application of different algorithms to extract features from facial images. we have found out that some of previous works that related to the method we proposed and constructed either similar or different algorithms to extract features as well as in performing preprocessing and standardization as below.

[9]Aimed to build universal human age estimation from facial images. To proceed with plan, a multi-instance regression learning algorithm is proposed while several human face detectors of different implementation are used for robust face detection including principal component analysis to remove outliers.The process require noize removal during preprocessing and labeling of images. The datasets from FG_NET and MORPH are randomly partitioned to 600 images for training and 402 for testing. The performance recorded in MAE is 9.49 and 7.42 respectively.

[2] A novel hierarchical feature composition and selection model of facial estimation is proposed to proceed to the research process. The two-layer classifiers applied while features are extracted by using the boosting algorithms. For this methodology, FG-NET

databases of 1002 images for 82 individual images are used and the images are distributed in the range from 0 to 69 years of age. Moreover, MORPH database of 55,000 images for 13,000 individual images ages between 16-77 years from diverse races are also used for the study process.The feature selection is computed in four layers such as S1, C1, first boosting and second boosting. For the input images, the bounding box to detect five facial points such as eyes, centers and nose tips and two mouth corners applied and then the images were cropped. The boosting layers are applied to lower dimensional features from high-dimensional feature space and to combine the features. Some feature descriptors such as LBP (local Binary pattern) and SIFT (Scale-invariant transform) are used in this face recognition process. The performance result recorded for FG-NET database measurement is 3.76 in MAE and while 5.31 for MORPH.

[3]Automatic age estimation AGES (aging pattern) Subspace is proposed and the method uses each aging pattern as sample and modeling the aging pattern as representative of Subspace where aging is considered as face pattern that changes with time order. The aging pattern is formulated for images by projection of Subspace if their images are unseen to construct the face image.The algorithms that is used in this process is principal component analysis (PCA) to construct a Subspace that can capture certain variation in the dataset used for training and testing. As methodology, the appearance model is used to extract 200 features of victors. FG-NET and MORPH databases of 1002 images of 82 individuals and 1,724 face images of 515 individuals are used respectively. The age range estimation is applied to predict the significance for each range.The significance standard for age estimation is evaluated based on criterion and Mean Absolute error (MAE). This method (AGES) has FG-NET measurement in terms of MAE 6.77

[4] A hierarchical classifier facial features method based on local and global facial features is proposed. One type of methods used in feature extraction is local feature extraction. This feature extraction deals with external skin wrinkles, hairs. The other type of method is global feature extraction. This extraction deals with overall facial characteristics that is needed to estimate the detailed of age for all methods. This method is used Active appearance models(AAM) and Gabor Wavelet transform (GWT) hybrid features to combine both local and global features to bridge the gap.The experiment is set up based on FG-NET, BERC AND PAL aging database and fine cross validation performed in the process. The age gaps and gender for all images are evenly distributed for 390 individual images from age 3 to 83 years of age. For each age gap, the MAE computed for at 4 scales and 5 scales and recorded the result for MAE 4.657

[5]Kullback-leiber divergence is applied to embed the age difference information. This method exploits age information through age differences and apply the algorithms of loss functions such as entropy loss, cross entropy loss and K-L divergence distance to

force the probability distribution of age class to have a single peak value but also make the probability distribution locate within the correct range.This method applied preprocessing to over 150,000 images of 4000 celebrities and filtered those images using face detection algorithms. The next step is to normalize the size of images and train as object for Age-labeled image and non-age labeled images age as a variable using three loss functions and c-dimensional probability distribution to infer probability values.To obtain the result from the data, an open source Caffe framework model of deep neural network. It trains the age estimation model based on the difference of CNN architecture.The final result for the procedure is 2.8 for FG-NET and 2.78 for MORPH in MAE.

III METHOD

The data used for this analysis comes from FG-NET ageing database. The face and gesture recognition network (FG-NET) ageing database was released in 2004 in an attempt to support research activities aimed at understanding the changes in facial appearance caused by ageing. The FG-NET contains 1002 images from 82 different subjects with ages ranging between newborns to 69 years old subjects. However, ages up to 40 years are the most populated in the database. FG-NET images display considerable variability in resolution, image sharpness, illumination in combination with face viewpoint and expression variation. Occlusions in the form of spectacles, facial hair and hats are also present in a number of images. An example of one subject is shown in Figure 1.

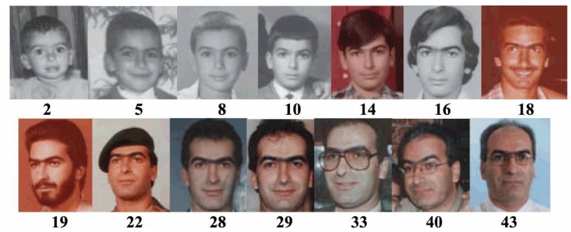


Figure 1: FG-NET Dataset Examples

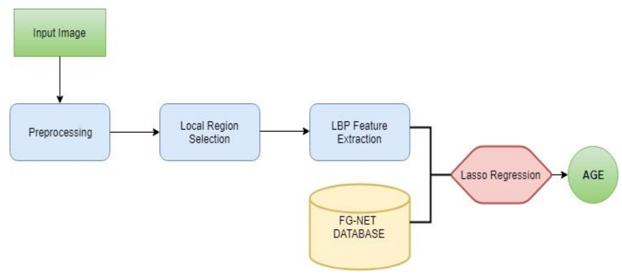


Figure 2: Age estimation flowchart

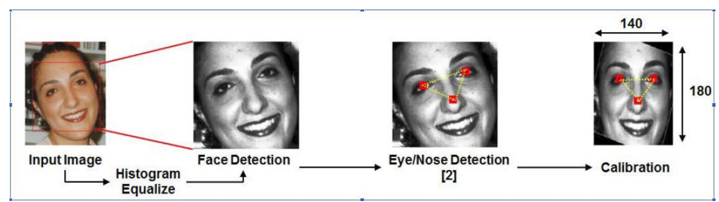


Figure 3: Pre-processing

The implementation of our proposed human age estimation system follows the flowchart as shown in Fig.2.

Pre-processing is the first stage to make the input image compatible with the system.Particularly certain local regions on the face images was selected for LBP feature extraction.Then we fit the features in a regression model to predict a exact age for test images.The rest of the section will be organized in three parts: A. pre-processing stage, B. feature extraction and C.estimation stage.

A. Pre-processing Stage

The pre-processing stage essentially transforms input face images of varying size to calibrated images ready for feature extraction.

Step 1: Face detection

Given an image containing a face to be estimated, the system first needs to detect face location and alleviate the effects of non-face regions.The Viola-Jones algorithm was used for facial detection which is based on four stages: Haar feature selection, creating an integral image, AdaBoost training, and cascading classifiers [1].

Then, all images were converted to grayscale. As noted previously, some images from the FG-NET dataset are in color. As this method is based on pixel intensities, it is important to have all images in grayscale. Next, histogram equalization is applied for contrast enhancement.

Step 2: Eye/nose detection

After face detection, we need to rotate the image so that face can be aligned in vertical direction. This can be done by first detect eyes' positions. The method we used here for eye detection is given in paper [4], where an enhanced pictorial structure model is used for precise eye localization. Details for this method can be found in the two papers. The resulting image after eye/nose detection is shown in Fig. 3 , where two eyes as well as nose are highlighted by red square boxes.

Step 3: Calibration

As we know the positions of two eyes, we are able to rotate the image. Assume the two eyes' positions are (x_1, y_1) and (x_2, y_2) , then the rotation angle θ can be determined by:

$$\theta = tan^{-1}(\frac{y_1 - y_2}{x_1 - x_2})$$

We also set a standard image size of 140x180 for all images sent to the regression model to make it easy for database training. Upon completion of image rotating, image resizing based on bilinear interpolation is applied to images that are not compatible with this standard. Image after the calibration stage is shown in Fig. 4(d). Note that calibrated eye/nose positions can be calculated from non-rotated image with equation below:

$$\begin{bmatrix} x_{rot} \\ y_{rot} \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix} \times \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} x_{cent} \\ y_{cen} \end{bmatrix} \right) + \begin{bmatrix} x_{rot.cent} \\ y_{rot.cent} \end{bmatrix}$$

After calibration stage, image is ready for feature extraction.

B. Feature extraction

Due to the fact the ageing process will change human's skin feature significantly by growing blemishes, we decide to use LBP to capture the skin texture. Before applying LBP the key problem is to select proper region of interest.

After further study, we decided to adapt the region selection method proposed in [5]. It has been proved that aging patterns are especially significant in the labeled regions.

To obtain region of interest,we first divide the entire face image into 8x8 cells. Since the eyes and nose positions are (x_1, y_1), (x_2, y_2) and (x_n, y_n) respectively, we can determine the cell positions of these key components. Assume left eye corresponds to cell $[x_{c1}, y_{c1}]$ right eye corresponds to cell $[x_{c2}, y_{c2}]$, and nose corresponds to cell $[x_{c3}, y_{c3}]$, we pick the following regions for skin feature extraction:

$$Region\ 1: [(x_{c1} - 2):(x_{c1} + 1), (y_{cn}: y_{c1})]$$

$$Region\ 2: [(x_{c2} - 1):(x_{c2} + 2), (y_{cn}: y_{c2})]$$

$$Region\ 3: [x_{cn}, (y_{cn}: y_{c1})]$$

Fig. 4 shows the region selection results of our implementation.

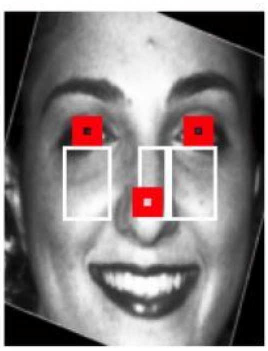


Figure 4: Example of region of interest

After region selection is completed, LBP feature extraction is applied to these regions.

For completeness,we briefly introduce LBP algorithms implemented in this age estimation system.

LBP operator is one of the best performing texture descriptors and used widely in texture classification, segmentation, and face detection/recognition. It has the advantage of computational simplicity while still offering good performance. As a person is getting older, facial blemishes such as freckles, age spots and fine wrinkles increase on the face skin. These micro-structures can be detected efficiently using LBP method.

The basic concept of LBP is to assign a code to each pixel comparing it to its neighbors. The creation of the LBP code is expressed by the equation below:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \text{, where } s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

where P is the number of neighboring pixels, R is the distance from the center to the neighboring pixels, g_c is the gray value of the center pixel, g_p (p = 1, …, P-1) are the gray values of the p equally spaced pixels on the circle of radius R that forms a circularly symmetric neighbor set, and s is the threshold function of x.The LBP feature vector can be created in the following manner:

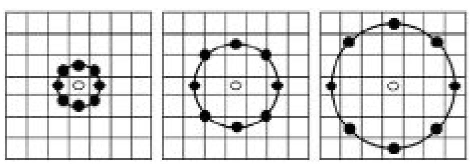
1.Divide the specified examined region into m×m cells (eg. 8×8 pixels for each cell).

2.For each pixel in a cell, compare the pixel to each of its n (eg. 8) neighbors. Follow the pixels along a circle. Examples are shown in Fig.5 for different combinations.

3.If the center pixel's value is greater than the neighbor's value, assign a "1". Otherwise, assign a "0". This gives an n digits binary number. An example of n = 8 is shown below in Fig. 6.

4.Compute the histogram, over the cell, of the frequency of each "number" occurring.

5.Concatenate histograms of all cells. This gives the feature vector for the examined region.



P = 8, R = 1 P = 8, R = 2 P = 8, R = 3

Fig. 5 Positions of neighboring according to P and R

Fig.6 An example of generating n-digit binary number (n = 8)

The resulting LBP vectors we obtained using Matlab from training images are 59 length long feature vectors.

C) Estimation stage

Similar study [10]on this problem using this dataset when doing classification has shown proportional relationship between age and LBP histogram mean. Therefore we assume that there is a linear relationship between age and LBP features. So we fit the features using LASSO regression to estimate an exact age from the model.

For the regression modeling, k-fold cross-validation was used rather than a training and testing set. As the datasets are not very large, a standard training and testing set was not appropriate 10 folds were selected k-fold cross-validation, as suggested by Hastie et. al in Elements of Statistical Modeling [2].

IV RESULTS

The output of an age estimation system can be an estimate of the exact age of a person or the age group of a person. For exact age estimation, the performance of an age estimator is usually based on the mean average error (MAE) between actual and estimated ages over a test set. The lower MAE indicates a better age estimation system.

Table 1 shows the accuracy results of our implementation using different scales of LBP and comparison with PCA method.

Table 1: MAE Comparison

feature extraction	Regression model	validation	MAE
PCA	Multi-instance regression	600/402	9.49
LBP(P=8, r=1)	LASSO Regression	10 fold cross validation	9.23
LBP(P=12, r=3)	LASSO Regression	10 fold cross validation	9.38

We can see that LBP has a slightly better prediction rate than the global method.And different scales of LBP capture the skin feature at a different level and may result different MAE.

V CONCLUSION

In this paper, the technique of local binary pattern for the problem of age estimation based on face images has been thoroughly studied and implemented. At region of interest, LBP provide good representations of ageing features and obtain a decent degree of accuracy for age estimation.Using LBP to build age estimation system has the advantage of computation efficiency.It is easy to implement and is robust to image invariants.However, the structural information captured by LBP is limited. Only pixel difference is used, magnitude information is ignored. We can conclude that our system is able to estimate age,and better results can be achieved by combining using gabor features etc.

ACKNOWLEDGEMENT

Thanks for Prof. Jacob D Furst for teaching this wonderful class.Thanks to Dr. Xiaoyang Tan for sharing ideas and implementation details on accurate eye/nose detection. Special thanks to Jonathan DeRuiter and Arthur Christoph for giving us project advice.

REFERENCES

[1] Viola, P. and Jone, M. "Rapid Object Detection using Boosted Cascade of Simple Features", Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference.

[2] Hastie, T., Tibshirani, R., and Friedman, J., "The Elements of Statistical Learning: Data Mining, Inference, and Prediction", New York: Springer, 2001

[3] X.Tan, S.Song, Z-H.Zhou and S.Chen. Enhanced Pictorial Structures for Precise Eye Localization under Uncontrolled Conditions, In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'09),pp. 1621–1628, Miami, Florida, USA, June 2009.

[4] X.Tan and B.Triggs. Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions, In: Proceedings of the 2007 IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'07), LNCS 4778, pp.168-182, 2007.

[5] Li, Y., Peng, Z., Liang, D., Chang, H. and Cai, Z. (2015). Facial age estimation by using stacked feature composition and selection. *The Visual Computer*, 32(12), pp.1525-1536.

[6],Geng, X., Zhou, Z. and Smith-Miles, K. (2007). Automatic Age Estimation Based on Facial Aging Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12), pp.2234-2240.

[7] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, J. Kim, "Age estimation using a hierarchical classifier based on global and local face features", Journal of Pattern Recognition, vol .44, no.6, p1262-1281, Feb 2011.

[8] Hu, Z., Wen, Y., Wang, J., Wang, M., Hong, R. and Yan, S. (2017). Facial Age Estimation With Age Difference. *IEEE Transactions on Image Processing*, 26(7), pp.3087-3097.

[9] Ni, Bingbing, et al. "Web Image Mining towards Universal Age Estimator." Proceedings of the Seventeen ACM International Conference on Multimedia - MM '09, 2009

[10] Tao, Y. (2014). Automated Estimation of Human Age, Gender and Expression. [online] Available at: https://web.stanford.edu/class/ee368/Project_Spring_1415/Reports/Tao.pdf [Accessed 16 Mar. 2018].