Automated Clustering and Estimation of Age Groups from Face Images using the Local Binary Pattern Operator

Imed Bouchrika, Ammar Ladjailia, Nouzha Harrati and Sofiane Khedairia
Faculty of Science and Technology
University of Souk Ahras
Souk Ahras, 41000 , Algeria
imed@imed.ws

***Abstract*—Automated age estimation from facial images has recently drawn a lot of attention from the research community emerging as a key technology with numerous applications ranging from access control to human machine interaction. In this research study, we explore a vision-based approach for the estimation of age groups from face images. The local binary pattern operator is applied to derive a set of hybrid features composed local and global characteristics from the face. A histogram of features is constructed based on the concatenation of locally produced histogram vectors from grid cells of face images. Hierarchical feature selection is described for the classification process where age ranges determined automatically in a tree-based fashion. Feature selection is based on the proximity of instances belonging to the same range is applied to obtain the most discriminative traits at each level of the defined age range. Experimental results carried out on a publicly available dataset confirmed the efficiency for the method to better cluster and estimate different age groups for different face images.**

***Keywords*-Age Estimation, Hierarchical Feature Selection, Local Binary Pattern**

I. Introduction

The human face conveys important amount of perceptible information and attributes such as expressions, ethnicity, gender and age. Automated age estimation from facial images has recently attracted remarkable interest from the research community emerging as a vital technology with numerous applications ranging from access control, human machine interaction, person identification and image-based data indexing and retrieval. Typical applications for the categories just mentioned include, age-restricted security control and surveillance monitoring [10], [6]. An age estimator application can generate a warning sound or alarm when a person who is below the legal age is entering bars or other restricted areas such as as casinos or even purchasing tobacco products from self-vending machines where face-based age estimation can be used as a primary check point. In the same way, knowing the age of the person automatically can have some merits in controlling access to the web through denying children from browsing internet pages with adult and inappropriate materials. Interestingly, automatic age estimation can be used within the area of human computer interaction to determine the age of a computer user in order to automatically adjust the user interface in order to suit the needs of their age group. For instance, an icon-based interface can be activated for young children whereas text with large font can be activated when dealing with elderly users.

There were a number of early studies in the literature concerning age progression through simulating the aging effects which is considered the inverse process of age estimation [1]. This includes the work of [2] where they artificially simulated aging variations through subjecting facial images to typical changes in shape and color. There are other research studies which are partly related to age estimation but are directed into exploring the mapping between face biometrics and age. Chellappa *et al* [3] described an approach for face verification across age based on a Bayesian classifier. The first study found in the literature which concerns age classification from facial images using image processing methods is published by Kwon *et al* [4]. They had extracted and used natural wrinkles for age classification of facial images into three main groups: baby, young adult, and senior adult. This is based on cranio facial changes in feature-position ratios in addition to skin wrinkle analysis.

The first major work was proposed by Lanitis [5] who claims to have the first attempt on automated age estimation. Statistical face models are generated by applying Principal Component Analysis on an a set of images which are subsequently used as the basis for obtaining a compact parametric description for facial images. Various classifiers are used to judge the performance of the proposed method such as the quadratic function, shortest distance classifier, neural networks and self-organizing map. For each age group, a different classifier is being employed based on another procedure to choose the most suitable classifier. Lanitis argued that the obtained results is an indication that machines can estimate the age of the person almost as reliable as humans. Choi *et al* [7] worked on the estimation of age from facial data using a hierarchical classifier based on support vector machine. Age features are constructed as a combination of global and local features derived using the local binary pattern operator together with Gabor filters to extract the wrinkle pattern. Ylioinas *et al* [8] have used a combination of local binary patterns for encoding the structure of the elongated facial patterns and their strength. The method was tested on images recorded in unconstrained conditions.

Because of the prime importance of age prediction in various applications ranging from human computer interaction to smart security applications, we describe in this research study a clustering approach for the estimation of age group from visual features of face images. The local binary pattern operator is applied to extract a hybrid set of features taken from local and global characteristics of the face. A histogram of features is constructed based on the concatenation of locally produced histogram vectors taken from grid cells. Hierarchical feature selection is described for the clustering process where age ranges are hierarchically estimated in a tree-based fashion. Feature selection is based on the proximity of instances belonging to the same class is applied to derive the most discriminative features at each level of the defined age range. Experimental results carried out on a publicly available dataset confirmed the ability of the proposed method to better estimate the age group for different face images.

II. Low-Level Feature Extraction

A. Local Binary Patterns

The Local Binary Pattern (LBP) operator was first introduced for texture analysis by Ojala *et al*. [9]. The LBP can be efficiently and swiftly computed in a single image scan offering facial recognition capabilities even for lower resolution images. The operator sets the pixels of a given image by thresholding each number of the neighboring pixels against the centre pixel within a 3x3 matrix and therefore, resulting a series of values of consecutive 1 or 0. By reading in the same direction of the arrow , a binary number is formulated which is converted to a decimal number i.e. a label where the binary number: 11010011 is converted to 211. The 256-bin histogram of the resulting labels is computed and employed as a texture descriptor for facial-based applications.

The main setback of the basic local binary operator is its small neighborhood area of (3*3) whereby it may ignore or disregard prominent features for larger structures. An extended version of the LBP operator is outlined in recent research studies by Ojala et al. [9] to use neighborhoods of different sizes. The extended LBP operator is represented by a circular neighborhood area written as (P, R) where P is the number of pixel points in the circular neighborhood whilst R is the radius of the circular area. The value of the LBP for the pixel point having the coordinate (x_c, y_c) is computed as shown in Equation (1):

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c)2^i \qquad (1)$$

Where g_i is the grayscale value of the pixel point i. c is the centre pixel. The function $s(a)$ is a thresholding function returning 1 for the case of $a \geq 0$ and 0 otherwise.

An $LBP(P, R)$ can produce 2^p different output values according to the 2^p different patterns formed by P which is the number of points in the chosen circular neighborhood. Research studies have shown that some patterns contain more discriminative information than the others. Further extension of the LBP has been introduced to take into account only uniform patterns which are defined as the patterns containing at most two bitwise transitions from 0 to 1 or vice versa, i.e.: the number of times that a digit alters from 0 to 1 or vice versa. As an example, the following binary numbers 00000000 and 00111000 and 11100001 are uniform patterns. An LBP operator written as (LBP^{U2_P}, R) meaning that this LBP operator is using the (P, R) circular neighborhood area with only uniform patterns being considered.

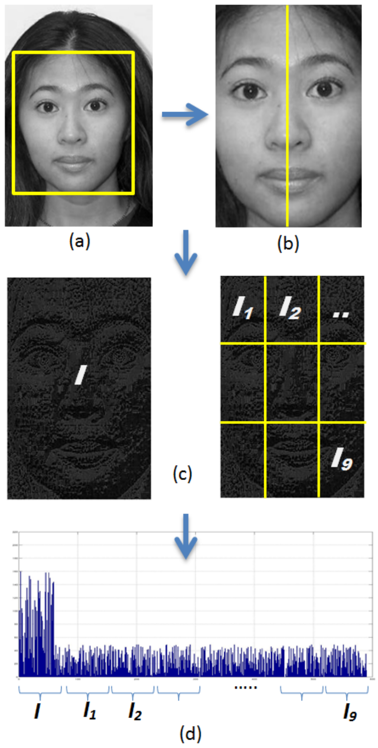

Figure 1. Face Detection & Feature Vector: (a) Face detection using Viola & Jones Algorithm. (b) Symmetry Analysis to tune face detection. (c) LBP Images (d) Concatenated Histogram of Features.

The histogram with an n-bin for the Local Binary Pattern operator is derived from a labeled image f_i as defined in Equation (2):

$$H_i = \sum_{x,y} b(f_i(x, y) == i) \qquad i = 0, 1, , , , n-1 \qquad (2)$$

Where b is a Boolean function returning 1 for true cases and 0 for false conditions. In [9], an improved spatially-based histogram is described which divides the image into m smaller regions R for the aim to retain spatial features where the histogram is computed as set in Equation (3):

$$H_{i,j} = \sum_{x,y} b(f_i(x, y) == i) b((x, y) \in R_j) \qquad (3)$$

where R_j is the j^{th} region of an image divided into an m region.

B. Facial Feature Selection

The feature vector for a facial age estimation is constructed initially via detecting the face from a single static image using the Viola and Jones from an algorithm. The implementation is provided within the computer vision

Matlab toolbox. To further refine and tune the detection accuracy, we apply a symmetrical analysis so that non-prominent features of the face such as hair are ignored. Figure (1) shows the steps being performed from the detection of the face to the formulation of the feature vector using a histogram construction based on the local binary operator discussed in the previous section. The technique of feature selection is considered in this study to extract the most distinctive features and remove the redundant and irrelevant facial components which may affect the estimation accuracy. The Adaptive Sequential Forward Floating Selection (ASFFS) search algorithm is deployed to reduce the number of features.

The feature subset selection procedure is purely dependent on an evaluation function that examines the distinctiveness of each feature or set of features in order to construct the best subset for the recognition or classification process. In this work, two validation criteria are employed. The first function is to approximate the the different clusters of age groups. The system implemented in this research uses a variation of the Bhattacharyya distance measure. The Bhattacharyya distance metric is a measure of the separation score $S_{i,j}$ between class i and j given by:

$$S_{i,j} = (m_i - m_j)\left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1}(m_i - m_j)^T \quad (4)$$

such that m_i and Σ_i are the mean and covariance of class i. For the case of N classes, the separation score is computed using the following:

$$J = \frac{1}{N^2}\sum_{a=1}^{N}\sum_{b=1}^{N} S_{a,b} \quad (5)$$

For the estimation of the age group for a given image based on the clustered groups, a validation-based evaluation criterion is proposed to choose the subset of features that would lesson the classification errors and ensure good inter-class separability between the different groups. As opposed to the voting paradigm used by the *KNN*, the evaluation criterion utilises coefficients w that reflect the importance of most nearest neighbours of the same class. The probability score for a candidate s_c to belong to a cluster c is expressed in the following Equation (6):

$$f(s_c) = \frac{\sum_{i=1}^{N_c-1} z_i w_i}{\sum_{i=1}^{N_c-1} w_i} \quad (6)$$

where N_c is the number of instances within cluster c, and the coefficient w_i for the i^{th} nearest instance is inversely related to proximity as given:

$$w_i = (N_c - i)^2 \quad (7)$$

The value of z_i is defined as:

$$z_i = \begin{cases} 1 & \text{if } nearest(s_c, i) \in c \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where the $nearest(s_c, i)$ function returns the i^{th} nearest instance to the instance s_c. The Euclidean distance metric is used to infer the nearest neighbours from the same class.

Instead of performing the feature selection process once for all final classes i.e. specific age groups, final classes are regrouped into higher classes in a hierarchical fashion producing an alike of a tree-based classification where the leaves are the final classes. Nodes correspond to higher level regrouped classes or age ranges. Feature selection is therefore applied recursively to generate at each level the appropriate subset of features. The regrouping of age ranges is performed incrementally at each level to compose two distinctive clusters of ages containing adjacent classes having higher separation. Figure (2) shows the hierarchical regrouping of classes deployed for age estimation where two higher classes of age groups are used as a start.

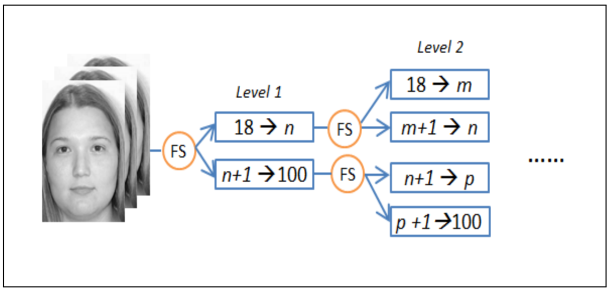

Figure 2. Hierarchical subset feature selection process for estimating age groups

III. Experimental Results

To investigate the potency of the proposed method to automatically predict the age range from facial features, experiments are carried out on the BW Kennedy database [11] provided by the University of Texas, Dallas. The dataset contains 180 face images with even distributions of both men and women. The dataset is constructed to include people with diversified age, sex and ethnicity. Face images are divided into 3 groups of 60 pictures each equally balanced on age group, sex, and ethnic diversity. Face pictures are all from the forward facing profile with neutral expression recorded at a resolution of 300×450 pixels. The dataset contains annotations conducted by 108 undergraduate student to collect data related to semantic attributes such as the perceived age of an individual from their face, perceived mood and memorability distinctiveness of the picture.

After running the feature selection procedure on the obtained raw features which contains 2,550 features, age signatures are constructed at each level of the binary tree after automated clustering of age groups. The Correct Classification Rate (CCR) is measured using the K-nearest neighbour (KNN) rule with $k = 3$ using the leave-one-out cross-validation procedure. Using the Cumulative Match Score (CMS) evaluation method proposed by Phillips during the FERET protocol, a poor correct classification rate (CCR) of 48.3% is achieved when using the normal feature selection procedure at once for all classes. Interestingly upon running the hierarchical classification process, a

Figure 3. BW Kennedy database for Age Estimation

high CCR of 96.1% is attained to differentiate successfully between the two age groups: 18-43 and 44-100. The overall correct classification rate for the hierarchical process reaches 87.8% for the rank of $R = 1$. Figure (4) shows the CMS curve for the classification process for the two cases. The achieved results are promising because the age estimation is based purely on texture-based information and this can be boosted through adding geometric properties of the face.

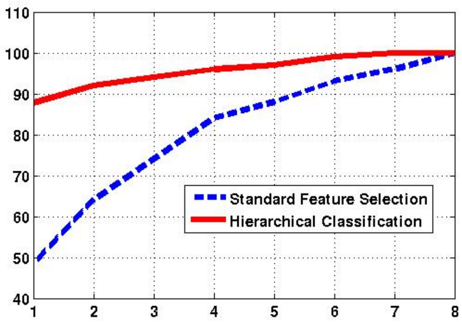

Figure 4. Cumulative Match Score for Age Estimation

IV. Conclusions

Automated age estimation from facial images has recently attracted considerable body of work from the computer vision community emerging as a key research topic with numerous applications ranging from access control to image-based data indexing and retrieval. We explore in this study a vision-based method for the clustering and estimation of age groups from facial features. The local binary pattern is applied to extract a hybrid set of features including local and global characteristics from the face. Hierarchical feature selection is described for the classification process where age ranges are grouped in a tree-based fashion. Experimental results carried out on a publicly available dataset confirmed the potentials for the proposed method to better estimate the age range for different face images.

Table I
Hierarchical Clustering and Estimation of Age Groups

<table>
<tr><th colspan="2">Level 1</th><th colspan="2">Level 2</th><th colspan="2">Level 3</th></tr>
<tr><td rowspan="4">18-43</td><td rowspan="8">96.1%</td><td rowspan="2">18-32</td><td rowspan="4">93.3%</td><td>18-24</td><td rowspan="2">91.5%</td></tr>
<tr><td>25-32</td></tr>
<tr><td rowspan="2">33-43</td><td>33-37</td><td rowspan="2">92.1%</td></tr>
<tr><td>38-43</td></tr>
<tr><td rowspan="4">44-100</td><td rowspan="2">44-56</td><td rowspan="4">97.2%</td><td>44-48</td><td rowspan="2">90.7%</td></tr>
<tr><td>49-56</td></tr>
<tr><td rowspan="2">57-100</td><td>57-72</td><td rowspan="2">94.6%</td></tr>
<tr><td>73-100</td></tr>
</table>

References

[1] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 12, pp. 2234–2240, 2007.

[2] D. M. Burt and D. I. Perrett, "Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 259, no. 1355, pp. 137–143, 1995.

[3] N. Ramanathan and R. Chellappa, "Face verification across age progression," *Image Processing, IEEE Transactions on*, vol. 15, no. 11, pp. 3349–3361, 2006.

[4] Y. H. Kwon and N. da Vitoria Lobo, "Age classification from facial images," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, Jun 1994, pp. 762–767.

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

[6] I. Bouchrika, "Gait analysis and recognition for automated visual surveillance," Ph.D. dissertation, University of Southampton, 2008.

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

[8] J. Ylioinas, A. Hadid, X. Hong, and M. Pietikäinen, "Age estimation using local binary pattern kernel density estimate," in *Image Analysis and Processing–ICIAP 2013*. Springer, 2013, pp. 141–150.

[9] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[10] I. Bouchrika, J. N. Carter, M. S. Nixon, R. Morzinger, and G. Thallinger, "Using gait features for improving walking people detection," in *20th International Conference on Pattern Recognition (ICPR)*. IEEE, 2010, pp. 3097–3100.

[11] K. M. Kennedy, K. Hope, and N. Raz, "Life span adult faces: Norms for age, familiarity, memorability, mood, and picture quality," *Experimental aging research*, vol. 35, no. 2, pp. 268–275, 2009.