Age Estimation from Facial Images based on Hierarchical Feature Selection

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

***Abstract*—The human face conveys remarkable amount of perceptible information and traits such as the emotional state, ethnicity, age and gender. In this paper, we explore a vision-based approach for the estimation of age range for an individual via facial features. The local binary pattern operator is applied to derive a hybrid set of features including 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. Hierarchical feature selection is described for the classification process where age ranges are classified in a tree-based fashion. Feature selection is based on the proximity of data instances belonging to the same age 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 potency for the proposed method to better estimate the age range for different face images.**

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

I. Introduction

Substantial amount of perceptible information and traits can be deduced from the human face including the emotional state, ethnicity, age and gender. This information serves a pivotal role in face-to-face interaction among people [1]. The majority of individuals are capable to determine with ease human traits such as the emotional state where they can infer when someone is sad, happy or angry [8] purely from the face Similarly, the gender of a whether male or female can be recognized spontaneously from facial traits. Contrarily, knowing the individual's age through looking or analyzing pictures is often a challenging task even for the human himself. Age estimation techniques by computer vision methods aim to approximate automatically the age range or exact age of an individual from their face [2]. Fu *et al* [3] listed different terminologies that describe the human age : *Actual age:* The real age of a person; *Perceived age:* This is the age predicted by a human based on the visual appearance of an individual; *Estimated age*: the age approximated by machine through automated methods.

Automated age estimation from facial data has recently gained a lot of attention from the research community emerging as a key technology with numerous applications ranging from human machine interaction, access control, biometrics and image-based data indexing and retrieval. Typical scenarios for the categories just mentioned include, age-restricted security control and surveillance monitoring. 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 buying 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, automated estimation of age can be deployed within the area of human computer interaction to determine the age of the user in order to automatically re-customize or adjust the graphical interface to meet the needs and requirements of user based on their age group [11]. For instance, an icon or image-oriented interface can be enabled for younger users meanwhile text with large font can be displayed for elderly users.

Age estimation systems are broadly designed to consists of two phases: Feature extraction and classification. The first step is central to the success of the classification stage as the extracted features affect largely the performance of the age prediction process [2]. Given a face image of an individual, the age range is estimated based on a set of low-level features whether automatically extracted or manually annotated. However, automated marker-less estimation of age from static images is proven to be a cumbersome process. The difficulties stem from a number of challenging factors related either to the individual or acquisition environment. The aging progress is uncontrollable and irreversible such that different people undergo different aging rate which is steered not only by the human gene, but also many other external factors that influence the aging process including health conditions, life style and working environment [4], [5], [6]. Challenges related to the acquisition environment may include background clutter, illumination, camera movement and viewpoint as well as occlusion.

Because of the pivotal importance of age prediction in various applications ranging from human computer interaction to smart security applications, we describe in this research study a vision-based approach for the estimation of age range from a face image. The local binary pattern is applied to extract a hybrid set of features including local and global characteristics from 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 classification process where age ranges are grouped in a tree-based fashion. Feature selection is based on the proximity and nearness of data instances belonging to similar classes 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 potentials for the proposed method to better estimate the age range for different face images.

This research paper is structured as follows. The following section outlines the previous methods for automated age estimation. The texture-based method for age prediction is detailed in Section 3. Subsequently, the experimental results are presented followed by conclusions.

II. Related Work

Compared with the considerable body of studies on face-related research area as biometrics and face synthesis, there are relatively few publications devoted for automated age estimation. Choi *et al* [2] surveyed the different types of features used for predicting the age from facial images grouping them into two major categories; local and global features. For the first type of features, they are taken from the depth or amount of wrinkles, facial hair or skin in addition to the geometric properties of facial elements. The local characteristics are commonly known to better classify people into age groups as they embed particular characteristics that discriminate different age groups. As opposed to the local features, Choi [2] argued that the global counterparts are better for estimating more accurate age information and contain not only the age attributes, but more further individual-related traits as identity, emotion and ethnic background. Hybrid features which are produced through combining local and global features are found to offer superior performance for various face-related applications. This is because of the inefficiency found in each type of feature can be compensated leading to the conclusion that hybrid features are desirable for accurate age estimation.

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 [7]. This includes the work of [9], [10] 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* [12] described an approach for face biomerics 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* [14]. They had extracted and used natural wrinkles for age classification of facial images into three main groups: baby, young adult, and senior adult.

The first major work was proposed by Lanitis [1] who claims to have the first attempt on automated age estimation. Statistical models of the face are constructed through applying Principal Component Analysis on an a set of images which are utilized subsequently as the basis for generating a compact parametric representation of the face. Various classifiers are considered to judge the performance rates of the presented 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* [2] worked on the estimation of age from facial data in a hierarchical way using the support vector machine classifier. Age features are constructed as a combination of local and global features derived using the local binary pattern operator together with Gabor filters to extract the wrinkle pattern. Ylioinas *et al* [15] have used a combination of local binary patterns for representing the structure of the facial patterns with their strength. The method was tested on images recorded in unconstrained conditions.

III. Low-Level Feature Extraction

A. Local Binary Patterns

The Local Binary Pattern (LBP) operator was first introduced for texture analysis by Ojala et al. [17] in 1996 in their editorial “a comparative study of texture measures with classification based on featured distribution”. 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 as shown in Figure (1). By reading in the same direction of the arrow shown in Figure (1), 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.

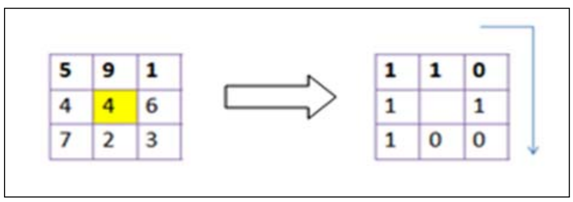

Figure 1. The basic Local Binary Pattern Operator (LBP)

The main drawback 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.* [17] to utilize neighborhood areas of varying sizes. The extended LBP operator is represented by a circular neighborhood area written as (P, R) where P refers to the number of pixels in the circular neighborhood whilst R is the radius of the circular area as shown in Figure (2). The value of the LBP

for pixel point having the coordinate (x_c, y_c) is computed as shown in Equation (1):

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

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.

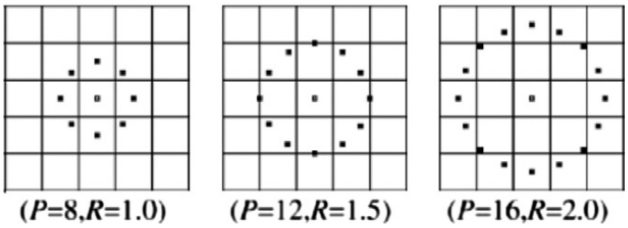

Figure 2. The Extended Local Binary Pattern Operator [18].

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 conveys more distinctive information than 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 transitions from 0 to 1 or vice versa, i.e.: the number of times that a digit alters between 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.

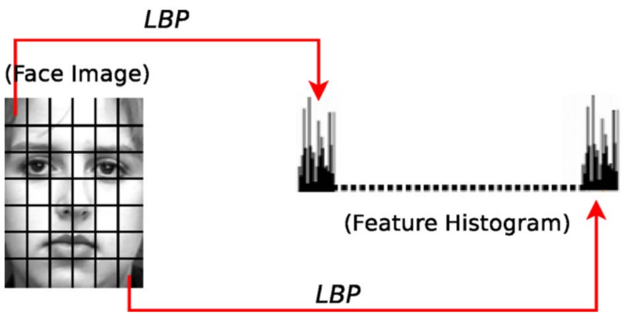

Figure 3. The Local Binary Pattern Histogram [18]

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

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

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 as shown in Figure (3) where the histogram is computed as set in Equation (3):

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

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 the detection of the face from a single static image using the Viola and Jones method. 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 filter out the redundant and irrelevant facial information which may affect poorly the estimation accuracy. The Adaptive Sequential Forward Floating Selection procedure is deployed in order to reduce the original feature space. The feature selection method has been used successfully in biometrics and surveillance applications [13], [19], [16].

The feature subset selection procedure is dependent on an evaluation criterion or objective function that examines the distinctiveness of a feature dimension or group of features to derive the ideal subset for the classification task. 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 derive the feature subset that would reduce the classification errors as well guarantee higher inter-class separability across the different groups. In contrast to the voting scheme utilized by the *KNN* classifier, the objective function employs coefficients w that reflect the importance of nearest neighbours belonging to the same class. The score value for an instance s_c

to belong to an age cluster c is given 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:

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

Such that the $nearest(s_c, i)$ function returns back the i^{th} nearest sample to the instance s_c. The Euclidean distance is computed 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 (4) shows the hierarchical regrouping of classes deployed for age estimation where two higher classes of age groups are used as a start.

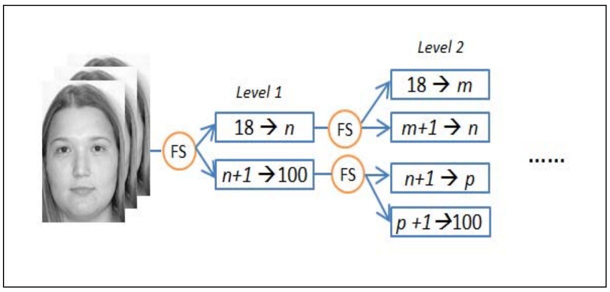

Figure 4. Hierarchical subset feature selection process for estimating age groups

IV. Experimental Results

To investigate the potency of the proposed method to automatically predict the age range from facial features, experiments are conducted using the BW Kennedy database [20] 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 images each equally balanced on sex, age group and ethnicity. 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.

Figure 5. BW Kennedy database for Age Estimation

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 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 (6) 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.

V. 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

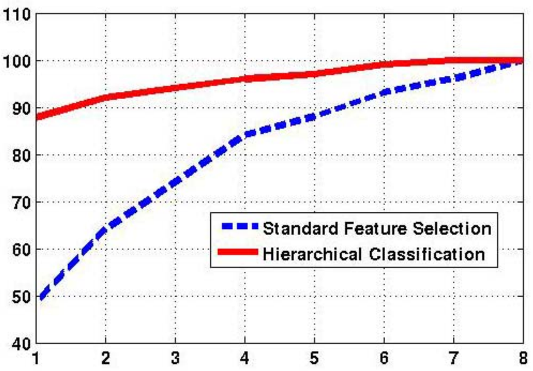

Figure 6. Cumulative Match Score for Age Estimation

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>

a publicly available dataset confirmed the potentials for the proposed method to better estimate the age range for different face images.

REFERENCES

[1] 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.

[2] 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.

[3] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 11, pp. 1955–1976, 2010.

[4] D. S. Berry and L. Z. McArthur, "Perceiving character in faces: the impact of age-related craniofacial changes on social perception." *Psychological bulletin*, vol. 100, no. 1, p. 3, 1986.

[5] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang, "Image-based human age estimation by manifold learning and locally adjusted robust regression," *Image Processing, IEEE Transactions on*, vol. 17, no. 7, pp. 1178–1188, 2008.

[6] G. Guo, G. Mu, Y. Fu, and T. S. Huang, "Human age estimation using bio-inspired features," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 112–119.

[7] 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.

[8] N. Harrati, I. Bouchrika, A. Tari, and A. Ladjailia, "Automating the Automated classification of facial expressions using bag of visual words and texture-based features," in *16 International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, 2015.

[9] 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.

[10] A. J. O'toole, T. Vetter, H. Volz, E. M. Salter *et al.*, "Three-dimensional caricatures of human heads: distinctiveness and the perception of facial age," *Perception*, vol. 26, pp. 719–732, 1997.

[11] N. Harrati, I. Bouchrika, A. Tari, and A. Ladjailia, "Automating the evaluation of usability remotely for web applications via a model-based approach," in *International Conference on New Technologies of Information and Communication (NTIC)*, 2015.

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

[13] I. Bouchrika, J. N. Carter, and M. S. Nixon, "Towards automated visual surveillance using gait for identity recognition and tracking across multiple non-intersecting cameras," *Multimedia Tools and Applications*, pp. 1–21, 2014.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009.

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

[20] 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.