 ISBN 978-3-200-02328-4

GENDER RECOGNITION FROM FACE IMAGES WITH LOCAL WLD DESCRIPTOR

Ihsan Ullah [1], *Muhammad Hussain* [1,a], *Ghulam Muhammad* [2], *Hatim Aboalsamh* [1], *George Bebis* [1,3] and *Anwar M. Mirza* [2]
[1]Department of Computer Science,[2]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[3]Department of Computer Science and Engineering, University of Nevada at Reno
[a]Email: mhussain@ksu.edu.sa

ABSTRACT

In various biometric applications, gender recognition from facial images plays an important role. In this paper, we investigate Weber's Local Descriptor (WLD) for gender recognition. WLD is a texture descriptor that performs better than other similar descriptors but it is holistic due to its very construction. We extend it by introducing local spatial information; divide an image into a number of blocks, calculate WLD descriptor for each block and concatenate them. This spatial WLD descriptor has better discriminatory power. Spatial WLD descriptor has three parameters. Through a large number of experiments performed on FERET database, we report the best combination of these parameters and that our proposed spatial WLD descriptor with simplest classifier gives much better accuracy i.e. 99.08% with lesser algorithmic complexity than state-of-the-art gender recognition approaches.

***Index Terms*—** Gender recognition, Weber's Law Descriptor, Face Recognition, Local descriptors

1. INTRODUCTION

Gender classification is an important task which in turn can enhance the performance of a wide range of applications including identity authentication, human-computer interaction, access control, and surveillance, involving frontal facial images. A large majority of gender classification approaches are based on extracting features from face images and then giving these features to a binary classifier. The feature extraction phase has been carried out by using either (a) appearance based methods or (b) geometric methods. In appearance based methods, the whole image is considered rather than the local features corresponding to different parts of the face. While, in geometric based methods, the geometric features like distance between eyes, face length and width, etc., are considered. For classification purposes, mostly neural networks, nearest neighbor method, linear discriminant analysis, and other binary classifiers are used.

Artificial Neural Networks (ANNs) [1-2] and the method of Principal Component Analysis (PCA) [3] were initially studied for gender classification. Gutta et. al. [4] proposed a hybrid gender classifier consisting of an ensemble of Radial Basis Functions and C4.5 decision trees. They used 3006 images of 1009 subjects using cross validation on manually segmented and normalized images of size 64 x 72 pixels from FERET database [5], reporting an accuracy of 96%. Moghaddam et. al. [6] proposed to classify gender from facial images (of 21 x 21 pixels) using SVMs and reported recognition rate of 97% on the color FERET database. Nakano et. al. [7] computed the edge information and exploited a neural network classifier for gender recognition. Lu et. al. [8] exploited the range and intensity information of human faces for ethnicity and gender identification using a Support Vector Machine (SVM). Kim et. al. [9] based their gender recognition system on a Gaussian Process Classifier. Yang et. al. [10] improved gender classification using texture normalization. Baluja and Rowley [11] combined several weak classifiers based on pixel value comparisons on low resolution gray scale images in their AdaBoost based gender classifier. Tests carried out on the 20 x 20 normalized images from FERET database showed an overall accuracy of 90%. Lu and Shi [12] employed the fusion of left eye, upper face region and nose in their gender classification approach. Their results showed that their fusion of face region approach out performs the whole face approach. Extending this idea, Alexandre [13] used a fusion approach based on features from multiple scales. They worked on normalized resized images (20 x 20, 36 x 36 and 128 x 128) to extract shape and texture features. For texture features, they used Local Binary Pattern [14] approach for the whole image.

In this paper we introduce a novel technique for enhancing the gender classification rate using the textural properties of the faces. The idea of using textural properties of faces is not new, however we employ a new texture descriptor WLD (Weber Local Descriptor) [15], which has never been tested for gender recognition. Chen et. al. [15] have demonstrated that WLD outperforms in texture recognition than stat-of-the-art best descriptors like LBP, Gabor, and SIFT. The basic WLD descriptor is a histogram where differential excitation values are integrated according their gradient orientations. The differential excitation values are concatenated irrespective of their spatial location and so WLD behaves like a holistic descriptor. We extend it to enhance its discriminatory power by embedding the local spatial information and call it Spatial WLD (SWLD) and will refer to the basic WLD as holistic WLD (HWLD) in our onward discussion. The paper mainly contributes in: (i) Exploiting WLD as local feature extractor (ii) finding the best combination of parameters and (iii) finding optimal gender recognition rate.

The rest of the paper is organized as follows. Section 2 presents an overview of WLD descriptor. Gender recognition system is discussed in Section 3. Section 4 presents experimental results. Section 5 concludes the paper.

2. WEBER'S LAW DESCRIPTOR FOR IMAGE REPRESENTATION

In this section we give an overview of basic WLD descriptor [15] and its extension. This descriptor represents an image as a histogram of differential excitations and gradient orientations, and has several interesting properties like robustness to noise and illumination changes, elegant detection of edges and powerful image representation.

WLD descriptor is based on Weber's Law. According to this law the ratio of the increment threshold to the background intensity is constant. Inspired by this law, Chen et.al [15] proposed WLD descriptor for texture representation. The computation of WLD descriptor involves three steps i.e. finding differential excitations, gradient orientations and building the histogram.

2.1. Differential Excitation

For calculating differential excitation $\mathcal{E}(x_c)$ of a pixel x_c first intensity differences of x_c with its neighbors x_i, i = 1, 2, …, p are calculated as follows:

$$\Delta I_i = I_i - I_c \,. \quad (1)$$

Then the ratio of total intensity difference of x_c with its neighbors x_i to the intensity of x_c is determined as follows:

$$f_{ratio} = \sum_{i=0}^{P-1} \left(\frac{\Delta I_i}{I_c}\right). \quad (2)$$

Arctangent function is used as a filter on *Eq* (2) to enhance the robustness of WLD against noise which results in:

$$\varepsilon(x_c) = \arctan\left[\sum_{i=0}^{P-1} \left(\frac{\Delta I_i}{I_c}\right)\right]. \quad (3)$$

The differential excitation $\varepsilon(x_c)$ may be positive or negative. The positive value indicates that the current pixel is darker than its surroundings and negative value means that the current pixel is lighter than the surroundings.

2.2. Gradient Orientation

Next main component of WLD is gradient orientation. For a pixel x_c the gradient orientation is calculated as follows:

$$\theta(x_c) = \arctan\left[\frac{I_{73}}{I_{51}}\right] \quad (4)$$

where $I_{73} = I_7 - I_3$ is the intensity difference of two pixels on the left and right of the current pixel x_c, and $I_{51} = I_5 - I_1$ is the intensity difference of two pixels directly below and above the current pixel, $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

The gradient orientations are quantized into T dominant orientations as:

$$\phi_t = \frac{2t}{T}\pi \text{ where } t = mod\left(\left\lfloor\frac{\theta'}{2\pi/T} + \frac{1}{2}\right\rfloor, T\right) \quad (5)$$

where $\theta' \in [0, 2\pi]$ and is defined in terms of gradient orientation computed by Eq. (4).

In case $T = 8$, the dominant orientations are $\phi_t = \frac{t\pi}{4}$, $t = 0,1, \ldots, T-1$; all orientations located in the interval $\left[\phi_t - \left(\frac{t\pi}{4}\right), \phi_t + \left(\frac{t\pi}{4}\right)\right]$ are quantized as ϕ_t.

2.3. Holistic WLD Descriptor

After calculating differential excitation and dominant orientation, WLD descriptor is build. Corresponding to each dominant orientation ϕ_t: t = 0, 1, 2, …, T-1 differential excitations are organized as a histogram H_t. Then each histogram H_t: t = 0, 1, 2, …, T-1 is evenly divided into M subhistograms $H_{m,t}$: m = 0, 1, 2, …, M-1, each with S bins. These histograms form a histogram matrix, where each column corresponds to a dominant direction ϕ_t. Each row of this matrix is concatenated as a histogram H_m = { $H_{m,t}$: t = 0, 1, 2, …, T-1}. Subsequently, histograms H_m: m = 0, 1, 2, …, M-1 are concatenated into a histogram H = { H_m: m = 0, 1, 2, …, M-1}. This histogram is referred to as WLD descriptor. This descriptor involves three free parameters: T, the number of dominant orientations, M the number of segments of each histogram corresponding to a dominant orientation and S, the number of bins in each segment.

2.4. Spatial WLD Descriptor

The basic WLD descriptor described in previous subsections represents an image as a histogram of differential excitations organized according to dominant gradient orientations. In this histogram differential excitations are collected according to their values and gradient orientations irrespective of their spatial location. Spatial location is also an important factor for better description. For example, two different regions in an image with similar differential excitations and gradient orientations will contribute to the same bins in the histogram, and will not be discriminated by WLD descriptor. To enhance the discriminatory power of WLD descriptor, we introduce spatial information into the descriptor. We divide each image into a number of blocks, compute WLD histogram for each block and concatenate them to form a Spatial WLD descriptor (SWLD). SWLD involves four parameters: T, M, S and the number blocks.

This performs better because it captures the local information in a better way, which is important for recognition purpose. But this approach introduces another parameter: the size of blocks. The optimal value of this parameter can lead to better recognition results.

3. GENDER RECOGNITION

The block diagram of the recognition system which we used in gender recognition is shown in Figure.1. The two main components of the system are feature extraction and classification. Various existing approaches differ in the choice of feature extraction and classification. For feature extraction we used spatial WLD descriptor described in previous section. The need to extract most discriminative features for gender recognition motivated us to use SWLD.

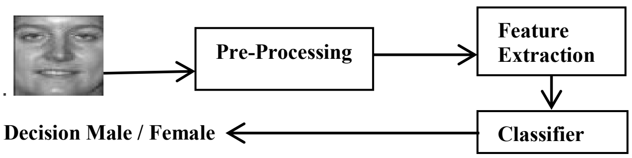

Figure.1. Gender recognition system

3.1. Classifier

In the literature most of the authors used highly sophisticated classifiers like SVM and Ada-boost but to keep the system simple, we preferred to employ minimum distance classifier for getting optimal results for gender classification. SWLD can give better or comparable result to many of stat of the art techniques using city block distance (L1), Euclidean Distance (L2), and Chi-Square (CS). The accuracy of a gender recognition system depends on choice of selecting suitable metric with each feature extractor.

3.2. Facial Datasets

For experiments, we used two dataset with low (20x16) and high (60x48) resolutions from FERET database [5], which is one of the challenging databases for face recognition. This database contains image corpus that is collected to evaluate the algorithms of face recognition by standardized procedures and test. The images in the database are frontal, left or right profiles and could have some variations in pose, expression and lightning. For experiments, we used two sets: *fa(training)* and *fb(testing)*. The set *fa*, that is usually used as training set, contains 1204 (746+458) images of 403 male subjects and 403 female subjects. The set *fb*, contains 1196 (740 male + 456 female) images which were taken seconds after the corresponding *fa* images but with different face expression, illumination and pose.

4. EXPERIMENTS AND DISSCUSION

In our experiments we tested various combinations of three parameters (T, M, S). These combinations were applied on different block size images. We performed experiments with T = 4, 6, 8, M = 4, 6, and S = 4, 8 to find optimal results. Block Sizes of 20 x 12, 15 x 12, 12 x 12, and 6 x 12 were used for high resolution images. Figure.4 shows the effect of block-sizes and WLD parameters. The result obtained using holistic WLD is not promising and is less than 90%. Reducing the block size up to some extent improves recognition rate. With block size of 15 x 12 overall result reached to 98.83% with metrics L1 and CS. At block-size of 12 x 12, our system generate maximum overall result of 99.08% with CS metric and T, M and S values of 8, 4 and 4 respectively which in not only bester than holistic WLD but also other state-of-the-art techniques results for gender recognition up to date.

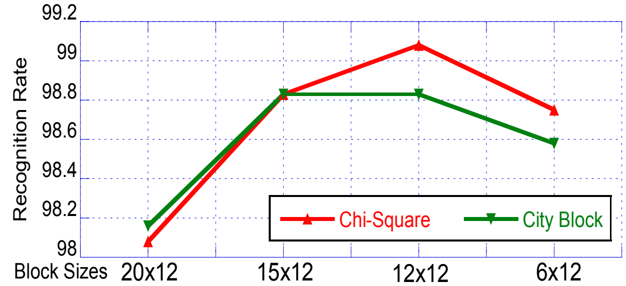

Figure.2. Comparison of Chi-square and City Block

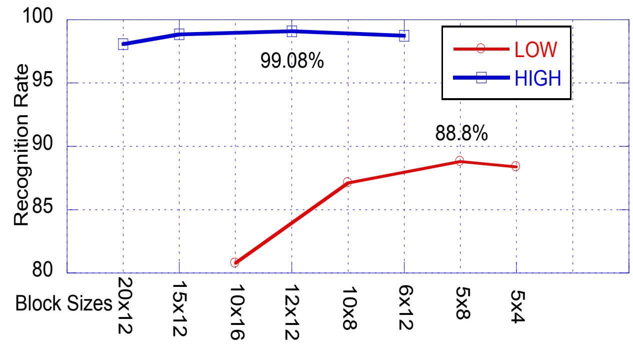

Figure.3. Effect of resolution on recognition rate

In case of low resolution database we used block size of 10 x 16, 10 x 8, 5 x 16, 5 x 8, 5 x 4, 4 x 16, and 4 x 8with same parameters for T, M and S as for high resolution datasets. The overall results are very poor with all combinations of parameters and block sizes with a maximum of 88.80% with block size of 5 x 4 and T, M, and S values of 8, 6, and 4 respectively, see Figure.3. This indicates that the resolution of images also performs key role in achieving best result for gender recognition. Also, note that T = 8 give better results in both cases: low and high resolution.

We checked the recognition rate with four different distance measures: City-block distance (L1), Euclidean distance (L2), Cosine (COS), and Chi-Square distance (CS). CS is on the top and excels L1 with fractional difference and gives the best accuracy as is clear in Figure.2.

In Figure.3 the effect of different block sizes on low and high resolution images with CS metric is shown. A huge block size does not show good result while reducing the block size up to some extent improves the recognition rate which starts decreasing after that. Figure.3 shows that the maximum recognition rate for low resolution database is given by block size 5 x 4 with T, M, and S values of 8, 6, and 4 using CS metric. On high resolution images optimal result is given on 12 x 12 block size and using CS metric with T, M, and S as 8, 4, and 4 respectively.

At the end we compared proposed spatial WLD results with state-of-the-art best techniques: Multi-resolution Decision Fusion method (*MDF*) [13], Local Gabor Binary Pattern with LDA and SVMAC method (*LGBP-LDA SVMAC*) [16] and Local Gabor Binary Pattern with LDA and SVM method *(LGBP-LDA SVM*) [16]. Also we compared it with holistic WLD and PCA. Figure.5 shows that despite being much simpler than other three methods, Spatial WLD gives comparable recognition rate and is much better than holistic WLD and PCA.

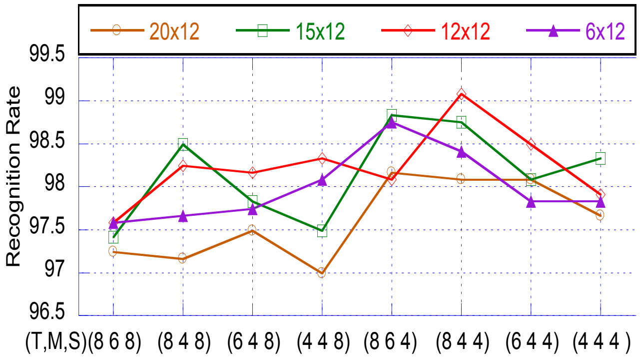

Figure.4. Effect of different combinations of T, M, S and Block Sizes on recognition accuracy.

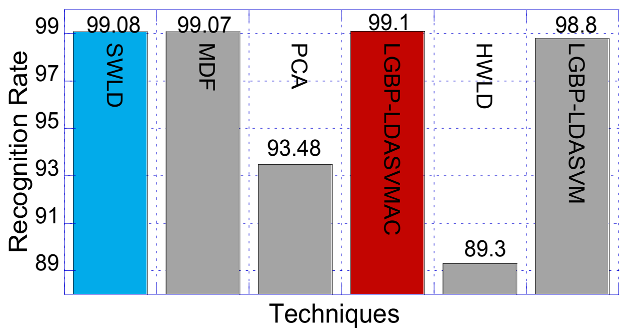

Figure.5. Comparison with Other Methods

5. CONCLUSION

WLD as a local descriptor results in much improvement in recognition accuracy for gender recognition problem. The best result is obtained with block size of 12 x 12 and T, M, and S values of 8, 4, and 4 respectively while using simple classifier with Chi-square distance. Despite its simplicity, the proposed system can produce as good results as complicated systems. In our future work we will check for two things. We will explore WLD with sophisticated classifiers like SVM. As we decrease block size, SWLD feature histogram dimensionality increases, which increase time complexity. This can be reduced by feature selection techniques. We will investigate this also as a future work.

Acknowledgement

This work is supported by the National Plan for Science and Technology, King Saud University, Riyadh, Saudi Arabia under project number 10-INF1044-02.

REFERENCES

[1] A. Golom, D. T. Lawrence, and T. J. Sejnowski, “SEXNET: A neural network identifies gender from human faces”, Advances in Neural Information Processing Systems, pp. 572-577, 1991

[2] B. Edelman, D. Valentin, and H. Adbi, “Sex classification of face areas: how well can a linear neural network predict human performance”, Journal of Biological System, Vol. 6(3), pp. 241-264, 1998.

[3] Z. Sun, G, Bebis, X. Yuan and S. J. Louis, “Genetic feature subset selection for gender classification: a comparison study”, Proc. IEEE Conference on Applications of Computer Vision, pp. 165-170, 2009.

[4] S. Gutta, H. Wechsler, and P. Phillips, “Gender and ethnic classification of face images”, Third IEEE International Conference on Automatic Face and Gesture Recognition (FG’98), pp. 194-199, 1998

[5] Phillips, P.J.,Hyeonjoon M.,Rizvi, S.A. and Rauss, P.J., “The FERET evaluation methodology for face-recognition algorithms,” IEEE Trans. Pattern Analysis and Machine Intelligence, , vol. 22, no. 10, pp. 1090-1104, Oct. 2000.

[6] B. Moghaddam and M.-H.Yang, “Gender classification with support vector machines” Proc. IEEE International Conference on Automatic Face and Gesture Recognition, pp. 306-311, March 2000.

[7] M. Nakano, F. Yasukata and M. Fukumi, “Age and gender classification from face images using neural networks” Proc. of Signal and Image Processing, 2004.

[8] X. Lu, H. Chen and A. K. Jain, “Multimodal facial gender and ethnicity identification”, Proc. Advances in Biometrics. Vol. 3832, pp. 554-561, 2005.

[9] H.-C. Kim, et al., “Appearance based gender classification with Gaussian processes”, Pattern Recognition Letters, 27(6), pp. 618-626, April 2006.

[10] Z. Yang, M. Li and H. Ai, “An experimental study on automatic face gender classification”, Proc. IEEE Int. Conf. on Pattern Recognition, pp. 1099-1102, 2006.

[11] S. Baluja, and H. Rowley, “ Boosting sex identification performance”, International Journal of Computer Vision, 71(1), pp. 111-119, January 2007.

[12] L. Lu and P. Shi, “Fusion of multiple facial regions for expression-invariant gender classification”, in IEICE Electron. Exp. 6(10), pp. 587-593, 2009

[13] L. A. Alexandre, “Gender recognition: A multiscale decision fusion approach” Pattern Recognition Letters, 31, pp. 1422-1427, 2010.

[14] T. Ojala, M. Pietkainen, and T. Maenpaa, “Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, July 2002.

[15] Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, Matti Pietikainen, Xilin Chen, and Wen Gao, “WLD: A Robust Local Image Descriptor”, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 32, No. 9, September 2010

[16] J. Zang and B. L. Lu, “A support vector machine classifier with automatic confidence and its application to gender classification”, Neurocomputing vol. 74, pp.1926-35, 2011