

Gender Classification from Facial Images Using Texture Descriptors

Ihsan Ullah¹, Hatim Aboalsamh¹, Muhammad Hussain², Ghulam Muhammad³, George Bebis⁴

¹*Department of Computer Science, King Saud University, KSA*

²*Department of Software Engineering, King Saud University, KSA*

³*Department of Computer Engineering, King Saud University, KSA*

⁴*Department of Computer Science and Engineering, University of Nevada at Reno, USA*

{ihsanullah, hatim, mhussain, ghulam}@ksu.edu.sa, bebis@cse.unr.edu

Abstract

Face recognition performance can be improved when face images are first classified into categories and then analysed with category-specific descriptors. One such category is gender. The face image is a type of texture that can be represented using texture descriptors. We employ two state-of-the-art texture descriptors, the local binary pattern (LBP) and Weber's law descriptor (WLD), and investigate their spatially enhanced versions (SLBP and SWLD) for gender classification. A suitable choice of parameters used in these descriptors leads to significant improvement. The best combination of parameters is found through a large number of experiments performed on the FERET and Multi-PIE databases. Using these parameters, the SLBP and SWLD perform much better with less algorithmic complexity compared to commonly used gender recognition approaches.

Keywords: Gender recognition, Local binary pattern, Weber's law descriptor, Face recognition.

1 Introduction

Face recognition continues to be a demanding problem. In this study, we approach the problem from the viewpoint of category-specific face recognition. Faces are first categorized based on visual cues, and then face recognition is performed using category-specific features. One such visual cue is gender. For an initial evaluation, we performed face recognition with the FERET database [1] (sets fa and fb), achieving a recognition rate of 95.8%. Then, we separated the faces into male and female categories and performed face recognition on each category; the recognition rate for the male category was 98.78% and that for the female category was 92.98% using the same face descriptor i.e., Weber's law descriptor (WLD) [2]. Thus, although the overall result using WLD for combined males and females is fairly good, it does not reveal the poorer female results compared to the male results, which are shown only when applying categorization based on gender.

The initial experiment indicated that the same face descriptor does not perform well for both gender

categories, and therefore, face recognition performance can be improved if first the faces are categorized and then category-specific descriptors are used for feature extraction. In this approach, the first problem is to classify faces based on category. Gender classification is important from other perspectives as well. In this study, we address the issue of gender classification. Significant work has been performed in the domain of gender classification over the last two decades [3-6].

Humans can easily identify the gender of a subject, but computer vision continues to fall short in the development of gender classification systems for biometrics, criminology, etc. Robust gender classification systems could assist in the surveillance of prohibited areas, searches for specific individuals at airports and borders, and crowd surveillance during sporting events to detect disturbances caused by different groups. In addition, reliable gender classification systems could improve commercial effectiveness in places such as parks and markets where people are entertained according to their gender.

Many systems have been proposed in the literature to classify images by gender; however, images are often misclassified due to variant illumination, different rotations, or varying poses. The steps in gender recognition are similar to those in face recognition, i.e., preprocessing, feature extraction, and classification. In the system, a descriptor is required to extract the most discriminating features from each image. Two types of features can be extracted and utilized, causing the gender classification methods to be roughly divided into two categories, geometrical feature-based and appearance-based. In the former category, geometrical features, such as the dimensions of the face, and the location and size of important facial features, such as the eyes, nose and mouth, are considered. The distances between these features are used as a unique measurement set for each subject. In the latter category, non-geometrical features are extracted from the facial images through various feature extraction techniques such as PCA, DCT, Gabor, and Wavelets [6-7][26]. To date, appearance-based methods are considered more accurate than geometrical-based methods. The second phase is the selection of the classifier. The classifier must be selected to provide maximum accuracy when combined with the feature

descriptor. Appropriate classifier selection is an important part, this help in boosting performance recognition system [27].

In the literature, LBP [3-4][6] and WLD [2][8] descriptors are the current state-of-the art descriptors used for texture classification. Because a face image is a type of texture, these descriptors can be used to represent face images that can lead to a better recognition rate in combination with simple minimum distance classifiers in comparison to sophisticated classifiers such as SVM [6].

In our work, we report the first use of WLD in the extraction of features specific for gender classification as well as an enhancement of its performance obtained by maintaining spatial information while creating a histogram. We also investigated LBP features to enhance gender classification performance compared with previous results reported using LBP. The minimum distance classifiers city-block (L1), Euclidean (L2), cosine (COS), and chi-square (CS) are used, which are less complicated than neural network techniques or SVM.

Section 2 presents a review of the current literature in the field of gender classification. In Section 3, a methodology for gender recognition is proposed, including the feature extraction and selection details. Section 4 explains how the features are fused for better performance. The results and a subsequent discussion are presented in Section 5. In Section 6, the results are compared with the latest reports in the literature. Finally, conclusions are presented in Section 7.

2 Literature Review

Gender recognition by humans is easy. However, for a computer, it is not a simple task. Similar to face recognition, gender classification from facial images encounters many problems due to the variant illumination, rotations, and poses present in the images. Another problem arises from the large number of features extracted by feature descriptors. To classify images for gender, many researchers have developed complex systems based on either geometrical features or appearance features.

With respect to systems using geometrical features, a combination of principal component analysis (PCA) [9] and artificial neural networks (ANNs) [3][10] was first used 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. This system was more robust because of the consensus provided by RBF, and it performed well because of the benefits provided by the flexible and adaptive threshold provided by the decision trees. In their experiments, they used 3,006 images of 1,009 subjects, using cross validation on manually segmented

and normalized images, 64×72 pixels in size, from the FERET database and reported an accuracy of 96%. Moghaddam and Yang [11] proposed gender classification of facial images (21×12 pixels) using SVMs and reported a recognition rate of 97% on the colour FERET database. Nakano et al. [12] computed edge information and exploited a neural network classifier for gender recognition. Lu and Shi [13] exploited the range and intensity information of human faces for ethnicity and gender identification using a support vector machine (SVM).

With respect to systems using appearance features, Kim et al. [14] based their gender recognition system on a Gaussian process classifier and showed that the proposed system outperformed SVM with cross validation for most of the data tested. Yang et al. [15] improved gender classification using texture normalization on images of up to 32×32 pixels, reporting a 15% error rate with SVM. Baluja and Rowley [16] combined several weak classifiers based on pixel value comparisons of low-resolution grey scale images in their AdaBoost gender classifier. Tests performed on 20×20 pixel normalized images from the FERET database showed an overall accuracy of 90%. Lu and Shi [13] fused the left eye, upper face region, and nose in their gender classification approach. Their results showed that the fused face region approach out-performs the whole face approach. Extending upon this idea, Alexandre [5] used a fusion approach based on features from multiple scales. The author used normalized resized images (20×20 pixels, 36×36 pixels, and 128×128 pixels) to extract shape and texture features. For texture features, a local binary pattern [10] approach was used for the whole image. Jabid et al. [17], inspired by LBP, proposed a feature extraction technique called the local directional pattern (LDP) technique with an SVM classifier, reporting a gender recognition rate of 95.05% on 2,000 images from the FERET database. Sun et al. [18] used LBP for gender classification with the chi-square and AdaBoost algorithms to classify 2,000 images from the FERET dataset, including 1,200 male and 800 female images. They reported a 95.75% overall recognition rate. Luis [5] proposed a multiscale decision fusion approach tested on the FERET dataset using 411 images. He reported an accuracy of 93.46% when using a block size of 16×16 pixels for the LBP descriptor. Zheng and Lu [6] used several descriptors to compare performances with and without SVM for gender classification. In their experiments, they used the FERET dataset with only 992 male and female images. Using Gabor, LBP, and the multi-resolution LBP, they reported 91.6%, 93.1%, and 93.8% accuracy, respectively, for SVM. However, the local Gabor binary pattern combined with local discriminative analysis as the feature subset selection achieved 99.1% accuracy using SVM.

Recently, Chen et al. [2] introduced a new feature descriptor based on Weber's law, Weber's law descriptor (WLD). WLD showed good results for texture recognition. Wang et al. [8] applied Weber's law to normalized face images with illumination for face recognition problem and obtained promising results. They achieved a recognition rate of 94.7% using the CMU-PIE database. In this work, we explore and enhance WLD to extract the most discriminative features for gender classification.

When using geometrical features, only a few features are employed. In contrast, one of the main issues encountered in using appearance features is the large number of features that must be managed. To overcome this issue, researchers have used feature subset selection techniques to reduce the number of features as well as the ambiguity within the features. Researchers have used PCA, LDA, genetic programming, and SVM to select the optimal set of features. Lian and Lu [20] proposed a feature subset selection algorithm that resulted in a large reduction of features in two class problems.

3 Methodology

We employ a commonly used architecture for recognition systems, as shown in Figure 1.

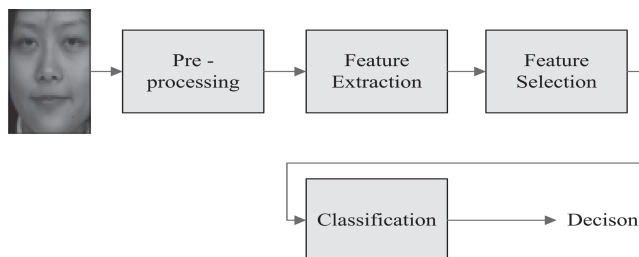


Figure 1 Gender Recognition System

Our primary focus is to find the best face description. In our initial experiments, we used the model shown in Figure 1, but without the feature selection phase. In pre-processing, we align the images in the centre so that only the face region is analysed. The techniques used in the feature extraction phase are presented in Section 3.1, where we explain the LBP descriptor in depth and the WLD is introduced and enhanced for gender classification. We used their spatially enhanced versions, SLPB and SWLD, in the feature extraction phase of our gender recognition system. In each case, we empirically explored the parameter space to find the best combination of parameters to achieve the maximum accuracy.

In the last phase of our system, we used the minimum distance classifiers L1, L2, COS, and CS to classify images into male or female. To reduce the dimensionality

of the feature space and to eliminate redundant features, we applied Sun's algorithm to select only the most discriminating features after the feature extraction step. This methodology will be discussed in Section 3.2.

3.1 Feature Extraction

In a gender recognition system, feature extraction plays an important role. In our system, we employ LBP and WLD because of their discriminative nature.

3.1.1 LBP Descriptor

LBP, introduced by Ojala et al. [10], is a widely used texture descriptor and has shown promising results in many applications. Ahonen et al. [3] used LBP for face recognition; Lu and Shi [13] and Sun et al. [18] employed it for gender recognition. In this section, we give a brief review of LBP and introduce its extension, the spatially enhanced LBP, which has a higher discriminatory power.

• Basic LBP

The basic LBP operator $LBP_{8,1}$ associates a label with each pixel of an image; the label is obtained by converting each pixel value in the 3×3 neighbourhood of a pixel into a binary digit (0 or 1) using the centre value as a threshold and concatenating the bits as shown in Figure 2. This operator has also been extended to general neighbourhood sizes as $LBP_{P,R}$ [10], where P is the number of points in the neighbourhood and R is its radius. The histogram of the labels is then used as a texture descriptor.

The rotation-invariant LBP operator, $LBP_{P,R}^{ri}$, assigns the P-bit number with the minimum value, obtained by applying circular bit-wise right shifts to the P-bit number of the pixel obtained using $LBP_{P,R}$, to each pixel. In the case of an (8, R) neighbourhood, there are 36 unique rotation-invariant binary patterns, and the dimension of the rotation-invariant LBP descriptor is 36.

Uniform microstructures, such as spots, flat areas, and edges, are binary patterns with up to two-bit transitions in their binary code. Usually, greater than 90% of the fundamental micro-features of texture are captured by them. A uniform LBP operator is denoted by $LBP_{P,R}^{u2}$. In the case of an (8, R) neighbourhood, there are 58 unique rotation-invariant binary patterns and the dimension of the uniform LBP descriptor is 58.

The general LBP operator, $LBP_{P,R}^{riu2}$, has three

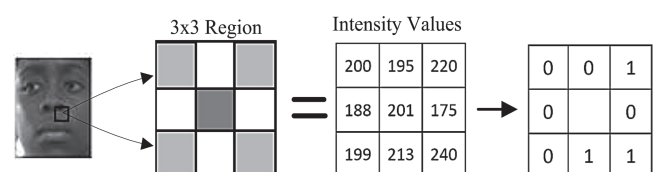


Figure 2 LBP Code with a 3×3 Neighbourhood of Central Pixels, i.e., 00101100 (Binary Code) or 44 (Decimal Form)

parameters: the circular neighbourhood (P, R), rotation invariance (ri), and uniformity (u2). For a particular application, this parameter space must be explored to find the best combination of these parameters.

• Spatially Enhanced LBP

As the name indicates, LBP is a local feature. However, the LBP descriptor is holistic because it is a histogram of labels, i.e., the labels are separated into bins according to their values, irrespective of their spatial location. The spatial locations of the micro-patterns, which are also important for recognition, are lost. Local salient patterns can also be lost when an image has different texture patterns at different locations. For example, two similar microstructures occurring in two different patterns with different spatial locations will contribute to the same bins in the histogram and will not be discriminated by the LBP descriptor. To enhance the discriminatory power of the LBP descriptor, we incorporate spatial locality into the descriptor. Each image is divided into n blocks, B_1, B_2, \dots, B_n . The LBP histogram, HB_i , is computed for each block B_i , where $i = 1, \dots, n$ and then the histograms are concatenated to form a spatial LBP descriptor (SLBP) $H = \{HB_i; i = 1, 2, \dots, n\}$ [9] and [18], as shown in Figure 3. SLBP descriptor performs better because it captures the spatial influence of the features. This approach introduces another parameter, the number of blocks or the size of blocks used to divide the image. For a particular application, using the optimal value of this block parameter leads to better recognition results.

There are six parameters to set in SLBP: ri, u2, riu2, (P, R), and n , the number of blocks. When SLBP is used for gender recognition with a suitable set of these parameters, it can yield results as good as those obtained with other more complicated approaches.

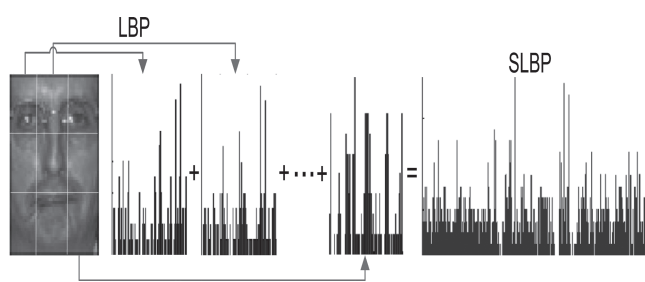


Figure 3 SLBP Descriptor for a Facial Image

3.1.2 WLD Descriptor

In this section, we provide an overview of the basic WLD descriptor [2] and its extension. This descriptor represents an image as a histogram of differential excitations and gradient orientations and has several interesting properties, such as robustness to noise and illumination changes, elegant edge detection, and powerful image representation. The WLD descriptor, proposed

by Chen et al. [2] for texture representation, is based on Weber's law. The computation of the WLD descriptor involves three steps: calculating the differential excitations, computing the gradient orientations, and building the histogram.

• Differential Excitations

The differential excitation $\varepsilon(x_c)$ of a pixel x_c is calculated as follows:

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

where P is the number of pixels x_i in the neighbourhood (P, R) of the pixel x_c and I_c and I_i are the intensity values of x_c and x_i , $i = 1, 2, \dots, P - 1$, respectively. The differential excitation $\varepsilon(x_c)$ may be positive or negative. A positive value indicates that the current pixel is darker than its surroundings, and a negative value indicates that the current pixel is lighter than its surroundings.

• Gradient Orientation

For a pixel x_c , the gradient orientation is calculated as follows:

$$\theta(x_c) = \arctan \left[\frac{I_{73}}{I_{51}} \right] \quad (2)$$

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

The gradient orientations are quantised into T dominant orientations as follows:

$$\phi_i = \frac{2t}{T} \pi \text{ where } t = \text{mod} \left(\left\lfloor \frac{\theta}{2\pi/T} + \frac{1}{2} \right\rfloor, T \right) \quad (3)$$

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

• Histogram Construction

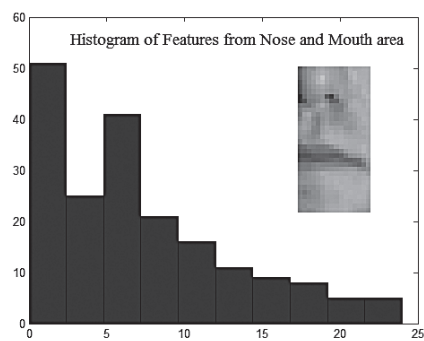
The differential excitation and dominant gradient orientation are calculated for each pixel. Using these features, the WLD histogram is constructed, for detail consult [2]. The descriptor involves three free parameters as follows:

- T , the number of dominant orientations of ϕ_i : $t = 0, 1, 2, \dots, T-1$,
- M , the number of segments $H_{m,t}$ of each sub-histogram H_t corresponding to a dominant orientation ϕ_i , and
- S , the number of bins in each sub-histogram, $H_{m,t}$.

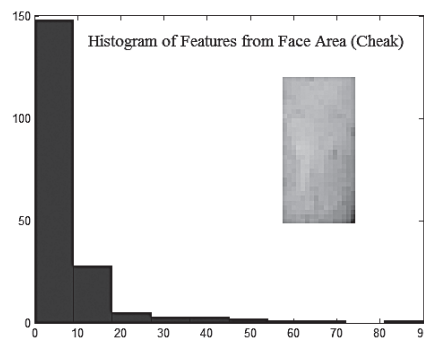
• Spatially Enhanced WLD

The WLD is a local feature; for similar reasons described for the SLBP descriptor in Section 2 and similar

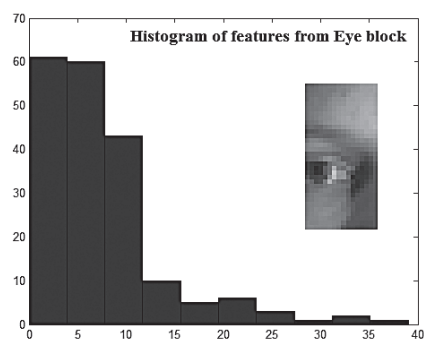
to the approach used for SLBP, we build a spatially enhanced WLD descriptor (SWLD). In SWLD, the bins encode not only the gradient orientation information but also the spatial locality of salient micro-patterns. An appropriate choice of the additional parameter, the number of blocks, can lead to better recognition results. To examine the performance of SWLD, four parameters must be tuned: T, M, S, and n, the number of blocks. Figure 4(a), (b) and (c) show the effect of SLBP on different blocks corresponding to different areas of the face. Figure 4(a) shows nose and mouth region along with histogram generated by feature extractor SLBP. Similarly Figure 4(b) and (c) depict cheek and eye, respectively, along with their LBP histograms. These figures show that less important features are extracted from the block of non-discriminative area (cheek) of face while more features are extracted from discriminative parts like eyes, nose and mouth.



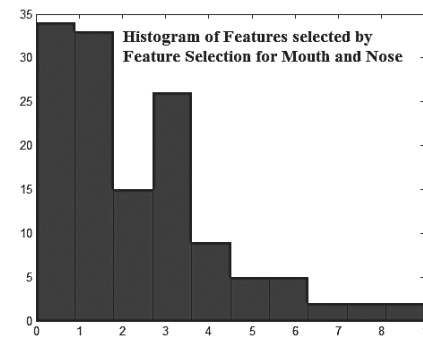
(a)



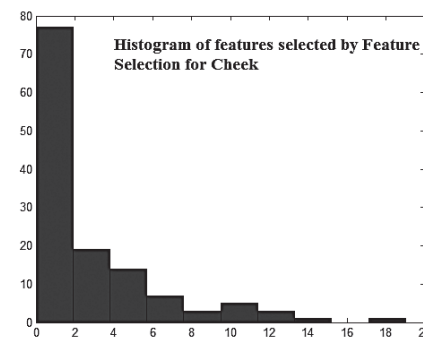
(b)



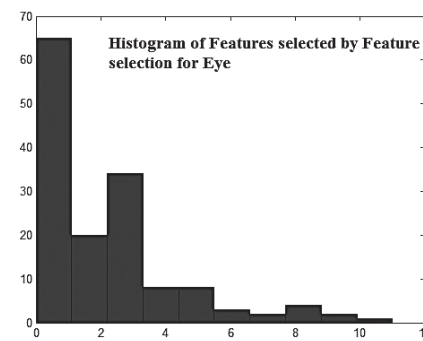
(c)



(d)



(e)



(f)

Figure 4 (a), (b) and (c) Blocks Containing Important Areas and Their Respective Histograms Generated by Feature Extraction Techniques while (d), (e) and (f) Shows Histograms Generated by Features Selected by Feature Selection Algorithm

3.2 Feature Selection

The spatially enhanced descriptors involve redundant features, which not only increase the dimension of the feature space, the “curse of dimensionality,” but may also result in a decrease in detection accuracy. We employ the method proposed by Sun et al. [19] for feature selection, which has shown good results in both reducing the dimensions of the feature space and increasing detection accuracy. This method is simple, powerful, and robust. It uses two free parameters, the kernel width, σ , and the regularisation parameter, λ . The proper choice of these parameters is imperative to obtain the best results. We used

a grid search, described below, to find the optimal values of these parameters.

```

For  $\sigma = 0.1:0.2:2$ 
  For  $\lambda = 0.1:0.2:2$ 
    SelectedF = Sel_Features_Sun_Met( $\sigma, \lambda$ );
    acc = Classify_with_NN(SelectedF);
  End
End

```

Although Sun's technique is a filter method, we employed it as a wrapper method in our system to increase performance. Figure 4(d), (e) and (f) show LBP histograms after feature selection using Sun's algorithm. It is obvious from these figures that Sun's algorithm selects more features from blocks containing more discriminative facial parts (eye, nose mouth) and fewer features from blocks containing less discriminative parts like cheeks. We gave freedom to Sun's algorithm by providing concatenated features generated by all the blocks of an image and let it select subset of most discriminative features. For blocks shown in Figure 4, it extracts 147 and 133 features out of 192 features for eye block and mouth and nose block, while only 130 from cheek area.

4 Fusion of SLBP and SWLD

The results obtained from testing and tuning the different parameters of SLBP, SWLD, and the feature subset selection technique for male and female images were not the same. One method works well for female recognition, and the other works well for male recognition. Therefore, we decided to fuse the features extracted for the best cases and to then use the fused version once again for the final gender recognition. The model used is shown below in Figure 5.

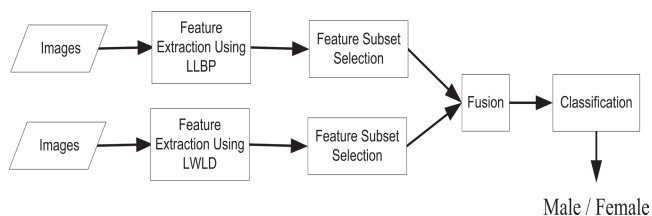


Figure 5 Fusion of SLBP and SWLD with FSS

In the classification stage, we employ only the L1 and CS distance measures because these measures yielded the best results from the previous experiments.

5 Experiments and Discussion

Face Recognition faces challenge due to various

numbers of factors which includes illumination, pose, expression and occlusions. For this reason, to compare and evaluate performance of different feature extraction techniques, several face databases have been collected over the last decade such as Multi-PIE [21], FERET [1], AR [22], XM2VTS [23], Cohn-Kanade [24], and Yale B [25]. However, we selected Multi-Pie and FERET for our experiments due to its variation and complexity in order to evaluate and compare our system with existing ones.

• FERET

We used FERET database [1] in our experiments, which is one of the challenging databases for face recognition; it contains a large number of images acquired during different photo sessions and has a good variety of gender groups. The lighting conditions, face orientation, and time of capture vary. All faces were normalised in terms of orientation, position, and size prior to experimentation. They were also masked to include only the face region (i.e., the upper body and background were cropped), yielding an image size of 60×48 pixels. In total, there are 1,486 and 914 images of 403 male and 403 female subjects, respectively. We used a set of 1,204 images for training and a set of 1,196 images for testing. The number of test images of males and females was 740 and 456, respectively. A representative sample is shown in Figure 6.



Figure 6 Samples of Male and Female Face Images from the FERET Database

The dataset is sampled again for another experiment but with a low resolution of 20×16 pixels. This low-resolution dataset is used to show the effect of image resolution on recognition rate.

• Mutli-PIE

Multi-PIE (Carnegie Mellon University Multi Pose Illumination and Expression) Multi-PIE [21] contains 337 subjects of different ethnicity and age, imaged under 15 view points and 19 illumination conditions in up to four recording sessions. The lighting conditions, face orientation, and time of capture vary. All faces were

normalised in terms of orientation, position, and size prior to experimentation. They were also masked to include only the face region (i.e., the upper body and background were cropped), yielding an image size of 64×64 pixels. In total we have taken 1,776 images of 1,020 male (102 subjects) and 756 female images (95 subjects). We used a set of 985 images for training and a set of 791 images for testing. Number of male and female images in training set is 510 and 475 respectively. Whereas remaining 510 and 281 images of male and female respectively were used in testing dataset. A representative sample is shown in Figure 7.

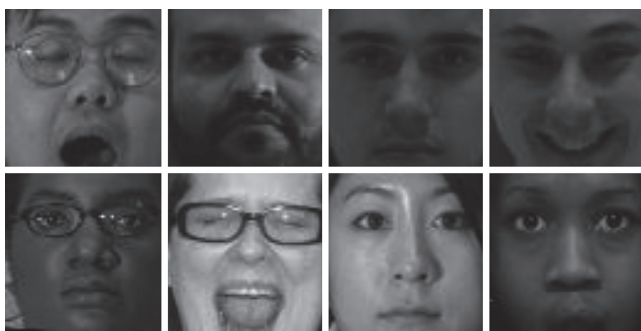


Figure 7 Samples of Male and Female Face Images from Multi-PIE Database

5.1 Experiments with LBP and SLBP Descriptors

In this section, we show the performance of LBP, SLBP, and SLBP with Sun's feature selector on FERET and Multi-PIE databases for gender recognition.

5.1.1 Basic LBP

The basic LBP has previously been used in gender recognition, but the performance was not satisfactory, even in the presence of sophisticated classifiers and a small number of images. In our experiments, first we tested basic LBP on FERET dataset with minimum distance classifiers L1 and CS. In this case, LBP histogram is computed without dividing an image into blocks. It resulted in maximum of 86.21% and 85.79% accuracy with L1 and CS respectively. However when tested with Multi-PIE database, it resulted in maximum of 91.44% with CS classifier and uniform LBP.

5.1.2 SLBP

Then we used SLBP to check its performance as compared to basic LBP and in order to report the best set of parameters. We tuned all its parameters to achieve optimal recognition on different block sizes. The results in Figure 8(a) show impact of different block sizes and the variants of LBP on FERET dataset; the recognition rate with basic LBP descriptor is very poor (less than 90%), as shown in the first column for full images, whereas SLBP yields better recognition with different block sizes. We obtained the best result (99.00%) with a block size of 6×12 pixels, $LBP_{8,1}$

operator and CS classifier shown in Figure 9. However, the second best (98.83%) result was found using the $LBP_{8,1}$ and $LBP_{8,1}^{u2}$ operators with L1 classifier, as shown in Figure 8(a). In all cases, the CS results were slightly better than L1 results. The performance of $LBP_{8,1}^{ri}$ and $LBP_{8,1}^{riu2}$ operators were not satisfactory.

In case of Multi-PIE, SLBP resulted even better than it did in case of FERET. We performed similar tests with Multi-PIE dataset as we did for FERET. We achieved 99.16% and 99.04% accuracy with L1 and CS classifiers respectively as shown in Figure 8(b) for block size of 6×12 , uniform mapping.

System was tested with different neighbourhood (P, R) values of (8, 1), (8, 2), (16, 2) and (24, 3). The effect of (P, R) with values of (8, 1) & (8, 2) only and their fusion is shown in Figure 9; we found that they perform better than (16, 2), and (24, 3). When descriptors with different (P, R) values like (8, 1) and (8, 2) are fused, the recognition rate improves slightly in case of $LBP_{8,1}^{ri}$ and $LBP_{8,1}^{riu2}$, but there is no improvement when other variants of LBP are used, as is obvious from Figure 9. Fusion with other values of (P, R) was also tested but the results were not better than the best results shown in Figure 9.

5.1.3 SLBP with Feature Selection (FS)

Using best case parameters which results in 99.00% and 99.16% for FERET and Multi-PIE respectively, we performed feature subset selection on its respective

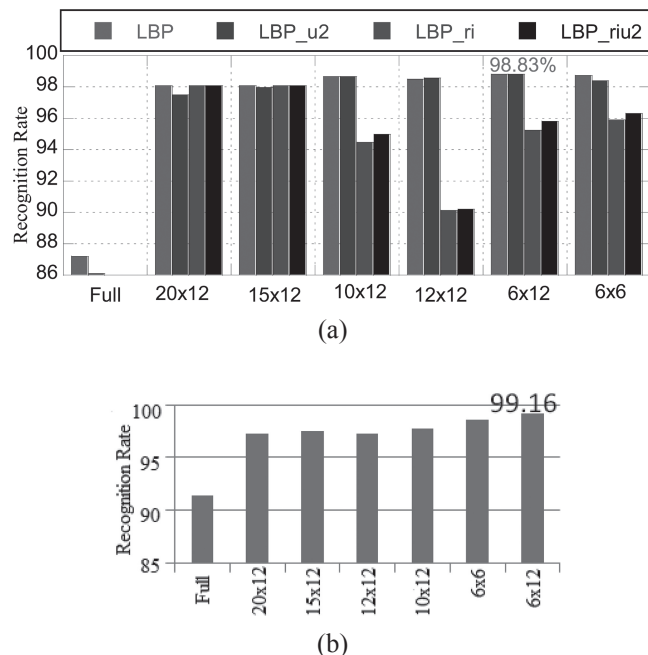


Figure 8 The Effect of Block Size on the Recognition Rate with L1 Classifier and the Different LBP Operators. The Horizontal Axis Represents the Block Sizes. "Full" Means the Full Image. "LBP" Means Basic LBP Operator (a) On FERET (b) On Multi-PIE Dataset

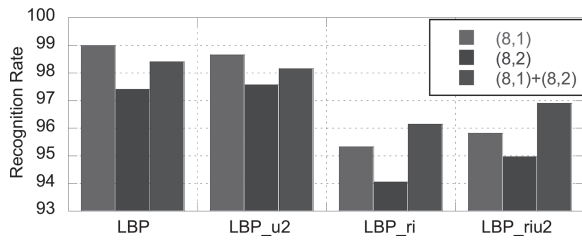


Figure 9 The Effect of Different Neighbourhoods (P, R) and Their Fusion on FERET Database

histograms with Sun’s algorithm for σ and λ values from 0.1 to 2 with an increment of 0.2. At $\sigma = 1.3$ and $\lambda = 0.9$, we achieved a recognition rate of 99.33% on FERET dataset. Whereas on Multi-PIE we achieved 99.38% with $\sigma = 1.3$ and $\lambda = 1.1$; FS not only increased the accuracy but also decreased the dimension of the feature space enormously from (10,240 to 6,321) and (6,490 to 916) in case of FERET and Multi-PIE dataset respectively.

5.2 Experiments with WLD and SWLD Descriptors

In this section, we will show performance of WLD, SWLD, and SWLD with Sun’s feature selector on FERET and Multi-PIE for gender recognition.

5.2.1 Basic WLD

First, we tested basic WLD on FERET and Multi-PIE datasets. To validate the performance of WLD for gender recognition, we tested various combinations of WLD parameters (T, M, and S), i.e., T = 4, 6, 8; M = 4, 6; S = 4, 8. The optimal result achieved was 89.3% with (T, M, and S) values of (8, 4, and 4) on FERET dataset while on Multi-PIE, it resulted in 94.71% with CS having 384 features and (T, M and S) values of (12, 4 and 8) respectively.

5.2.2 SWLD

After basic WLD, we used block sizes of 20×12 , 15×12 , 12×12 , 6×12 pixels to validate performance of SWLD for gender recognition. Figure 10 shows the results obtained from the SWLD experiments. For basic WLD, the accuracy is less than 90%. By using SWLD, a block-size of

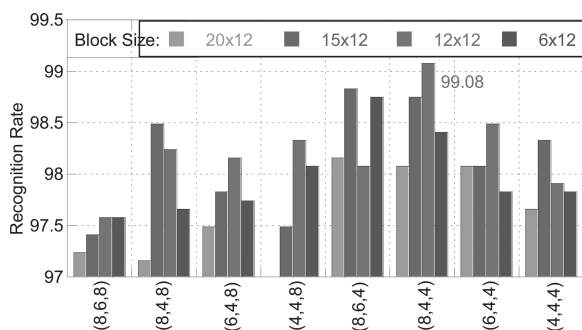


Figure 10 The Effect of Different Combinations of (T, M and S) and Block Sizes on Recognition Accuracy with FERET Dataset

12×12 pixels with (T, M and S) values of (8, 4 and 4) and CS classifier, best recognition rate of 99.08% was achieved for FERET with only 2,560 extracted features. On the other hand in case of Multi-PIE dataset, (T, M, and S) values of (12, 4, and 4) resulted in maximum accuracy of 98.8% having 6,144 extracted features.

Figure 11 shows the effect of image resolution on accuracy. Due to variations in the image sizes, the block sizes vary accordingly, i.e., the sizes of 10×16 , 10×8 , 5×8 , and 5×4 pixels are used. The low-resolution dataset achieved a maximum recognition rate of 88.8% with (T, M and S) values of (8, 6 and 4), and a block size of 5×8 pixels using L1 classifier.

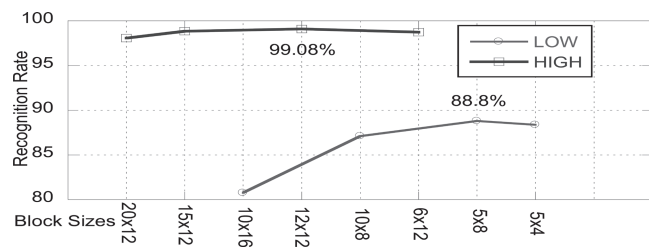


Figure 11 Effect of Low and High Resolution on Accuracy with FERET Dataset

Similarly, we tested different block size for same or different (T, M and S) values but result was below 99%.

5.2.3 SWLD with Feature Selection:

The feature subset selection (Sun’s) algorithm was then applied to the best case parameters (99.08% in case of FERET and 98.8% in case of Multi-PIE) scenarios. With Sun’s algorithm parameters values of $\sigma = 1.3$ and $\lambda = 0.7$, the accuracy in case of FERET increases to 99.25% and the dimension of its feature space reduces from 2,560 to 2,060. In case of Multi-PIE, we achieved 99.21% and 99.16% accuracy with L1 and CS respectively, having only 1878 features instead of 6144 for sun’s parameters of $\sigma = 0.7$ and $\lambda = 0.3$.

5.3 SLBP and SWLD Fusion Motivation and Results

In our experiments, we observed two issues concerning decreasing block size. First, decreasing the sizes of the blocks increases the number of features. Second, decreasing the block size after a certain limit reduces the overall accuracy. In addition, we observed that the images misclassified by both SLBP and SWLD with feature subset selection differed. Thus, we fused the features from both descriptors. We used the two best cases from the application of SLBP and SWLD. After computing the best cases with the SLBP and SWLD descriptors and then independently applying feature selection with $\sigma = 0.9$ and $\lambda = 0.7$, we fused the resulting features. The fusion does not increase the overall accuracy; but it does reduce the dimension of

the feature space. SLBP yields 99.33% accuracy with 6,321 features, and SWLD yields 99.25% with 2,060 features. After fusion, the system yielded an accuracy of 99.33% with 2,900 (1,306 from SLBP + 1,594 from SWLD) features, i.e., the dimension is reduced by 54%.

In case of Multi-PIE, after employing optimal parameters set in the same manner as we did for FERET, we achieved 99.55% and 99.61% with L1 and CS classifiers, respectively, $\sigma = 0.3$ and $\lambda = 0.3$, and 1640 features instead of 2,794 (916 from SLBP + 1,878 from SWLD without fusion) features. Whereas second best results of 99.38% and 99.61% were obtained with L1 and CS classifiers, respectively, and $\sigma = 0.9$ and $\lambda = 0.5$ and 1,006 features i.e., with more than 60% reduction in features.

These results show that the proposed system is a general system, which produces optimal results for various databases, if parameters are tuned properly.

6 Comparison

We compared the proposed method with the multi-resolution decision fusion method (MDF) [5], which also uses texture descriptors and geometric techniques such as PCA (self-tested) and LDA-SVM [6], and the state-of-the-art techniques presented in [7] i.e., the local Gabor binary pattern with LDA and SVMAC method (LGBP-LDA SVMAC) and the local Gabor binary pattern with LDA and SVM method (LGBP-LDA SVM). Even though our approach is simpler than these other methods, it results in better recognition accuracy, as shown in Figure 12.

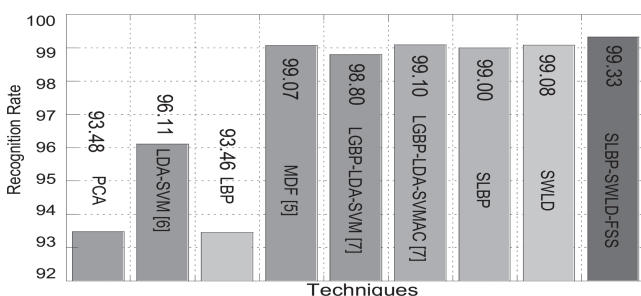


Figure 12 Comparison of the Proposed Method with Other Techniques

Our method not only surpasses other reported techniques in accuracy but also uses fewer features. LGBP-LDA-SVMAC [6] was reported to obtain a maximum of 99.1% on the FERET database with 992 male and female images when using 4,000 features per image. Our system achieved 99.33% using only 2,900 features. Gender recognition on Multi-PIE has not been reported in literature according to our knowledge; however the result achieved on Multi-PIE by our approach i.e., 99.61% with only 1,006 features is optimal till date.

7 Conclusion

Although LBP (previously used) and WLD (applied for the first time) operators capture local information from face images, the corresponding descriptors are holistic and do not produce adequate recognition results. The spatially enhanced versions of these descriptors resulted in an improvement in accuracy when applied to gender recognition. We investigated the parameter space for each operator and found that SLBP without any mapping and uniform mapping for FERET and Multi-PIE respectively, with a block size of 6×12 pixels performs best. Whereas SWLD gives the best results for T, M and S values of 8, 4, 4 and 12, 4, 4 for FERET and Multi-PIE respectively with block size of 12×12 pixels. Feature selection with Sun's algorithm (introduced for the first time for gender classification) improves the recognition accuracy and significantly reduces the dimension of the feature space. The overall recognition rates are 99.33% and 99.61%, which are the best results so far on the FERET and Multi-PIE databases, respectively. Fusion of the two descriptors reduces the dimension of the feature space in case of the FERET and Multi-PIE databases whereas in case of Multi-PIE it also increases accuracy. The same parameters do not produce optimal result for both datasets in some cases. As the internal structures of the images in the two databases are different due to device noises and other factors, so the tuning of parameters is needed for each database. When the parameters are tuned, the proposed method gives almost similar result.

There is still need of further investigation for fusion in order to find better method for optimal accuracy. Employing blocks of multiple sizes via, for example, spatial pyramid model can further improve the recognition accuracy. These ideas will be explored in future work.

Acknowledgements

This work is supported by NSTIP strategic technologies programs, under grant number 08-INF1044-02 in the Kingdom of Saudi Arabia.

References

- [1] P. Jonathon Phillips, Hyeonjoon Moon, Syed A. Rizvi and Patrick J. Rauss, *The FERET Evaluation Methodology for Face-Recognition Algorithms*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, No.10, 2000, pp.1090-1104.
- [2] Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, M. Pietikainen, Xilin Chen and Wen Gao, *WLD: A Robust Local Image Descriptor*, *IEEE Transactions*

- on *Pattern Analysis and Machine Intelligence*, Vol.32, No.9, 2010, pp.1705-1720.
- [3] T. Ahonen, A. Hadid and M. Pietikäinen, *Face Description with Local Binary Patterns: Application to Face Recognition*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.28, No.12, 2006, pp.2037-2041.
- [4] Ning Sun, Wenming Zheng, Changyin Sun, Cairong Zou and Li Zhao, *Gender Classification Based on Boosting Local Binary Pattern*, *Proc. of ISNN 2006*, Chengdu, China, May/June, 2006, pp.194-201.
- [5] Luis A. Alexandre, *Gender Recognition: A Multiscale Decision Fusion Approach*, *Pattern Recognition Letters*, Vol.31, No.11, 2010, pp.1422-1427.
- [6] Ji Zang and Bao-Liang Lu, *A Support Vector Machine Classifier with Automatic Confidence and Its Application to Gender Classification*, *Neurocomputing*, Vol.74, No.11, 2011, pp.1926-1935.
- [7] Juan Bekios-Calfa, Jose M. Buenaposada and Luis Baumela, *Revisiting Linear Discriminant Techniques in Gender Recognition*, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.33, No.4, 2011, pp.858-864.
- [8] Biao Wang, Weifeng Li, Wenming Yang and Qingmin Liao, *Illumination Normalization Based on Weber's Law with Application to Face Recognition*, *IEEE Signal Processing Letters*, Vol.18, No.18, 2011, pp.462-465.
- [9] Hui-Cheng Lian and Bao-Liang Lu, *Multi-view Gender Classification Using Local Binary Patterns and Support Vector Machines*, *Proc. of ISNN 2006*, Chengdu, China, May/June, 2006, pp.202-209.
- [10] Timo Ojala, Matti Pietikainen and Topi Maenpaa, *Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns*, *IEEE Transactions Pattern Analysis and Machine Intelligence*, Vol.24, No.7, 2002, pp.971-987.
- [11] B. Moghaddam and Ming-Hsuan Yang, *Gender Classification with Support Vector Machines*, *Proc. of IEEE International Conference on Automatic Face and Gesture Recognition*, Grenoble, France, March, 2000, pp.306-311.
- [12] M. Nakano, F. Yasukata and M. Fukumi, *Age and Gender Classification from Face Images Using Neural Networks*, *Proc. of Signal and Image Processing*, Honolulu, HI, August, 2004, pp.23-25.
- [13] Li Lu and Pengfei Shi, *Fusion of Multiple Facial Regions for Expression-Invariant Gender Classification*, *IEICE Electronics Express*, Vol.6, No.10, 2009, pp.587-593.
- [14] Hyun-Chul Kim, Daijin Kim, Zoubin Ghahramani and Sung Yang Bang, *Appearance Based Gender Classification with Gaussian Processes*, *Pattern Recognition Letters*, Vol.27, No.6, 2006, pp.618-626.
- [15] Zhiguang Yang, Ming Li and Haizhou Ai, *An Experimental Study on Automatic Face Gender Classification*, *Proc. IEEE Int. Conf. on Pattern Recognition*, Hong Kong, China, August, 2006, pp.1099-1102.
- [16] Shumeet Baluja and Henry A. Rowley, *Boosting Sex Identification Performance*, *International Journal of Computer Vision*, Vol.71, No.1, 2007, pp.111-119.
- [17] T. Jabid, H. Kabir and Oksam Chae, *Gender Classification Using Local Directional Pattern (LDP)*, *Proc. of 20th IEEE International Conference on Pattern Recognition (ICPR)*, Istanbul, Turkey, August, 2010, pp.2162-2165.
- [18] Ning Sun, Wenming Zheng, Changyin Sun, Cairong Zou and Li Zhao, *Gender Classification Based on Boosting Local Binary Pattern*, *Proc. of ISNN 2006*, Chengdu, China, May/June, 2006, pp.194-201.
- [19] Y. Sun, S. Todorovic and S. Goodison, *Local-Learning-Based Feature Selection for High-Dimensional Data Analysis*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.32, No.9, 2010, pp.1610-1626.
- [20] Hui-Cheng Lian and Bao-Liang Lu, *Multi-view Gender Classification Using Local Binary Patterns and Support Vector Machines*, *Proc. of ISNN 2006*, Chengdu, China, May/June, 2006, pp.202-209.
- [21] Ralph Gross, Iain Matthews, Jeffrey Cohn, Takeo Kanade and Simon Baker, *Multi-PIE, Image and Vision Computing*, Vol.28, No.5, 2010, pp.807-813.
- [22] Aleix Martinez and Robert Benavente, *The AR face database*, June, 1998. Computer Vision Center (CVC) Technical Report No. 24.
- [23] K. Messer, J. Matas, J. Kittler, J. Luetttin and G. Maitre, *XM2VTSDB: The Extended M2VTS Database*, *Proc. of Second International Conference on Audio and Video-Based Biometric Person Authentication (AVBPA)*, Washington, DC, March, 1999, pp.72-77.
- [24] Takeo Kanade, Yingli Tian and Jeffrey F. Cohn, *Comprehensive Database for Facial Expression Analysis*, *Proc. of 4th IEEE International Conference on Automatic Face and Gesture Recognition*, Grenoble, France, March, 2000, pp.46-53.
- [25] Athinodoros S. Georghiades, Peter N. Belhumeur and David J. Kriegman, *From Few to Many: Generative Models for Recognition under Variable Pose and Illumination*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.23, No.6, 2001, pp.643-660.
- [26] Lijun Yan, Weimin Zheng, Shu-Chuan Chu and John F. Roddick, *Neighborhood Discriminant Nearest*

Feature Line Analysis and Its Application to Face Recognition, Journal of Internet Technology, Vol.14, No.1, 2013, pp.127-132.

- [27] Yanfeng Zhang, Xiaofei Xu, Yingqun Liu, Xutao Li and Yunming Ye, *A Novel Decision Cluster Classifier with Nested Agglomerative K-Means, Journal of Internet Technology, Vol.14, No.1, 2013, pp.145-152.*

Biographies



Ihsan Ullah has been recently awarded PhD scholarship from University of Milan and started his PhD in CVPR lab of University of Parthenope Naples Italy. He did his Bachelor in IT from DCS University of Peshawar in March 2008. He got Scholarship from King Saud University and secured his MS (CS) degree. This Paper is one of the outcomes of his thesis. He has published papers in well reputed conferences and journals. His research interests include Computer Vision, Machine Learning and Pattern Recognition.



Hatim Aboalsamh is a professor in the Department of Computer Science, King Saud University, Riyadh, Saudi Arabia. He is former dean of College of Computer & Information Sciences. He was also the Vice President for development and Quality of King Saud University. He is an IT consultant and a fellow member of British Computer Society (BCS), a senior member of ACM, and a member of many other professional societies. He has over 70 referred publications and co-editor of two books.



Muhammad Hussain is a professor in the Department of Software Engineering, King Saud University. He did his PhD in Computer Science, Kyushu University, Fukuoka, Japan. He worked as a researcher at Japan Science and Technology Agency from April 2003 to September 2005. In September 2005, he joined King Saud University. He worked on a number of funded projects in Kyushu University, Japan and King Saud University, Saudi Arabia. His current research interests include Multimedia Forensics, Digital Image Processing, and Pattern Recognition. He has published more than 100 publications in well reputed journals and conferences.



Ghulam Muhammad is an associate professor in the Department of Computer Engineering King Saud University. He did his PhD in March 2006 from Electronics and Information Engineering, Toyohashi University of Technology, Japan. His Thesis Title was “A study on auditory based feature extraction and HMM/SM based classification for robust automatic speech recognition.” He published several papers in well reputed conferences and journals. His main research interests are Multimedia Forensics, Speech Signal Processing and Digital Image Processing.



George Bebis received PhD degree in Electrical and Computer Engineering from the University of Central Florida in 1996. He is professor in the Department of Computer Science and Engineering at the University of Nevada, Reno (UNR) and director of UNR’s Computer Vision Laboratory (CVL). His research interests include computer vision, machine learning and pattern recognition in which he has several conference and journal publications.

