Illumination Invariant Face Recognition Using the Statistical Features of BDIP and Wavelet Transform

Amany Farag and Randa Atta

Abstract—**In this paper, an efficient feature extraction method based on local statistics features of block difference of inverse probabilities (BDIP) and the wavelet transform is proposed for face recognition. In the proposed method, the BDIPs are first computed in a face in order to overcome the variation of illumination and facial expressions. The obtained BDIP image is then decomposed into wavelet subbands. In order to reduce the dimensionality of the feature vector, each BDIP subband is partitioned into a set of blocks. The means and variances are then calculated from all the blocks in each subband and are fused into a feature vector. Experimental results on ORL and FERET databases show that the proposed method achieves higher recognition accuracies than the wavelet-based methods with higher dimensionality reduction of the feature vector. It also outperforms the other well known methods such as PCA and the DCT with the zigzag scanning.**

Index Terms—**Face recognition (FR), discrete wavelet transforms (DWT), wavelet packet decomposition (WPD), block difference of inverse probabilities (BDIP), support vector machine (SVM).**

I. Introduction

Face recognition has gained significant attention in the last two decades due to the increasing demand on its applications such as personal identification. Therefore, many face recognition methods have been proposed [1-6]. However, the performance of reliable methods varies due to the image variations caused by illumination conditions, facial expressions, poses, and other factors.

One of the most popular face recognition methods is eigenface technology [1], which is based on principal component analysis (PCA). It includes a linear core process that projects the high-dimensional data into a lower dimensional space. However, this method requires high computational cost in determining the basis space for a large number of training images. To overcome this problem, many face recognition methods based on the transforms such as the discrete cosine transform (DCT) and discrete wavelet transforms (DWT) have been proposed. In [2], the most relevant DCT features are extracted using the zigzag scan and the obtained features are fused either at the feature level or at the decision level for face recognition. On the other hand, the wavelet-based methods focus on the subbands that contain the most relevant information to better represent the face image as in [3]. Moreover, these methods have been combined with other techniques to achieve better face recognition performance [7].

One of the most important factors to degrade the performance of face recognition is known to be the illumination variation problem. Many methods have been proposed to solve this problem [8-10]. Zhang et al. [9] proposed a gradient faces method as an image preprocessing technique for face recognition under varying lighting. In [10], the BDIP and BVLC (block variation of local correlation coefficients) operators have been applied to face recognition to overcome the variation of illumination problem. Both of the operators are bounded and well locally normalized to be robust to illumination variation. BDIP is a kind of nonlinear gradient operator normalized by local maximum, which is known to effectively measure the local brightness variations so that edges and valleys are extracted well.

The cost of classification can be reduced by limiting the number of features which must be measured and stored. The feature selection has a considerable impact on the results of any classification algorithm. A number of approaches for feature selection have been proposed. Among them, a feature extraction method based on the embedded zero-tree of the DCT and wavelet transforms has been proposed in [4]. It allows selecting a subset of the most important coefficients to improve the recognition rates. The method in [11] uses the means and variances extracted from the wavelet subbands to obtain the feature vector with the minimal dimension.

In this paper, a face recognition method based on the BDIP features and the wavelet transform is proposed in order to reduce the effect of the variation of illumination. In the proposed method, to reduce the dimension of the feature vectors, the block-based statistical measures such as means and variances are then calculated from all blocks in each subband. After extracting the feature vectors, the Euclidean distance and the support vector machines (SVM) are used for classification. The remainder of the paper is organized as follows: In Section II, the proposed feature extraction method is described. Then in Section III, the basic theory of the Euclidean distance and the support vector machines are briefly reviewed. Experimental results and conclusions are given in Sections IV and V, respectively.

II. Facial Features

In this section, the conventional features which are partly used in the proposed face recognition system will be described.

A. BDIP in the Spatial Domain

BDIP for an image I is defined as

Manuscript received January 11, 2012; revised January 28, 2012.

A. Farag is with the Computer Science Department, Port Said University, Port Said, Egypt (e-mail: amanyfarag@hotmail.com).

R. Atta is with the Electrical Engineering Department, Faculty of Engineering, Port Said University, Port Said, Port Fouad 42523, Egypt (e-mail: r.atta@eng.psu.edu.eg).

$$BDIP(x,y)=\frac{\frac{1}{|R(x,y)|}\sum_{(p,q)\in R(x,y)}[I_{max}(x,y)-I(p,q)]}{I_{max}(x,y)} \quad (1)$$

where $I(x,y)$ denotes the value at a pixel (x,y) in the image I, and $R(x,y)$ is a local region whose center is the pixel (x,y). $|R(x,y)|$ and $I_{max}(x,y)$ are the number of pixels and the maximum value in the local region, respectively. The numerator is selected as a representative gradient in the local region, which is defined by the averaged difference between the maximum pixel value and each pixel value in the local region. The denominator is defined by the maximum pixel value. Therefore, (1) gives the result of gradient operator normalized by the representative, which yields a sketch-like image. Fig. 1 illustrates some original images and their BDIP images.

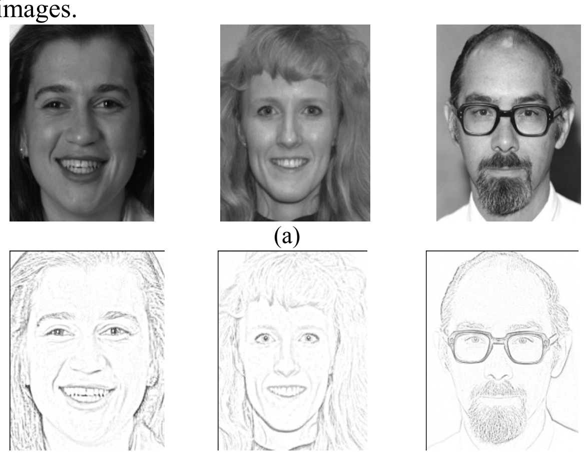


Fig. 1. Original images and their BDIP images. (a) Original images and (b) BDIP images.

B. Wavelet Transform Features

In wavelet transform, a two-dimensional image is filtered by LPF (low pass filter) and HPF (high pass filter) along the horizontal direction and vertical direction. Therefore, one level of decomposition yields four sub-bands: one smooth subband (LL) and three detail subbands (LH, HL, and HH). In the DWT, each level is calculated by passing only the previous approximation coefficients through low and high pass filters. However, in the wavelet packet decomposition (WPD), both the details and approximation coefficients are decomposed. Fig. 2 illustrates the effect of applying 2-level of decomposition of both DWT and WPD on an image and its BDIP image.

C. Proposed Feature Extraction Method

The block diagram of the proposed feature extraction method is shown in Fig. 3. The first step towards implementing the proposed feature extraction method is that BDIP image is extracted from an original image with 3x3 moving window for every pixel using (1). Therefore, the size of the extracted feature image is equal to that of an original image. BDIP can extract sketch-like feature images, where edges and valleys around the eyes and lips are more emphasized as shown in Fig. 1(b).

The extracted BDIP image is then decomposed into a set of subband images by using either DWT or WPD. Next, each subband is partitioned into a set of equally-sized blocks. Moreover, in order to determine the optimal block size among the decomposition levels, two types of partitions have been implemented:

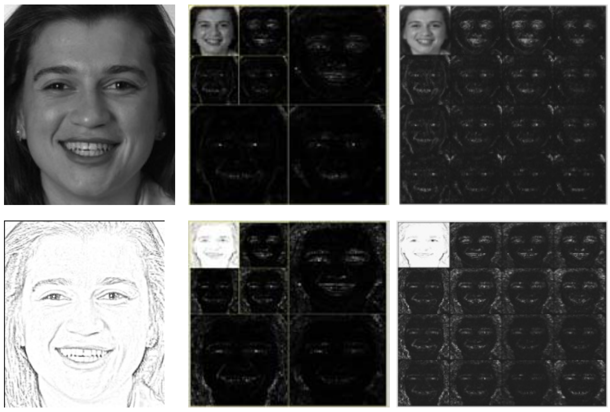


Fig. 2. The wavelet transform of an image (2-level of decomposition) (a) Original image & BDIP image, (b) DWT images, and (c) WPD images.

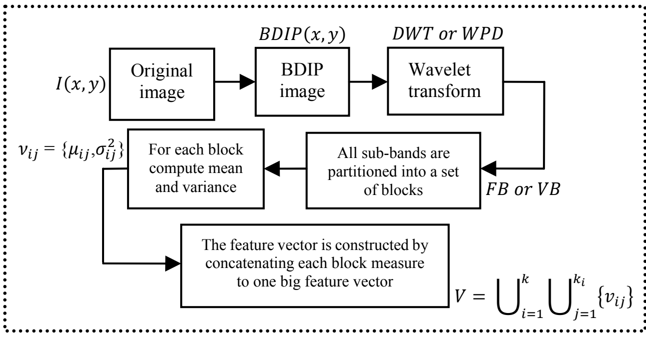


Fig. 3. Block diagram of the proposed feature extraction method.

1) Fixed Blocks size (FB)

All subband images in each decomposition level are partitioned into a set of equally-sized blocks such that the size of the blocks among the decomposition levels is fixed. The block size can be chosen for instance as 10×10 or 8×8.

2) Varying Blocks size (VB)

All subband images in each decomposition level are partitioned into a set of equally-sized blocks but the sizes of the blocks among the decomposition levels are variable. The reason behind that is the subband images at the lower decomposition level contain more important information than those at the higher levels. Therefore, the (LL) subband can be partitioned into a set of small-sized blocks while the other subbands (LH, HL and HH) are partitioned into a set of bigger-sized blocks. The (LL) subband is partitioned into blocks with a size chosen half of that used with (LH, HL and HH) subbands.

Finally, the statistical measures such as the mean and the variance calculated from each block at each decomposition level can be used to form the feature vectors. Let $I_{ij}(x,y)$ be the image at the specific block j of subband i, the resulting feature vector is $v_{ij}=\{\mu_{ij},\sigma_{ij}^2\}$, where μ_{ij} and σ_{ij}^2 are the mean and the variance, respectively, and are defined as:

$$\mu_{ij}=\frac{1}{M\times N}\sum_{x=1}^{M}\sum_{y=1}^{N}\left|I_{ij}(x,y)\right| \quad (2)$$

$$\sigma_{ij}=\frac{1}{M\times N}\sum_{x=1}^{M}\sum_{y=1}^{N}\left|I_{ij}(x,y)-\mu_{ij}\right|^2 \quad (3)$$

where M and N are the size of the block $I_{ij}(x, y)$. The feature vector of a face is then constructed by concatenating the feature vectors of all blocks to one big feature vector V.

$$V = \cup_{i=1}^{k} \cup_{j=1}^{k_i} \{v_{ij}\} \quad (4)$$

where k is the number of subbands and k_i is the number of blocks in the i th subband. Therefore, the best features can be extracted with more dimensionality reduction of the feature vector.

III. Classification

In this paper, the nearest neighbor classifier with Euclidean distance and support vector machines are used for classification, SVM is implemented on both authentication and identification applications.

A. Euclidean Distance (EUD)

The Euclidean distance is used to measure the similarity between the test feature vector and the reference feature vectors in the gallery. It is defined as the straight-line distance between two points. For N-dimensional space, the Euclidean distance between two any points p_i and q_i is given by:

$$D = \sqrt{\sum_{i=1}^{N}(p_i - q_i)^2} \quad (5)$$

where p_i (or q_i) is the coordinate of p (or q) in dimension i.

In the application of this approach for face recognition, distances in the feature space from a query image to every image in the database are calculated. The index of the image which has the smallest distance with the image under test is considered to be the required index.

B. Support Vector Machines (SVMs)

SVM-based algorithm is one of the most useful techniques in classification problems. A Support Vector Machine (SVM) performs classification by constructing an N-dimensional hyperplane that optimally separates the data into two categories. The vectors near the hyperplane are support vectors. There are two schemes for SVMs multi-class classification. One is the one-against-one strategy to classify between each pair of classes. The other is the one-against-all strategy to classify between each class and all the remaining classes. The latter one is adopted for the proposed face recognition system.

The input to the SVM algorithm is a set $\{(x_i, y_i)\}$ of labeled training data, where x_i is the feature vector of an image (which is obtained by using the proposed method) and $y_i = 1$ or -1 is the label. The output of the SVM algorithm is a set of N_s support vectors s_i, coefficient weights α_i, class labels y_i of the support vectors and a constant term b. The hyperplane which is called the optimal separating hyperplane (OSH) is given as [12]:

$$\sum_{i=1}^{N_s} \alpha_i y_i s_i - b = 0 \quad (6)$$

SVM can be extended to nonlinear decision surfaces by using a kernel function $k(x_i, x_j)$ .that satisfies Mercer's condition [13]. The nonlinear decision surface is given as:

$$\sum_{i=1}^{N_s} \alpha_i y_i K(s_i, z) - b = 0 \quad (7)$$

Radial Basis Function (RBF) is a kernel function and is used in the one-against-all SVM adopted in proposed system. The RBF is given by

$$k(x_i, x_j) = exp(-\gamma \left\|x_i - x_j\right\|^2) \quad (8)$$

SVM-based algorithm is demonstrated on both verification and identification applications. In identification, the algorithm is presented with an image of an unknown person. The algorithm reports its best estimate of the identity of an unknown person from a database of known individuals. In a more general response, the algorithm will report a list of the most similar individuals in the database. There is a gallery $\{g_j\}$ of m known individuals. The algorithm is presented with a probe p to be identified. The first step of the identification algorithm computes a similarity score between the probe and each of the gallery images. The similar score between p and g_j is

$$\delta_j = \sum_{i=1}^{N_s} \alpha_i y_i K(s_i, g_j - p) - b \quad (9)$$

In the second step, the probe is identified as person j that has minimum similarity score δ_j. An alternative method of reporting identification results is to order the gallery by the similarity measure δ_j .

In verification, the algorithm is presented with an image and a claimed identity of the person. The algorithm either accepts or rejects the claim. There is a gallery $\{g_j\}$ of m known individuals. The algorithm is presented with a probe p and a claim to be person j in the gallery. The first step of the verification algorithm computes the similarity score between p and g_j is

$$\delta = \sum_{i=1}^{N_s} \alpha_i y_i K(s_i, g_j - p) - b \quad (10)$$

The second step accepts the claim if $\delta < T$. Otherwise, the claim is rejected. The classifier is designed to maximize the verification performance which is usually measured by two statistics the probability of correct verification, P_V, and the probability of false acceptance, P_F. The value of the threshold (T) is set to meet the desired tradeoff between P_V and P_F.

IV. Experimental Results

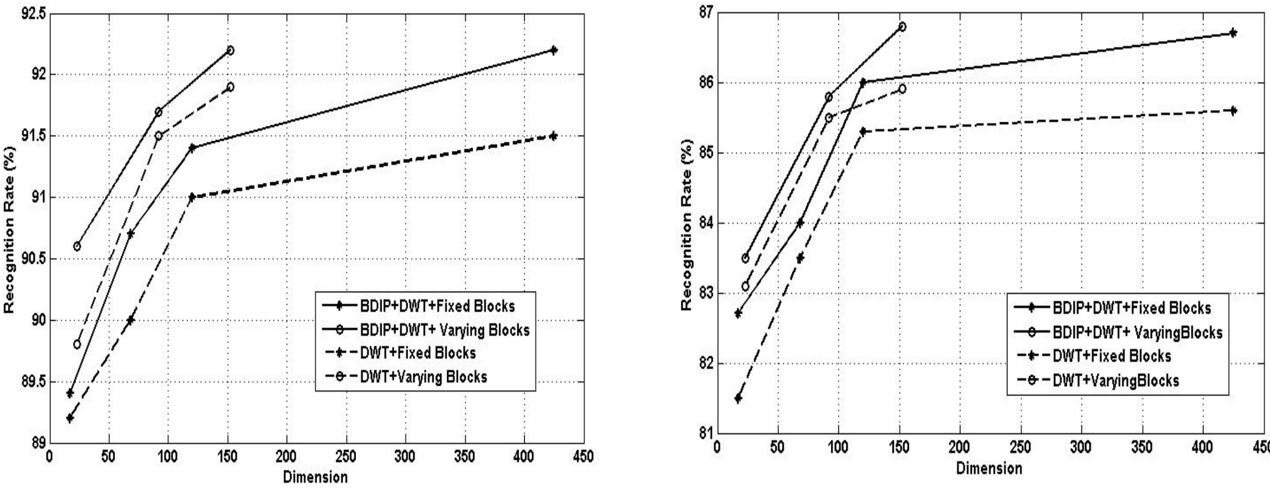
A. Database

The performance of the proposed feature extraction method is evaluated using two databases ORL and FERET. The ORL database contains images from 40 individuals, each has 10 different images. The FERET database is a popular database for testing and evaluating face recognition algorithms. In our experiments a subset of this database was used. It includes 100 individuals which were selected randomly such that each has 10 different images. The images in both databases vary in pose, illumination, facial expression and age. For both databases, four images per individual were chosen randomly for the training set and the remaining six images were used for the testing phase.

B. Results

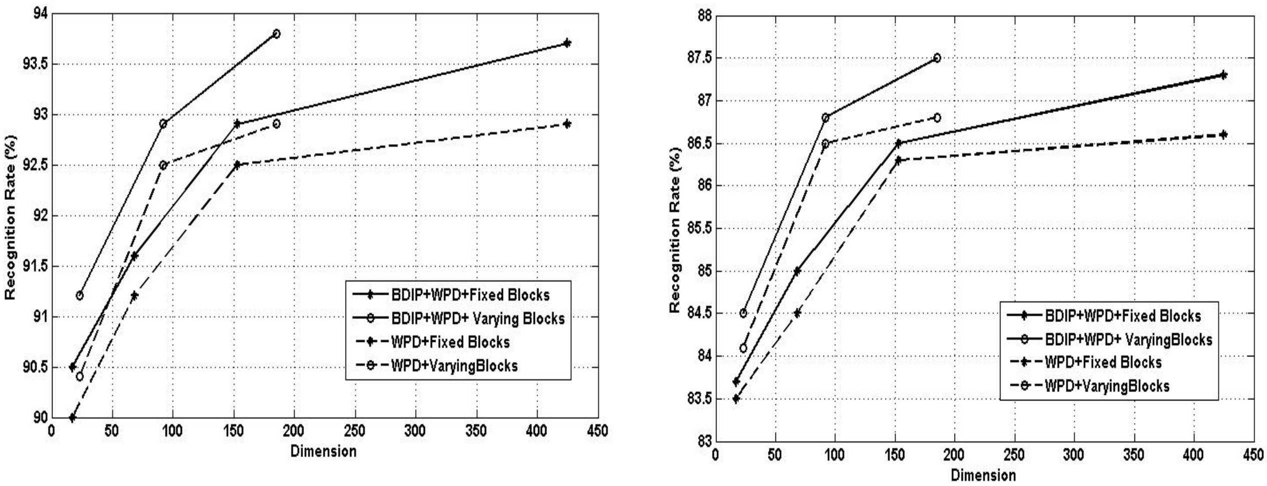
Two sets of experiments were carried out to investigate the performance of the proposed method. In our experiments, all images in the training and test sets were cropped to extract the facial region based on detection of the eyes and then resized to the 80×80 pixels. Furthermore, two classification techniques which are Euclidean distance (EUD) and SVMs based algorithms were used to evaluate the performance of the proposed face recognition system.

In the first set of experiments, the performance of the proposed feature extraction method based on BDIP and wavelet transform was compared to the wavelet-based methods (without using BDIP) such as DWT and WPD. With all methods, two-level of Haar wavelet decomposition was performed and the feature vectors were extracted using the block-based statistical measures described in Section II. Furthermore, the Euclidean distance based algorithm was used for classification. Figs. 4 and 5 show the average recognition rates versus the feature vector dimension for various feature extraction methods with fixed and varying blocks sizes. It can be seen that an improvement is gained when the BDIP is combined with wavelet transform. Moreover, all feature extraction methods with the block-based statistical measures (varying blocks size (VB)) achieves better recognition accuracy than those with the fixed blocks size (FB) with high dimensionality reduction for the feature vector. Furthermore, using the WPD with and without the BDIP exhibits better classification performance than using the DWT.



(a) (b)

Fig. 4. Average recognition rates versus the feature vector dimension for various feature extraction methods with and without BDIP and DWT. The EUD classifier is used. (a) ORL and (b) FERET databases.



(a) (b)

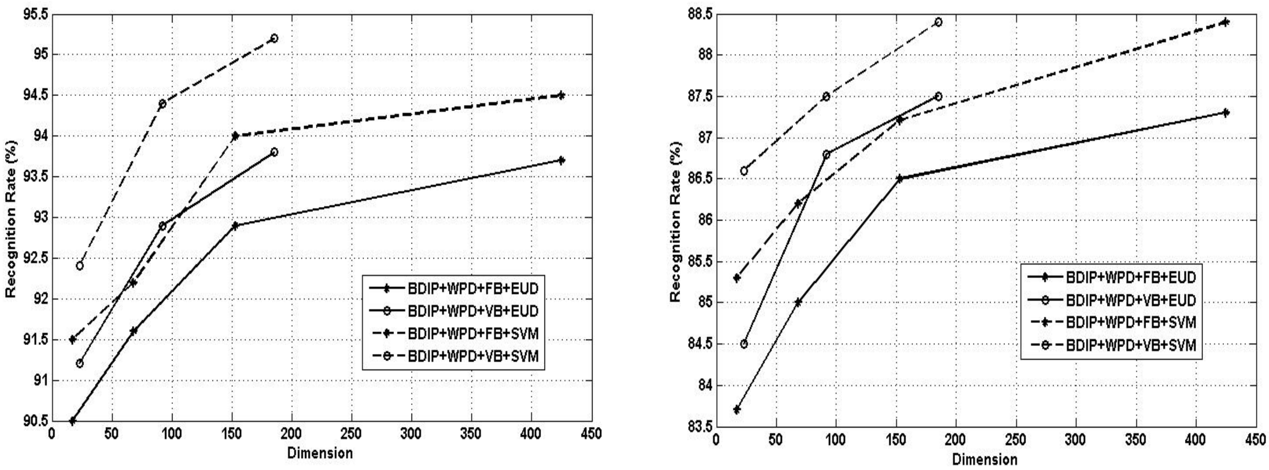
Fig. 5. Average recognition rates versus the feature vector dimension for various feature extraction methods with and without BDIP and WPD. The EUD classifier is used. (a) ORL and (b) FERET databases.

In order to further show the performance of the proposed method, comparisons were also carried out to the other methods such as PCA and DCT with the zigzag scan. The best recognition rates obtained using the various feature extraction methods are given in Table I. It can be seen that the proposed method gives the best results with high percentage of the dimension reduction.

TABLE I: THE BEST AVERAGE RECOGNITION RATES OF VARIOUS FEATURE EXTRACTION METHODS USING THE EUD CLASSIFER. THE NUMBER IN THE PARENTHESES IS THE FEATURE VECTOR DIMENSION.

Feature(s)	ORL	FERET
PCA	88.0% (150)	83.4% (400)
DCT with zigzag scan	91.0% (500)	86.5% (400)
WPD+FB	92.9% (425)	86.6% (425)
WPD+VB	92.9% (185)	86.8% (185)
WPD+BDIB+FB	93.7% (425)	87.3% (425)
WPD+BDIP+VB	93.8% (185)	87.5% (185)
WPD+FB	91.5% (425)	85.6% (425)
WPD+VB	91.9% (185)	85.9% (152)
WPD+BDIB+FB	92.2% (425)	86.7% (425)
WPD+BDIP+VB	92.2% (152)	86. 8% (152)

In the second set of experiments, the SVMs based algorithm was carried out to investigate the performance of the proposed method. It was also compared with the EUD classifier. It was implemented on both verification (authentication) and identification scenarios. For verification, the results were obtained under the Equal Error Rate (EER) by applying Receiver Operating Characteristic (ROC). For identification application using the EUD and SVMs classifiers, Fig. 6 shows the average recognition rates versus the feature vector dimension for the proposed feature extraction methods (BDIP+WPD+FB and BDIP+WPD+VB). The results show that the classification using SVMs is superior to the classification using the EUD with improvement rate about 0.95%−1.4% and 1.3% for (BDIP+WPD+FB) and (BDIP+WPD+VB) methods, respectively.



(a) (b)

Fig. 6. Average recognition rates versus the feature vector dimension for the proposed feature extraction method and using the EUD and SVM classifiers. (a) ORL and (b) FERET databases.

For verification (VR) application, the average recognition rates obtained using the proposed feature extraction method were compared to those obtained from identification (Id) application and are listed in Tables II and III. It is obvious that the proposed method with the SVMs implemented in the identification (Id) application performs better than that in the verification (VR) application. This is due to that the verification application has two types of errors which are the false acceptance and the false rejection.

TABLE II: AVERAGE RECOGNITION RATES OBTAINED USING THE (BDIP+WPD+FB) FEATURE EXTRACTION METHOD AND SVM IMPLEMENTED IN BOTH VERIFICATION (VR) AND IDENTIFICATION (ID) APPLICATIONS UNDER RBF KERNEL AND EQUAL ERROR RATE (EER).

Method	BDIP+WPD+FB			
Dimension	ORL		FERET	
	VR	Id	VR	Id
17	87.2%	91.5%	80.5%	84.5%
68	90.0%	92.2%	82.0%	85.5%
153	91.8%	94.0%	85.0%	87.3%
425	92.5%	94.5%	85.3%	87.6%

TABLE III: AVERAGE RECOGNITION RATES OBTAINED USING THE (BDIP+WPD+VB) FEATURE EXTRACTION METHOD AND SVM IMPLEMENTED IN BOTH VERIFICATION (VR) AND IDENTIFICATION (ID) APPLICATIONS UNDER RBF KERNEL AND EQUAL ERROR RATE (EER).

Method	BDIP+WPD+VB			
Dimension	ORL		FERET	
	VR	Id	VR	Id
23	89.6%	92.4%	82.0%	86.6%
92	91.1%	94.4%	84.5%	87.5%
185	92.8%	95.2%	85%	88.4%

V. CONCLUSION

This paper proposed a face recognition approach based on exploiting the BDIP features and the wavelet transform (WPD or DWT). A feature vector is constructed based on some block-based statistical measures such as the means and variances. In the block-based method, two types of partitions (FB or VB) are conducted to determine the optimal block size among the decomposition levels. The experimental results showed that the proposed feature extraction method, based on BDIP and wavelet transforms, outperforms the wavelet-based methods. All methods with varying blocks size statistical measures (VB) achieve better performance than those with the fixed blocks size (FB). Moreover, the proposed method outperforms the other well known methods such as the PCA and DCT with the zigzag scan. Furthermore, the proposed method achieves high recognition rates with a lower number of dimensions. The results also showed that the support vector machines (SVMs) are a better classification algorithm than the Euclidean distance (EUD) classifier for face recognition.

REFERENCES

[1] M. Turk and A. Pentland, "Face recognition using eigenfaces," *Computer Vision and Pattern Recognition*, pp. 586-591, Jun. 1991.

[2] H.K. Ekenel, and R. Stiefelhagen, "Local appearance based face recognition using discrete cosine transform", *EUSIPCO 2005*, *Antalya, Turkey*, 2005.

[3] J. T. Chien and C. C. Wu, "Discriminant wavelet faces and nearest feature classifiers for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1644–1649, 2002.

[4] A. Farag, R. Atta, and H. Mahdi, "Feature extraction based on the embedded zero-tree DCT for face recognition," in Proc. the 17th *IEEE International Conference on Image Processing* (*ICIP2010*), Hong Kong, pp. 3813-3816, September 26-29, 2010.

[5] R. Atta and M. Ghanbari, "Low-memory requirement and efficient face recognition system based on DCT pyramid," in *IEEE Trans. on Consumer Electronics*, vol. 56, issue 3, 2010.

[6] A. Farag, R. Atta, and H. Mahdi, "Face Recognition Based on the Statistical Features of BDIP and Wavelet Transform," *the 4th IEEE the International Conference on Computer Science and Information Technology (IEEE ICCSIT 2011)*, China, June 10-12, 2011.

[7] G.C. Feng, P.C. Yuen, and D.Q. Dai, "Human face recognition using PCA on wavelet subband," *SPIE Journal of Electronic Imaging*, vol. 9, no. 2, pp. 226–233, 2000.

[8] H. Wang, S. Z. Li, Y. Wang, and J. Zhang, "Self quotient image for face recognition," *in Proc. IEEE Int. Conf. Image Processing, Singapore*, vol. 2, pp. 1397-1400, Oct. 24-27, 2004.

[9] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under illumination using gradientfaces," *IEEE Trans. Image Processing*, vol. 18, no. 11, pp. 2599-2606, Nov. 2009.

[10] Y. Ju, H. J. So, N. C. Kim, and M. H. Kim, "Face recognition using local statistics of gradients and correlations," *in Proc. EUSIPCO2010, Aalborg, Denmark*, Aug. 23-27, 2010.

[11] M. El Aroussi, M. El Hassouni, S. Ghouzali, M. Rziza, D. Aboutajdine, "Novel face recognition approach based on steerable pyramid feature extraction," in Proc. the 16th *IEEE International Conference on Image Processing* (*ICIP2009*), Cairo, pp. 4165- 4168, Nov.7-10, 2009.

[12] Zhi-jie He, Lian-Wen Jin, "A new fast training algorithm for SVM", Machine Learning and Cybernetics, vol. 6, pp. 3451-3456, July 2008.

[13] Q. Tao, D. Chu, and J. Wang. "Recursive support vector machines for dimensionality reduction,"*IEEE Trans*. NN, 19(1):189–193, 2008.

Amany Farag received her B.Sc. with honors and M.Sc. degrees in Electrical Engineering from Suez Canal University, Port Said, Egypt, in 1998 and 2004, respectively. She received the Ph.D. degree in Electrical Engineering from Port Said University, Port Said, Egypt in 2011.
Currently, she is an assistant professor at computer science department, Port Said University. Her research interest is in the areas of pattern recognition, signal/image processing, and computer vision.

Randa Atta received the B.Sc. and M.Sc. degrees in Electrical Engineering from Suez Canal University, Port Said, Egypt, in 1991 and 1996, respectively. She received the Ph.D. degree in Electronic systems engineering from the University of Essex, England in 2004.
Currently, she is an Associate Professor at Port Said University. Her research interests are image/video coding and pattern recognition especially face recognition.