

Occlusion invariant face recognition using mean based weight matrix and support vector machine

G NIRMALA PRIYA^{1,*} and R S D WAHIDA BANU²

¹Department of Electronics and Communication Engineering,
Sona College of Technology, Salem 636 005, India

²Government College of Engineering, Salem 636 005, India
e-mail: nirmalaganeshs@yahoo.co.in

MS received 30 January 2013; revised 23 October 2013; accepted 14 November 2013

Abstract. In this paper, a novel occlusion invariant face recognition algorithm based on Mean based weight matrix (MBWM) technique is proposed. The proposed algorithm is composed of two phases—the occlusion detection phase and the MBWM based face recognition phase. A feature based approach is used to effectively detect partial occlusions for a given input face image. The input face image is first divided into a finite number of disjointed local patches, and features are extracted for each patch, and the occlusion present is detected. Features obtained from the corresponding occlusion-free patches of training images are used for face image recognition. The SVM classifier is used for occlusion detection for each patch. In the recognition phase, the MBWM bases of occlusion-free image patches are used for face recognition. Euclidean nearest neighbour rule is applied for the matching. GTAV face database that includes many occluded face images by sunglasses and hand are used for the experiment. The experimental results demonstrate that the proposed local patch-based occlusion detection technique works well and the MBWM based method shows superior performance to other conventional approaches.

Keywords. Face recognition; occlusion detection; MBWM; SVM; LBP; SLBM.

1. Introduction

The most important goal of computer vision is to achieve visual recognition equivalent to that of human recognition. And among many recognition subjects, the face recognition problem has been researched intensively during last few decades, due to its great potential to various practical applications such as HCI (Human Computer Interface), intelligent robot, surveillance, etc. A face image of a person gets changed due to many reasons. Different illumination conditions give different effects on the face image. Facial expression also changes the appearance

*For correspondence

of the face. Face in an image may have different poses (orientation) and of different sizes. It may be partially occluded. Extreme changes occur in the skin due to Aging. Among various problems associated with a face recognition system, handling partial occlusion is one of the most challenging problems. The problem of occlusion by other objects or apparels such as sunglasses, scarves or masks becomes eminent. Occluded parts in the face images usually degrade the recognition performance, and thus a robust algorithm for occluded faces is indispensable to real applications. Several approaches have been proposed in the literature for occlusion detection.

Leonardis & Bischof (2000) proposed a robust PCA approach in which the coefficients of Eigen images from partially degraded images are estimated and the coefficients are extracted by a robust hypothesize-and-test paradigm using subsets of image points. This technique presented successful reconstruction of partially occluded images, but the performance depends on the training set. Li *et al* (2001) proposed a local non-negative matrix factorization (LNMF) method for learning spatially localized, parts based subspace representation of visual patterns. In addition to the non-negativity constraint in the original NMF (1999), they imposed localization constraints to the objective function. The advantages of LNMF for occluded face recognition have been demonstrated experimentally compared with the NMF and PCA methods. Shang *et al* (2010) proposed an occlusion detection model in which the given face image is accomplished for each local patch independently by employing pattern classification framework. The high dimensional data set of the local patch is then reduced using PCA. Jia and Martinez (2009) proposed to use Partial Support Vector Machines (PSVM) in places where occlusions may occur. Gökberk *et al* (2004) investigated the effects of deception attacks (occlusions and camouflage) on two basic face recognition systems: a PCA based system and a Gabor wavelet-based recognizer. They proposed a distance called Asymmetric Trimmed Distance Measure that selects and combines only the regions where the similarities (between the probe and the gallery faces) are high. Martinez (2002) described a probabilistic approach that compensates the imprecisely localized, partially occluded, and expression-variant faces when only single training sample per class was available to the system. The occlusion problem is analysed by dividing the face image into k local regions and was analysed separately. The matching was done based on the probabilistic method that analysed better compared with the simple voting space. Martinez demonstrated experimentally that the suppression of $1/6$ th of the face does not decrease Recognition rate. Even when $1/3$ rd of the face is occluded, the identification results were close to those obtained in the occlusion-free case. Kim *et al* (2005) proposed a part-based local representation method called Locally Salient Independent Component Analysis (LS-ICA). Zhang *et al* (2007) proposed a method based on local Gabor Binary patterns. In this work, the face image is divided into rectangular region and, on the basis of their gray level histograms, probability of occlusion was estimated.

Tarres & Rama (2005) proposed a multiple PCA spaces for specific type of occluded faces using masks. But this approach as it deals with specific type of occlusion will not cope with wide variation of occlusions. This method also requires large memory and longer processing time. According to the experimental results, this description is quite robust with respect to various facial expression changes, varying illumination, and partial occlusions.

In this paper, a novel face recognition algorithm robust to occlusion using MBWM technique is proposed. The proposed algorithm is based on a local approach where face images are divided into a finite number of disjointed local patches. The occluded regions in the face images are detected by the SVM classifier. The recognition process is performed using the MBWM features corresponding to the occlusion-free patches. The algorithm is evaluated with the occluded

subset of the GTAV database (1999). The proposed algorithm produces superior performance to previous face recognition schemes.

2. Occlusion detection

The occlusion detection is based on the patch-based MBWM features and the SVM classifier. The features exhibit highly localized characteristics in spatial domain. So, the occlusion affects only the features of the corresponding region. Now, the error becomes local to the occluded region and not global. The MBWM features of occlusion-free regions are exclusively used for matching and face recognition. Thus, the occlusion detection step usually influences the overall performance of the face recognition system.

2.1 Local subdivision of a face image

In face images, Partial occlusions occur usually when subjects wear adornments like sunglasses or scarf, or when faces are covered with other objects such as hand, cup and so on, as shown in figure 1. To detect the locally occluded regions in a face image, the image is first divided into a finite number of local disjoint patches, and then each patch is individually examined. Generally, the configuration and the sizes of patches are important factors in overall recognition performance. An optimal division of face images was obtained empirically. In this paper, a symmetric 2 local patches division of a face image as in figure 2 is employed.

2.2 Local occlusion detection

Occlusion detection of a given face image is accomplished for each local patch independently using Support vector Machine. The local patches are still high-dimensional vector that is computationally infeasible. So, the statistical features like mean, standard deviation of the local patches are calculated and these features are used for finding the occluded region.

2.3 Occlusion detection using support vector machine (SVM)

As a powerful machine learning technique for data classification, SVM performs an implicit mapping of data into a higher (maybe infinite) dimensional feature space and then finds a linear separating hyperplane with the maximal margin to separate data in this higher dimensional space.



Figure 1. Examples of face images occluded with glass, hand, etc.



Figure 2. Subdivision on occluded face image.

Given a training set of labelled examples $\{(x_i, y_i), i = 1, 2, \dots, l\}$, where $x_i \in \mathbb{R}_n$ and $y_i \in \{-1, 1\}$ a new test example x is classified by the function as in Eq. (1):

$$f(x) = \text{sgn} \left(\sum_{i=1}^l \alpha_i y_i k(x_i, x) + b \right), \quad (1)$$

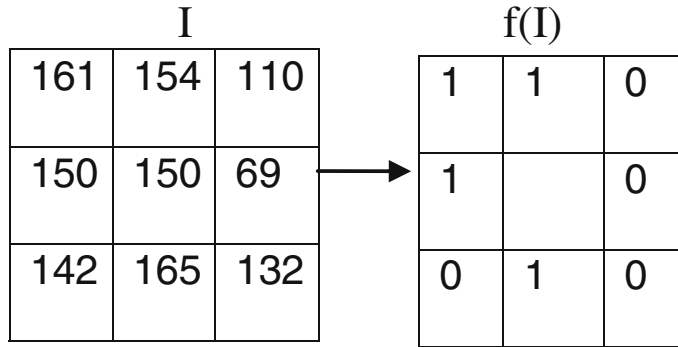
where, α_i is the Lagrange multiplier of a dual optimization problem that describes the separating hyperplane $K(x_i, x)$ denotes the kernel function and b is the threshold parameter of the hyperplane. The training sample x_i with $\alpha_i > 0$ is called support vectors and SVM finds the hyperplane that maximizes the distance between the support vectors and the hyperplane. Given a non-linear mapping Φ that embeds the input data into the high dimensional space, kernels have the form of $K(x_i, x_j) = (\Phi(x_i), \Phi(x_j))$. SVM allows domain-specific selection of the kernel function. Though new kernels are being proposed, the most frequently used kernel functions are the linear, polynomial and Radial Basis Function (RBF) kernels. SVM makes binary decisions. The effectiveness of SVM depends on the selection of kernel. The classification using linear and radial basis function (RBF) kernels are compared. The parameters are selected using cross validation and the parameters with best cross validation accuracy are used for classification. Training is done for the whole training set using the selected parameters. The generalization performances achieved by using two different kernels is discussed.

3. Face recognition

3.1 Local binary pattern (LBP)

LBP operator was originally designed for texture description (Ahoenon *et al* 2006) and it proved as a powerful means of texture description. The operator labels the pixels of an image by thresholding a 3×3 neighbourhood of each pixel with the center value and considering the results as a binary number. The extraction of Local Binary Pattern (LBPs) features is demonstrated in figure 3. The LBP features show good performance in the area of facial image description. The derived binary numbers (called Local Binary Patterns or LBP codes) represent the local primitives including different types of curved edges, spots, flat areas, etc. so each LBP code can be regarded as a micro-texton.

The basic LBP operator labels the pixels of an image (I_p) by thresholding each 3×3 pixel neighbourhood of the input image with the center pixel value (I_c), multiplying the threshold



$$LBP = (11000101)_2 = 2^0 + 2^1 + 2^5 + 2^7 = 163$$

Figure 3. LBP operator.

values by weight (powers of two) and summing them. Thresholding is done using the centre pixel as in Eq. (2):

$$f(I_P - I_C) = \begin{cases} 1 & I_P > I_C \\ 0 & \text{Otherwise} \end{cases} \quad (2)$$

Weights are then assigned and the LBP values are obtained using Eq. (3) and the values are summed to obtain the LBP values for the 3×3 matrix.

$$LBP = \sum_{p=0}^7 f(I_P - I_C) 2^p. \quad (3)$$

LBP's have been very effective for image representation as it is being applied to visual inspection, motion detection and outdoor scene analysis. The most important properties of LBP features are their tolerance against monotonic illumination changes and their computational simplicity. The LBP operator mainly detects many texture primitives and accumulates into a histogram over a region to capture local texture information. The statistical features as mean and standard deviation of the LBP features extracted are used for classification.

3.2 Simplified local binary mean (SLBM)

The simplified local binary mean (Priya & Banu 2012) overcomes the disadvantage of the LBP algorithm. The LBP produce long histograms with 256 bins, which slow down the speed. SLBM on the other hand, uses the mean of the 9 pixels for thresholding. This method involves three steps which include subdividing, thresholding and weighing. First, a 3×3 sub image is cropped. The pixel values are represented as I_p . Thresholding is done using the mean of the 9 elements of the 3×3 sub image (I_m). Thresholding is done based on the rule given in Eq. (4):

$$f(I_P - I_m) = \begin{cases} 1 & I_P > I_m \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

SLMB feature thus obtained are considered for classification purpose:

$$SLMB = \sum_{p=0}^7 f(I_P - I_m) 2^P. \quad (5)$$

The SLMB features are thus calculated. Many images of different types can have similar histograms. Histograms provide only a coarse characterization of an image. This is the main disadvantage of using histograms. So, the statistical features such as mean and standard deviation of the SLMB features are calculated.

3.3 Mean Based Weight Matrix (MBWM)

In LBP, thresholding is exactly at the value of the central pixel i_c . This makes it to be sensitive to noise, especially in near-uniform image regions. Many facial regions are relatively uniform. It is potentially useful to improve the robustness of the underlying descriptors in these areas. Also, in SLBM only two values are assigned and the weightage is not given to the mean value. So, the SLBM is extended to Mean Based Weight Matrix (MBWM) (Priya & Banu 2012). The mean based weight matrix involves three steps which include subdividing thresholding and weighing. All the pixels of the image are replaced by a 3-valued function as given in Eq. (6):

$$f(I_P - I_m) = \begin{cases} 2, & I_P > I_m \\ 1, & I_P = I_m \\ 0, & \text{Otherwise} \end{cases}. \quad (6)$$

Weights are then assigned and summed to obtain the MBWM values as in Eq. (7) for the 3×3 matrixes:

$$MBWM = \sum_{p=0}^7 f(I_P - I_m) 2^P. \quad (7)$$

The MBWM features are thus calculated. The first order statistical features such as mean and standard deviation of the MBWM features are calculated. These features are used for classification. The features are extracted for the patches 1 and 2 of the subdivided image and are used for the recognition of the face image.

3.4 Face recognition using MBWM features

Face recognition is performed using the MBWM features of the occlusion-free regions. Let $M = \{M_1, M_2, \dots\}$ and $S = \{S_1, S_2, \dots\}$ be the first order statistical features of the MBWM features of the original face image. For a given test face y , occlusion-free basis set associated with it is determined and denoted as $M_t = \{M_1, M_2, \dots\}$ and $S_t = \{S_1, S_2, \dots\}$. Similarly, each training face image x_i ($i = 1, 2, \dots, K$), where K is the total number of training faces, is projected into the same selected occlusion-free MBWM features with $M_T = \{M_1, M_2, \dots\}$ and $S_T = \{S_1, S_2, \dots\}$. Then, the recognition is performed by finding the closest training face in the feature space as follows.

$$\arg \min (||M_{Ti} - M_{ti}|| \&\& ||S_{Ti} - S_{ti}||), \quad (8)$$

where $i = 1, 2, \dots, K$.



Figure 4. GTAV face database.

4. Experimental results

4.1 The GTAV -face database

GTAV face database is used for the testing. This database has been created for testing the robustness of face recognition algorithms against strong pose, occlusion and illumination variations. This database includes a total of 44 persons with 27 pictures per person which corresponds to different pose views (0° , $\pm 0^\circ$, $\pm 45^\circ$, $\pm 60^\circ$ and $\pm 90^\circ$) under three different illuminations (environment or natural light, strong light source from an angle of 45° , and finally an almost frontal mid-strong light source). Furthermore, at least 10 more additional frontal view pictures are included with different occlusions and facial expression variations. The resolution of the images



Figure 5. Synthetic occluded face images.



Figure 6. Disjoint patches of occluded and non-occluded face image.

are 240×320 and they are in BMP format. A sample of the images from GTAV face database is shown in figure 4.

The frontal face images of all the 24 persons were used for training. For the occlusion classification, all 24 frontal face images and 48 occluded face images (24 sunglass images and 24 hand occluded images of both men and women) were used for the training the targetclass.

4.2 Synthetic occluded-face database

A face images occluded by medical mask in MIT face database is used for testing the algorithm. This database contains images of 40 distinct subjects with eleven different poses for each individual. A few additional images are also included for some individuals. All the images have a bright homogeneous background and the subjects are in an upright, frontal position. For each individual, the following pose for the face: looking front, looking left, looking right, looking up, and looking up towards left, and looking up towards right, looking down are included. In addition to the variation in pose, images with four emotions—neutral, smile, laughter, sad/disgust are also included for every individual. Figure 5 shows the synthetic occluded face database.

4.3 Occlusion detection results

Figure 6 shows the example of test occluded and non-occluded faces. Each training normal face image is divided into 2 disjoint patches and the corresponding features of the 2 regions are learned. Figure 6 shows the example of test occluded faces, in which patches 2 (lower half of

Table 1. Detection rate definition.

Definition of terms	Input image	Detection
Occlusion detection rate	Occluded	Occluded
False rejection rate	Occluded	Non-occluded
Rejection rate	Non-occluded	Non-occluded
False alarm rate	Non-occluded	Occluded

Table 2. Features of the non-occluded database for 2 segments.

Image	Segment 1		Segment 2	
	Mean	SD	Mean	SD
1	205.91	71.97	148.66	104.82
2	245.91	21.21	126.25	92.82
3	224.91	53.08	159.16	103.08
4	231.00	53.79	129.25	98.75
5	163.50	105.03	150.00	99.95
6	226.00	53.52	167.00	87.09
7	226.50	53.86	162.16	98.84
8	182.91	99.36	153.25	102.89
9	210.58	68.89	211.25	57.38
10	215.16	70.87	98.50	84.86

the face) is occluded. To evaluate the detection performance quantitatively, decision rates are summarized as definitions in table 1.

The detection of occluded image as occluded image is the right decision on the occluded data. This case is defined as detection rate as maximizing this rate is the goal of the occlusion detection. The wrong decision on normal data is defined as false alarm. The other cases are derived from detection rate and false alarm rate. The detection rate has trade-off relations with false alarm rate, so increase of only detection rate or decrease only false alarm rate is not possible. But, the misclassification of normal data as occluded can cause information loss for correct recognition, and also the misclassification of occluded data as non-occluded can lower the recognition performance seriously. Of two cases of misjudgment, false rejection rate is much more important than the false alarm rate for the performance of face recognition. Hence, as the performances of the classifiers, the false alarm rate is evaluated when the detection rate is 100%. The features of the occluded and non-occluded face image are subdivided into two equal halves with the upper segment focusing on the eyes and the lower segment focusing on the lips for the non-occluded image and the medical mask for the occluded image for the MBWM algorithm are presented in tables 2 and 3.

Table 3. Features of the occluded database for 2 segments.

Image	Segment 1		Segment 2	
	Mean	SD	Mean	SD
1	226.41	53.90	123.25	90.19
2	237.08	26.46	157.66	99.05
3	215.00	69.95	160.00	96.54
4	215.66	9.74	151.33	107.13
5	211.66	68.30	148.33	105.60
6	246.25	20.43	163.58	89.07
7	211.08	68.72	169.75	109.13
8	222.00	53.51	96.66	79.46
9	196.91	68.17	164.83	89.19
10	205.16	80.77	112.16	99.83

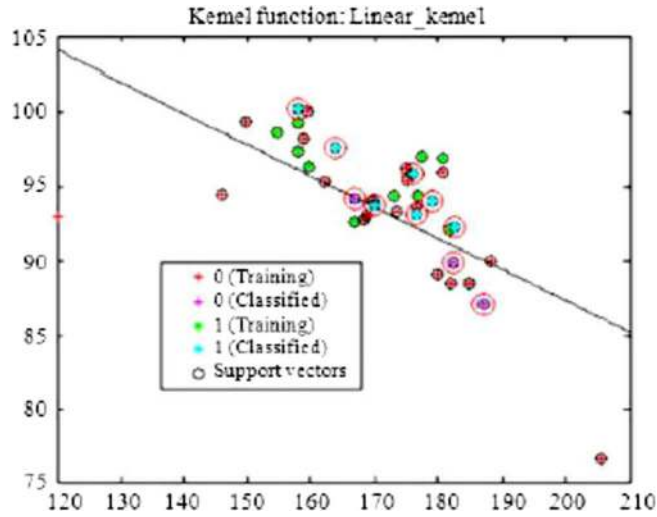


Figure 7. Linear SVM classified output for MBWM features.

The classified output using linear and RBF SVM classifier for the first order statistical features, mean and standard deviation are shown in figures 7 and 8. The + sign indicates test image classified as occluded face image and * sign indicates test image classified as non-occluded face image. The + sign indicates the occluded training face image and the * sign indicates the non-occluded training face image. The 'o' indicates the support vectors selected for the hyperplane. The support vector is generated using the equation (1) with the train and the test face image.

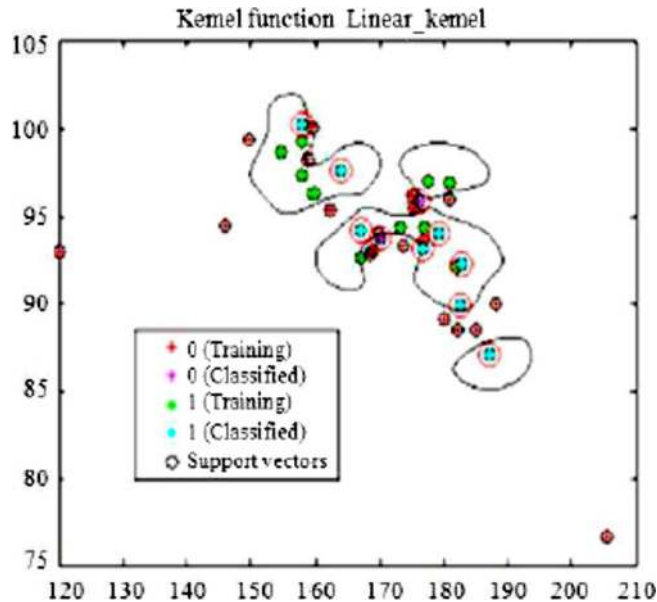


Figure 8. RBF SVM classified output for MBWM features.

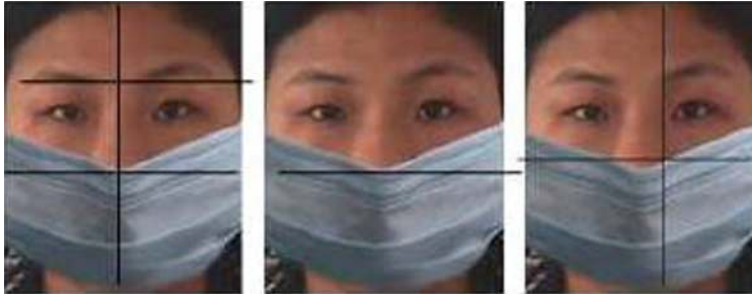


Figure 9. Subdivision of face image.

A better occlusion detection performance is obtained using the RBF Kernel and the MBWM features.

4.4 Subdivision of face images

Partial occlusion detection and face recognition algorithms are developed based on local patches of a face image. Therefore, different division methods may result in different performances on both occlusion detection and recognition. So, the optimal subdivision method of face images in an empirical sense is examined. Figure 9 shows 3 possible subdivision layouts considered for this experiment.

The SVM classifier was used for the comparison of the occlusion detection performances. And the detection rate for each subdivision method was calculated for the performance evaluation. The detection results of the 6-region-division, 4 region subdivision and 2 region subdivision methods are summarized in table 4.

As the results show, the 2 region division gives the better result compared with 4 and 6 region subdivision methods. This is because the 2 region division may contain both the eyes which will give the variation in the features calculated between the persons. When 4 region subdivision is considered, the features may be calculated for either one eye or a portion of the mouth which will not be a useful feature.

4.5 Face recognition results

100 Occlusion-free training face images from the GTAV face database have been trained. Mean and Standard deviation are the two statistical features of the MBWM features that are used for the face recognition. A threshold value is fixed for the finding the correct match. The threshold values are selected in a way such that the deviation between the original and occluded face image is minimum.

Table 4. Efficiency of overlapping RBF SVM methods.

Segments	Methods		
	LBP	SLBM	MBWM
2	94.5	96.5	98.75
4	93.0	94.0	95.45
6	90.0	93.5	94.50

Table 5. Recognition rate (%) on synthetic occlusion.

Methods	Medical mask	Upper mask	Lower mask
LBP	94	92	94.25
SLBM	96.25	94	96.75
MBWM	98.75	96.25	98.75

4.5a Experiments on synthetic occlusions: Initially, the face recognition algorithm was tested on synthetically occluded images. Occlusion-free images were used for training. The features of LBP and SLBM algorithm were also tested for the comparative performance evaluation. The recognition rate, defined by the percentage of correctly recognized faces, is used as the performance measure. Table 5 shows the recognition results. Experimental results show that the proposed algorithm achieved the highest recognition rate compared with LBP and SLBM algorithms.

4.5b Experiments on real occlusions: The face images occluded by sunglass and hand in GTAV database are used for testing. Among 72 occluded images (24 sunglass, 24 hand occluded images and 24 by both), 24 were used for training and the remaining images that were not included in the training set were used for probes, and remaining normal frontal faces were used for the gallery. Thus, in this experiment, the performance of the algorithm was investigated on the occluded face images exclusively unlike the synthetic occlusion test case where the occlusion-free parts of the gallery and the corresponding test images are exactly the same, in this case, those parts may differ since they are taken in different conditions. The performances of LBP, SLBM and MBWM were also evaluated and compared, and the results are summarized in table 6.

By comparing the results, we can conclude that the proposed algorithm is more robust than other algorithms especially for the occluded faces including sunglass and hand images. The hand should degrade the performance worse than sunglass as the area occluded by the hand in a face is larger than that by the sunglass, and most appearance-based approaches were affected by this fact. But, interestingly in the proposed method hand occluded image produce better results than sunglass. This is due to the fact that the eye region contains more important discriminative information than the mouth region for recognition.

In this experiment, only one occlusion-free face image per person is used for training. Moreover, although the face without occlusion was contained in the training set, the corresponding occlusion-free parts in the test images were different due to possible pose and expression variations, and localization error. The proposed algorithm produces better recognition results than the existing methods, as it minimizes recognition error by finding local occlusion areas explicitly, and uses only the local bases associated with occlusion-free parts. The processing time of the proposed algorithm is also less compared to the other existing algorithms.

Table 6. Recognition rate (%) on real occlusion.

Methods	Sunglass	Hand
LBP	92	94.25
SLBM	93.75	95.50
MBWM	96.50	98.25

5. Conclusion

In this paper, the occlusion problem which has been researched relatively less in face recognition than the illumination, expression and pose variation problems is dealt with. A new robust face recognition algorithm for the partial occlusion is proposed based on MBWM features. Local occluded areas in faces are first detected by SVM classifier and then matching is performed using the MBWM features of the occlusion-free regions. Experimental results demonstrated that the proposed algorithm could reliably recognize partially occluded faces with higher recognition rate than the existing methods.

References

- Ahonen T, Hadid A and Pietikainen M 2006 Face description with Local binary patterns: Application to face recognition. *IEEE Trans. Patt. Anal. Mach. Intell.* 28: 2037–2041
- Gökberk B, Akarun L and Aksan B 2004 How to deceive a face recognizer. In: *Proceedings of the Biometrics: Challenges arising from Theory to Practice Workshop*
- Jia H and Martinez A M 2009 Support Vector Machines in face recognition with occlusions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 20–25, IEEE Xplore Press, Miami, FL, pp 136–141
- Kim J, Choi J, Yi J and Turk M 2005 Effective representation using ica for face recognition robust to local distortion and partial occlusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 27(12): 1977–1981
- Leonardis A and Bischof H 2000 Robust recognition using eigenimages. *Comput. Vis. Image Understanding* 78: 99–118
- Li S Z, Hou X W, Zhang H J and Cheng Q S 2001 Learning spatially localized, part-based representation. *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.* 1:207–212
- Martinez A M 2002 Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. *IEEE Trans. Pattern Anal. Mach. Intel.* 24(6): 748–763
- Priya G N and Banu R S D W 2012 A Simplified Local Binary Mean (SLBM) based human gender classification. *Eur. J. Sci. Res.* 71: 435–442
- Shang L, Zhou C, Gu Y and Zhang Y 2010 Face recognition using the feature fusion technique based on LNMF and NNSC algorithms. *Adv. Intell. Comput. Theories Appl.* 6215: 547–554
- Tarres F and Rama A 2005 A novel method for face recognition under partial occlusion or facial expression variations. In: *Proc. 47th International Symposium ELMAR-2005, Multimedia Systems and Applications*
- Zhang W, Shan S, Chen X and Gao W 2007 Local Gabor binary patterns based on Kullback Óleibler divergence for partially occluded face recognition. *IEEE Signal Process. Lett.* 14(11): 875–878