

A Robust Method for Partially Occluded Face Recognition

Wenkai Xu¹, Suk-Hwan Lee², Eung-Joo Lee³

¹ Department of Information Communications and Engineering, University of Tongmyong
Busan, 608-711 Korea
[e-mail: xwk6298@hotmail.com]

² Department of Information Security, University of Tongmyong
Busan, 608-711 Korea
[e-mail: skylee@tu.ac.kr]

³ Department of Information Communications and Engineering, University of Tongmyong
Busan, 608-711 Korea
[e-mail: ejlee@tu.ac.kr]

*Corresponding author: Eung-Joo Lee

*Received April 9, 2015; revised June 3, 2015; revised June 9, 2015; accepted June 16, 2015;
published July 31, 2015*

Abstract

Due to the wide application of face recognition (FR) in information security, surveillance, access control and others, it has received significantly increased attention from both the academic and industrial communities during the past several decades. However, partial face occlusion is one of the most challenging problems in face recognition issue. In this paper, a novel method based on linear regression-based classification (LRC) algorithm is proposed to address this problem. After all images are downsampled and divided into several blocks, we exploit the evaluator of each block to determine the clear blocks of the test face image by using linear regression technique. Then, the remained uncontaminated blocks are utilized to partial occluded face recognition issue. Furthermore, an improved Distance-based Evidence Fusion approach is proposed to decide in favor of the class with average value of corresponding minimum distance. Since this occlusion removing process uses a simple linear regression approach, the completely computational cost approximately equals to LRC and much lower than sparse representation-based classification (SRC) and extended-SRC (eSRC). Based on the experimental results on both AR face database and extended Yale B face database, it demonstrates the effectiveness of the proposed method on issue of partial occluded face recognition and the performance is satisfactory. Through the comparison with the conventional methods (eigenface+NN, fisherfaces+NN) and the state-of-the-art methods (LRC, SRC and eSRC), the proposed method shows better performance and robustness.

Keywords: Face recognition, partial occlusion, sub-images, LRC

This research was supported by the Busan Metropolitan City, Korea, under the 2015 BB21 program grants and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2011-0023118), and the ICT R&D program of MSIP/IITP [R0126-15-1112, Development of Media Application Framework based on Multi-modality which enables Personal Media Reconstruction].

1. Introduction

Over the last decades, face recognition has emerged as an active research area in computer vision with numerous potential applications including video-mediated communication biometrics, human-computer interaction, surveillance and content-based access of images and video databases. Moreover, systems that rely on face recognition have gained great importance ever since terrorist threats imposed weakness among the implemented security systems.

Various face recognition algorithms have been developed and used as an application of access control and surveillance. On the analysis of face recognition system, some of the important studies on face recognition systems are discussed. Principal Component Analysis (PCA) is also known as Eigen face method, in PCA method the images are projected onto the facial value so called eigenspace [1, 2]. PCA approach reduces the dimension of the data by means of basic data compression method [3] and reveals the most effective low dimensional structure of facial patterns [4]. Local Feature Analysis (LFA) method of recognition is based on the analysis of the face in terms of local features e.g. eye, nose etc. by what is referred LFA kernels. References [5-7] is based on learning of the faces in an “Example Set” by the machine in the “Training Phase” and carrying out recognition in the “Generalization Phase”. The support vectors consist of a small subset of training data extracted by the algorithm given in [8]. This has resulted in the development of successful algorithms and the introduction of commercial products.

However, it remains still an unsolved problem which presents only successful results under controlled situations where the limited conditions like illumination, pose, facial expressions, and especially in the case of partial occlusion, the performance of recognition is abruptly degraded under these uncontrolled situations. Face recognition under controlled environments has been in scene for the past decades but recognition under uncontrolled situations like illumination, expression and partial occlusion is a recent issue. Partial face occlusion is one of the most challenging problems in face recognition. A face recognition system can confront occluded faces in real world applications very due to use of accessories, such as scarf or sunglasses, hands on the face, the objects that persons carry, and external sources that partially occlude the camera view, samples of partially occluded face image from AR face database are shown in Fig. 5. Therefore, the face recognition system has to be robust to occlusion in order to guarantee reliable real-world operation [9].

Numerous researches have been conducted in order to address this challenge [9-15]. In reference [9], the face image is first divided into k local regions and for each region an eigenspace is constructed. If a region is occluded, it is automatically detected. Moreover, weighting of the local regions are also proposed in order to provide robustness against expression variations. M. Aleix proposed a similar approach in [10], where a self-organizing map (SOM) is used to model the subspace. Face attributed relational graph (ARG) represented the face in [12], this representation contains a set of nodes and binary relations between these nodes. In [13] robustness against occlusion is provided by combining the subspace methods that aim at best reconstruction, such as principal component analysis, with the subspace methods that aim at discrimination, such as linear discriminant analysis. A sparse signal representation is used to analyze partially occluded face images in [14]. Another representation based approach is proposed in [15]. Different from the studies [9-14] in which occluded images are only included in the testing set, they are included both in the training and

testing sets in [15]. Overall, a major drawback of these methods is that they have a high computational cost in terms of training and testing the data.

To address the challenge of partial occluded face recognition, a novel method based on linear regression-based classification (LRC) algorithm is proposed in this paper. First, all images are downsampled and divided into several blocks, and then we exploit the evaluator of each block to determine the clear blocks of the test face image by using linear regression technique. Next, the remained uncontaminated blocks are utilized to partial occluded face recognition issue. Furthermore, an improved Distance-based Evidence Fusion approach is proposed to decide in favor of the class with average value of corresponding minimum distance.

The rest of this paper is organized as follow: In Section 2, we review the previous work in face recognition and partially occluded face recognition, especially. Section 3 focuses on determining the clear blocks of the test face image using B-LRC and recognizing face by using i-DEF algorithm. Section 4 presents experimental evaluation of classification accuracy using proposed method compared to other methods. Section 5 gives the conclusion for this study.

2. Related Work

Generally speaking, face recognition methods whether under controlled conditions or uncontrolled conditions can be classified into three groups i.e., feature based method that deal with features like eyes, mouth, nose and establish a geometrical correspondence between them [16]. The second category is appearance-based method that focus on the holistic features of face images by considering the whole face region. Another one deals with the hybrid local and global features of face images to be used for recognition purpose.

Nevertheless, It is difficult for appearance-based face recognition techniques based on the conventional component analysis techniques, such as principal component analysis (PCA) [1, 2], linear discriminate analysis (LDA) [4], and independent component analysis (ICA) [23] due to partial occlusions caused by sunglasses, scarf, hair or other factors, Basic methods used up to now for handling occlusion belongs to one of the following methods: fractal based method, feature based method and part based method.

For fractal based method, Marsico et al. presented Partitioned Iterated Function System (PIFS) based face recognition algorithm in [17]. This method computes the self-similarities among the image and establishes a relationship between the squares of grid region. But PIFS algorithm is sensitive to occlusions, so to avoid this individual component of face, parts are handled locally like eyes, mouth and nose. Further distortions are removed by ad-hoc distance measure.

Global and local features play an important role in occluded face recognition. The conventional methods used global features from one kernel [18, 19]. These features can be affected by noise or occlusions which reduces the robustness of these methods. The recognition techniques have also been used effectively using local features [15, 18, 20, 21, 22]. The partial occlusion affects the local features but the recognition methods can be made robust if these local features are merged together intelligently.

Recently, sparse representation technique has drawn wide attention in solving computer vision issues [24]. Wright et al. proposed the sparse representation-based classification (SRC) method for robust face recognition [25]. SRC uses the entire training set to represent a test sample with the coding vector subject to sparse constraint. The core of SRC is a l_1 -minimization problem (minimizing the l_1 -norm problem), which is approximately equal to

the related l_0 -minimization problem (minimizing the l_0 -norm problem) under some conditions. Although SRC can achieve very promising recognition results, the computational cost of solving the l_1 -minimization problem is quite high. Especially, when dealing with random corruptions and occlusions, an augmented sample matrix (ASM) is used to represent the test sample, which is referred to as the extended SRC (eSRC) method. ASM is very large compared with the original sample matrix, which greatly increases the scale of the optimization problem of eSRC and results in very heavy computational load. Another important relevant research called Linear regression-based classification (LRC) is proposed for robust face recognition [26]. LRC represents a test image as a linear combination of training images of class-specific samples. Therefore, face recognition can be cast as a problem of linear regression. Compared with SRC, LRC is much more time-saving to implement because it only uses least-squares estimation in computation while SRC needs iterative procedures to solve the l_1 -minimization. However, for LRC which essentially can be categorized as a nearest subspace [26], it estimates a given probe against class-specific models in the least-squares sense, and insufficient training samples will also make the estimation incorrect.

3. The Methodology of Partial Occluded Face Recognition

In this section, we first briefly describe the LRC algorithm and analyze that the error components derived from LRC are more suitable for identifying occlusion parts than original test face image, which motivates the work of this paper in section 3.2. The main steps of our method are as follows:

1. Downsampled and divide the face image into M blocks uniformly;
2. Construct the linear model to represent the input image based on LRC;
3. Exploit the evaluator of each block to determine the clear blocks of the test face image using B-LRC;
4. An improved Distance-based Evidence Fusion (i-DEF) method is proposed to final face recognition issue on the selected blocks.

3.1 Introduction of Linear Regression-based classification

Let there be N number of distinguished classes with p_i number of training images from the i -th class, $i = 1, 2, \dots, N$. Each gray-scale training images is of an order $a \times b$ and is represented as $u_i^{(m)} \in R^{a \times b}$, $i = 1, 2, \dots, N$ and $m = 1, 2, \dots, p_i$. Each gallery image is downsampled to an order $c \times d$ and transformed to vector through column concatenation such that $u_i^{(m)} \in R^{a \times b} \rightarrow w_i^{(m)} \in R^{q \times 1}$, where $q = cd$, $cd < ab$. Each image vector is normalized so that maximum pixel value is 1. Using the concept that patterns from the same class lie on a linear subspace [27], a class-specific model [26] X_i by stacking the q -dimensional image vectors as $X_i = [w_i^{(1)}, w_i^{(2)}, \dots, w_i^{(p_i)}] \in R^{q \times p_i}$, $i = 1, 2, \dots, N$. Each vector $w_i^{(m)}$ spans a subspace of IR^q also called the column space of X_i . Therefore, at the training level, each class i is represented by a vector subspace, X_i , which is also called the regressor or predictor for class i . Let z be an unlabeled test image and our problem is to classify z as one of the class $i = 1, 2, \dots, N$. The author transformed and normalized the gray-scale image z to an image vector $y \in R^{q \times 1}$ as discussed for the gallery. If y belongs to the i -th class, it should be represented as a linear combination of the training images from the same class (lying in the same subspace).

$$y = X_i \alpha_i, i = 1, 2, \dots, N \quad (1)$$

Where $\alpha_i \in IR^{p_i \times 1}$ is the vector of parameters. Given that $q \geq p_i$, the system of equations in (1) is well conditioned and α_i can be estimated using least-squares estimation [26],

$$\hat{\alpha}_i = (X_i^T X_i)^{-1} X_i^T y \quad (2)$$

The estimated vector of parameters, $\hat{\alpha}_i$, along with the predictors X_i are used to predict the response vector for each class i ,

$$\hat{y}_i = X_i \hat{\alpha}_i = X_i (X_i^T X_i)^{-1} X_i^T y = H_i y, i = 1, 2, \dots, N \quad (3)$$

Where the predicted vector $\hat{y}_i \in IR^{q \times 1}$ is the projection of y onto the i -th subspace. In another word, \hat{y}_i is the closest vector, in the i -th subspace, to the observation vector y in the euclidean sense. H is called hat matrix since it maps y into \hat{y}_i . The distance measure between the predicted response vector \hat{y}_i and the original response vector y is calculated as

$$d_i(y) = \|y - \hat{y}_i\|_2, i = 1, 2, \dots, N \quad (4)$$

And rule in favor of the class with minimum distance, i.e.,

$$\min_i d_i(y), i = 1, 2, 3 \dots, N \quad (5)$$

3.2 Detection of Occluded Region Based on Block-LRC (B-LRC)

A partial occluded part in face images usually occurs due to sunglasses, scarves, mustaches and so on. Such variations in facial appearance are commonly encountered in uncontrolled situations and may cause big trouble to the face recognition system. The challenge of partial occluded face recognition could be efficiently dealt with using the block (modular) representation approach [26]. Contiguous occlusion can safely be assumed local in nature in a sense that it corrupts only a portion of conterminous pixels of the image, the amount of contamination being unknown.

This is to say, the occlusion is not scattered over the images so that if we divide the face images into several blocks called sub-images, some of them remain clear and support distinctive information. Thus, local approaches try to extract meaningful partial facial features that can eliminate or compensate the difficulties brought by occlusion variations. Some previous algorithms [25, 26] processed each block individually and made the final decision by fusing each separate recognition result using majority voting or other principle. However, a major pitiful with majority voting is that it treats noisy and clean partitions equally. In this way, it is likely to be erroneous no matter how significant the clean partition may be in the context of facial features. To address this issue, a modular processing that a simple but effective approach is presented for handling the occlusion and corruption problem in this paper. The idea is to segment the face image into M non-overlap blocks uniformly ($M=8$ in this study [25, 26]). Then an occluded portion detection method based on block-linear regression classification (B-LRC) is proposed for removing the occluded blocks.

For partial occluded face recognition scenario, we assume there are N classes and p_i training samples in the i -th class, $i = 1, 2, \dots, N$. We transform images into vector $X_{i,j}^k$ ($k = 1, 2, \dots, M$, $j = 1, 2, \dots, p_i$), the vector of the k -th block from the i -th class can be described as

$$X_{i,j}^k = [x_{i,1}^k, x_{i,2}^k, \dots, x_{i,p_i}^k] \quad (6)$$

Based on the analysis of LRC, we still assume that the sub-images of the same block belonging to a subject lie on a certain subspace of the corresponding portion. Thus, given an input face image y belonging to the i -th class, the following equation is approximately reasonable.

$$y = \sum_{k=1}^M X_i^k \alpha_i^k \quad (7)$$

To a specified block (sub-image) y^k of the input test image

$$y^k = X_i^k \alpha_i^k \quad (8)$$

As the fundamental assumption of LRC that samples from a specific object class lie on a linear combination of training images of class-specific samples, namely, LRC represents a test image as a linear combination of training images of class-specific. In other words, if the k -th block of the test image y^k is contaminated by occlusion or corruption, it will not be represented as a linear combination of training images of class-specific. To evaluate the block based on the above analysis, the distance measure between the predicted response vector \widehat{y}^k and the original response vector y is calculated as

$$d_i(y^k) = \|y^k - \widehat{y}^k\|_2 \quad (9)$$

where

$$\widehat{y}^k = X_i^k \widehat{\alpha}_i^k = X_i^k (X_i^{kT} X_i^k)^{-1} X_i^{kT} y = H_i^k y^k \quad (10)$$

For estimating which portion is contaminated by occlusion, an evaluator is defined as

$$E(y^k) = [Ave_d(y^k) - \min_i d_i(y^k)] / Ave_d(y^k) \quad (11)$$

Here $\min_i d_i(y^k)$ indicates a corresponding minimum distance, and

$$Ave_d(y^k) = (\sum_{i=1}^N d_i^k) / N \quad (12)$$

In the case of a block without any occlusion, y^k , can be represented very well by one linear combination from its corresponding class of training sample X_i^k while models of other classes can not offer accurate representations. Therefore, the true class of the test sample could express the block image much better than the average of all the class, which indicates that this block has a big value of $E(y^k)$. While there are some blocks contaminated by occlusion, no class, generally, can accurately express this non-facial subimage or all classes offer approximately equally poor representation of this block, which means this block has a small

$E(y^k)$. In summary, a block with big $E(y^k)$ means it is very likely that block of the test sample is clear, while small $E(y^k)$ indicates this block is occluded.

Thus, the r ($r \leq M$) effective blocks with r largest $E(y^k)$, which clear correspondingly or without any occlusion are selected as

$$S(r) = \max_{(1 \sim r)} [E(y^1), E(y^2), \dots, E(y^M)] \quad (13)$$

So far we have detected which block should be removed and the remain r blocks ($r=4$ blocks are selected in this paper based on experience) would be utilized for face recognition in next section.

3.3 Face Recognition Using Improved Distance-based Evidence Fusion (i-DEF)

In reference [25], authors made use of the specific nature of distance classification to develop a fairly simple but efficient fusion strategy which implicitly deemphasizes corrupted subimages, significantly improving the overall classification accuracy. They proposed using the distance metric as evidence of their belief in the “goodness” of intermediate decisions taken on the subimages, which called “Distance-based Evidence Fusion” (DEF).

Inspired by aforementioned work, an improved Distance-based Evidence Fusion (i-DEF) method is proposed to final face recognition issue in this study. Unlike DEF method, we have approximately eliminate the effect of occlusion by removing the contaminated portions based on analysis above in order that the recognition can well proceed on the remaining parts of the image. So only the DEF on the remaining blocks should be calculated rather than all of that in the image.

Furthermore, as analyzed in reference [25], l_2 -norm reflects the overall difference between two vectors, and it is sensitive to large errors. On the other hand, we consider that if the occluded portion in the test image y^k is correctly removed, it can be well-approximated as a linear combination of $X_{i,j}^k$ of the real subject i , and the error between y^k and $X_i^k \alpha_i^k$ should be sparse. This error may affect any part of the image and may be arbitrarily large in magnitude. It is worth noting that l_1 -norm is typically used to characterize the sparsity of vectors [25]. Therefore, we attempt to make use of l_1 -norm instead of l_2 -norm for classification. For verifying this suppose, we compare l_1 -norm with l_2 -norm in Section 4.2.

So the l_1 -norm distance measure between the r estimated and the original response vector is computed as

$$d'_i(y^k) = \|y^k - \hat{y}^k\|_1, k = 1, 2, \dots, r \text{ and } i = 1, 2, \dots, N \quad (14)$$

Now, for the k -th partition, an intermediate decision called $j^{(k)}$ is reached with average value of corresponding minimum distance calculated as

$$d^{j^{(k)}} = \sum_{k=1}^r \min_{(i)} d'_i(y^k) / r, k = 1, 2, \dots, r \text{ and } i = 1, 2, \dots, N \quad (15)$$

Therefore, we now have r decisions $j^{(k)}$ with r corresponding distances $d^{j^{(k)}}$ and we decide in favor of the class with minimum average distance

$$\text{Decision} = \arg \min_{(i)} d^{j^{(k)}}, i = 1, 2, \dots, N \quad (16)$$

The overall procedure of considerations above can be summarized as **Algorithm**.

Algorithm. Block-Linear Regression Classification with improved DEF (i-DEF).

Inputs: Class models $X_{i,j}^k \in IR^{q \times p_i}$, and a test image $y^k \in IR^{q \times 1}$ ($i = 1, 2, \dots, N$, $k = 1, 2, \dots, M$, $j = 1, 2, \dots, p_i$).

1. $\widehat{\alpha}_i^k$ is evaluated against each class model, $\widehat{\alpha}_i^k = (X_i^{kT} X_i^k)^{-1} X_i^{kT} y$
2. \widehat{y}^k is computed for each $\widehat{\alpha}_i^k$, $\widehat{y}^k = X_i^k \widehat{\alpha}_i^k$
3. Distance calculation between original and predicted response variables $d_i(y^k) = \|y^k - \widehat{y}^k\|_2$
4. Evaluator is computed as $E(y^k) = [Ave_d(y^k) - \min_i d_i(y^k)] / Ave_d(y^k)$
5. The r ($r \leq M$) effective blocks with r largest $E(y^k)$ are selected
6. l_1 -norm distance measure: $d'_i(y^k) = \|y^k - \widehat{y}^k\|_1, k = 1, 2, \dots, r$
7. Compute average value of corresponding minimum distance $d^{j(k)}$
8. **end for**

Output: $Decision = arg \min_{(i)} d^{j(k)}$

4. Experimental Results and Analysis

The partial occluded face recognition system we designed in this study is running on the hardware environment of Intel (R) Core (TM) i5 (3.40 GHz) with 12.00GB RAM, and the software environment of Windows 7 and Matlab R2012b.

Extensive experiments are carried out to illustrate the efficiency and robustness of the proposed approach in this study. All images are divided into 8 blocks (4*2) [25, 26] and L1LS [28] approach is used for solving the l_1 -minimization problem for proposed i-DEF.

4.1 Complexity Analysis

The computational cost of the algorithm is an essential problem for face recognition application in real life. The core of SRC is a l_1 -minimization problem, which has shown good potential in handling occlusions that are spatially correlated, e.g., random pixel corruption. However, it is not robust to contiguous occlusion, e.g., sunglasses and scarf [26]. Although eSRC [25] has satisfactory performance to identify occluded subjects, it also has an obvious disadvantage as some useful information is lost caused by occlusion, the highest possible resolution version of images is used to exploit more information. However, this makes input image be a very large matrix. For instance, assume that the size of face images is 200×150 , $n = 1000$ images are used for training samples, so $X \in R^{(a \times b) \times (q+n)}$ ($a \times b = 200 \times 150, n = 1000$) is larger than $X \in R^{(a \times b) \times n}$ (SRC). Since l_1 -minimization solver [29] with the empirical computational complexity of SRC is $O((a \times b)^2 n^\epsilon)$, where $\epsilon \approx 1.5$. When dealing with eSRC, its complexity is $O((a \times b)^2 (a \times b + n)^\epsilon)$. It is very clear that eSRC has much higher computational cost than SRC so that it is restricted in reality for time-consuming even using modern high-end devices.

Compared with SRC and eSRC algorithm, our proposed method not only does not add much computational cost, but also can handle the problem of contiguous occlusion. As mentioned in section 3.1, face images are first downsampled to q -dimension ($q = c \times d, c \times d \ll a \times b$). In our method, y^k is transformed into a new subspace by $H^{(k)}$ and $H^{(k)}$ will not be changed with the probe images. The off-line computational complexity for all $H^{(k)}$ is

$O(NMn^2)$. To a probe image, the on-line cost for determining the clear blocks is $O(NMn)$. Therefore, the whole procedure has the approximate complexity with $O(NMn + (c \times d)^2 n^\epsilon)$. The computational cost of classification by proposed method will decrease correspondingly due to removing the contaminated blocks.

4.2 Partial Face Recognition with Random Block Occlusion

For this experiment, we test the effectiveness and robustness of the proposed method against random block occlusion. Extensive experiments are carried out using the Extend Yale B database [30]. The database consists of 2414 frontal face images of 38 individuals under various lighting conditions. The database was divided in five subsets; subset 1 consisting of 66 images (seven images per subject) under nominal lighting conditions was used as the gallery, while all others were used for validation (see Fig. 1). Subsets 2 and 3, each consisting of 12 images per subject, characterize slight-to-moderate luminance variations, while subset 4 (14 images per person) and subset 5 (19 images per person) depict severe light variations.

All experiments for the proposed method are conducted with images downsampled to an order 20×20 . Subset 1 is selected for training, and Subject 2 and 3 for testing, respectively. We simulate various levels of random contiguous occlusion, from 10% to 50%, by replacing a square block of each test image with a “baboon” image. The location of occlusion is randomly selected for Subset 2 and 3 (see Fig. 2).

Fig. 3 represents the recognition accuracy with random block occlusion under Subset 2 for testing based on different algorithms, including Eigenfaces+NN, Fisherfaces+NN, SOM, LRC, SRC, eSRC and proposed method. The proposed method comprehensively outperforms the other methods for all levels of occlusion. Up to 30% occlusion in this Subset, our method performs correctly identifying of 94.39%, 1.83 percent more than SRC and approximately equal to eSRC. In the case of 40% occlusion, the proposed method receives 87.73% correct rate, it is better than any these method. Even at 50% occlusion, none of these methods achieves higher than 60% recognition rate, while the proposed method achieves 65.4%.



Fig. 1. Five Subjects of Extended Yale B database, each row illustrates sample from Subsets

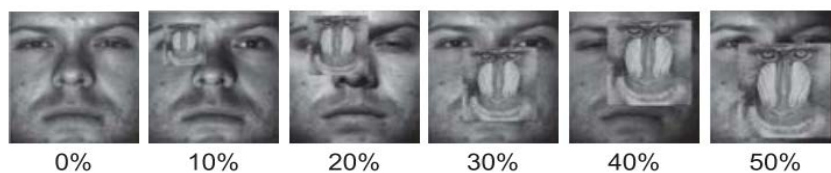


Fig. 2. Test images under varying level of contiguous occlusion with baboon image, from 0% to 50%

Fig. 4 shows the recognition rate of all compared algorithms using subset 3 for testing, the test images are under more various illumination conditions apart from random occlusion. The proposed method significantly outperforms the other seven methods for all levels of occlusion. The recognition rates of other methods drop to 90%, while our proposed method achieves 93.79% recognition rate and much better than SRC and eSRC. Even in the case of half face occluded, proposed method still achieves 65.4% recognition rate. Based on these two experiments, our proposed method is verified for effectiveness and a good tolerance on random occlusion problem. On the other hand, the proposed method with l_2 -norm might achieve good performance till 20 percent occlusion, the recognition rate drops drastically from 30 percent occlusion. Therefore, we will use the proposed method with l_1 -norm for extensive experiments later.

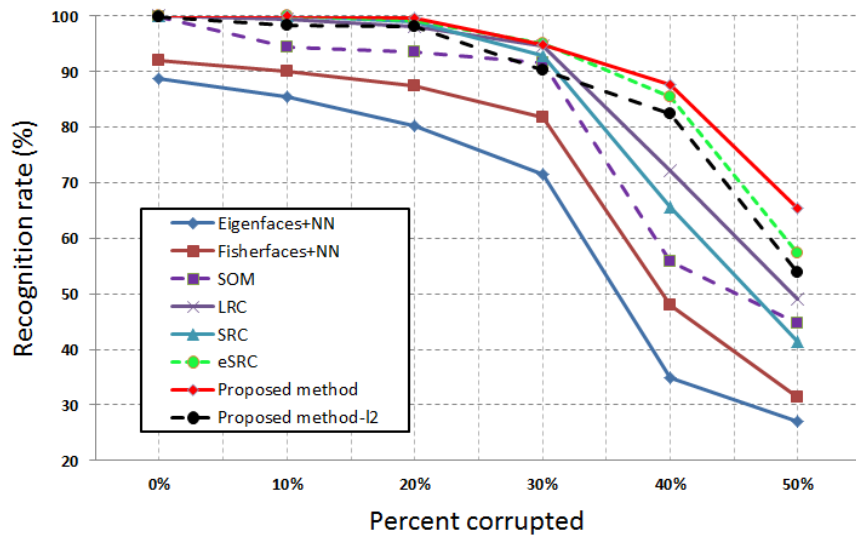


Fig. 3. Comparison result of random block occlusion on subset 2 of Extended Yale B database

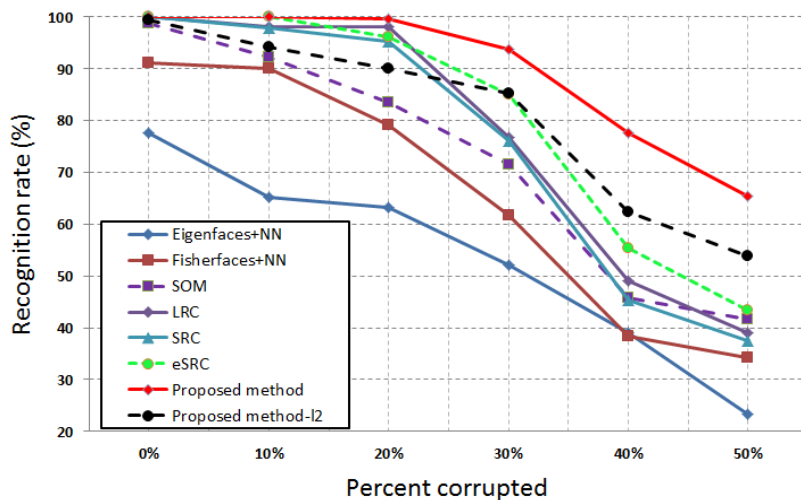


Fig. 4. Comparison result of random block occlusion on subset 3 of Extended Yale B database

4.3 Partial Face Recognition with Real Face Disguises

In this section, we test proposed method on AR database [31], which consists of over 4000-color image of 126 subjects (70 males and 56 females) with size of 576×768 . The database characterizes divergence from ideal conditions by incorporating various (neutral, smile, anger and scream), luminance alterations (left light on, right light on and all side light on), and occlusions modes (sunglass and scarf). There are 26 different images per person, recorded in two different sessions separated by two weeks, each session consists of 13 images. In each session, there are 3 images with different illumination conditions, 4 images with different expressions and 6 images with different facial disguises (3 images wearing sunglasses and 3 images wearing scarf, respectively). It has been used by researcher as a testbed to evaluate and benchmark face recognition algorithms.

As the problem of face identification in the presence of contiguous occlusion is arguably one of the most challenging paradigms in the context of robust face recognition. Commonly used objects, such as cap, sunglasses and scarves, tend to obstruct facial features, causing recognition errors. Therefore, we focus on sunglasses and scarves occlusion in this experiment for verifying the effectiveness of the proposed method. The AR database consists of two modes of contiguous occlusion, images with a pair of sunglasses and scarf, see Fig. 5. We only used a subset of the original database, which includes randomly selected 100 individuals (50 males and 50 females). For each subject, 8 images without occlusion are used as the training samples for recognizing the images with sunglasses and scarf. All images are cropped and downsampled to 40×50 pixels according to the eye positions (see Fig. 5 (b)).

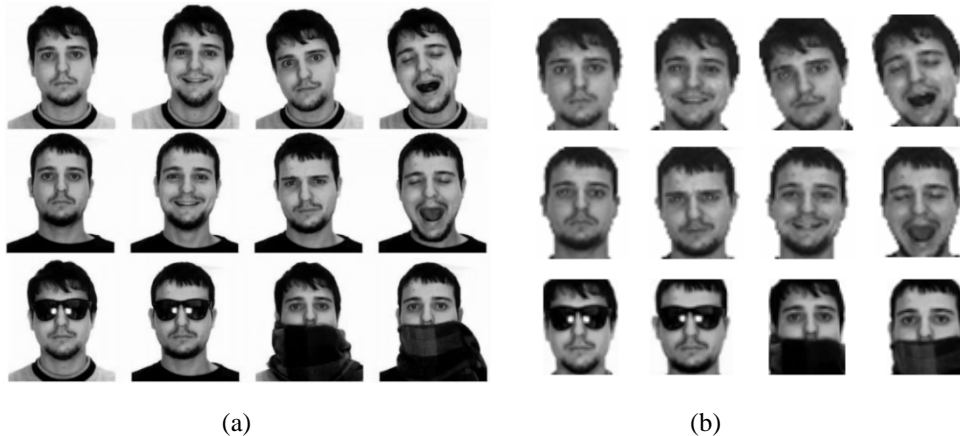


Fig. 5. Samples of AR database from one subject. (a) Original images from AR database, (b) Cropped images for training (first two rows) and testing (third row)

Table 1 depicts a detailed comparison of the proposed method with the other six representative and state-of-the-art algorithms, consisting of Eigenfaces+NN, Fisherfaces+NN, SOM, LRC, SRC and eSRC. For the scenario of sunglass occlusion, this occludes roughly 25 percent of the whole face image. Our proposed method achieves a successful recognition accuracy of 98.5 percent in a 200D feature space, which outperforms eSRC by a margin of 10 percent approximately. In addition, the other scenario considers images with the subject wearing a scarf, which occludes roughly 40% of the face image. It is easy to find that recognition accuracy drops substantially except our proposed method. The proposed method achieves the recognition rate of 78.9%, outperforming LRC, SRC and eSRC by 36.3%, 17.4% and 12.1%, respectively. It indicates that our proposed method is reasonable and effective. By

removing the occluded blocks exactly, the remained blocks can represent useful information for recognition issue.

Table 1. Recognition results (%) for sunglasses and scarf occlusion

Algorithm	Sunglasses	Scarf
Eigenfaces+NN	70.0	12.0
Fisherface+NN	38.5	67.5
SOM	77.8	36.8
LRC	96.5	42.6
SRC	87.5	61.5
eSRC	88.9	66.8
Proposed method	98.5	78.9

For further verifying the effect of feature dimension to recognition accuracy, we downsample the face images to 10×10 , 20×25 , 25×30 , 35×40 and 40×50 , respectively, which correspond to feature dimension of 100, 500, 750, 1400 and 2000. The comparison results are shown as Fig. 6 and Fig. 7.

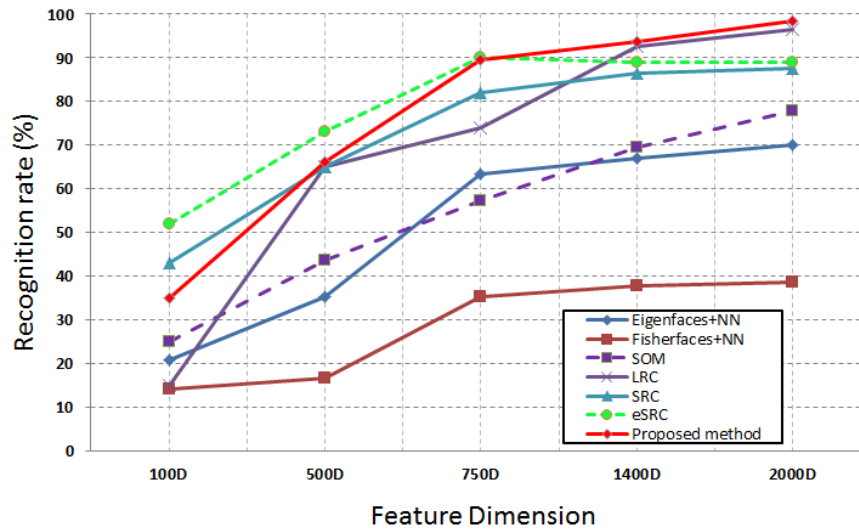


Fig. 6. Recognition rates of different methods on a subset of the AR database with the feature dimensions varying from 100 to 2000 for sunglass occlusion

From Fig. 6 and Fig. 7, global feature-based methods, such as eigenfaces+NN and fisherfaces+NN method cannot achieve a good performance on partial occluded face recognition issue under low dimension features. When images are downsampled to low dimension (such as 100D and 500D), the significant information of face is lost as well. eSRC algorithm has a better performance as it uses the raw image as input vectors, significant information is exploited fully rather than r “clear” blocks in the proposed method. However, the proposed method can handle this problem and achieve better performance with much faster than SRC and eSRC by using low dimension images. Even the images are downsample to 750D feature, the proposed method still outperforms SRC and eSRC by a margin of 8 percent and 4 percent approximately in the case of scarf occlusion. In other words, the proposed method might be suitable to corruption problem as well.

For verifying the performance of i-DEF, a comparison experiment is also conducted by using DEF and i-DEF, respectively, for partially occluded face recognition issue. We also used a subset of the original database of AR, which includes randomly selected 100 individuals (50 males and 50 females). For each subject, 8 images without occlusion are used as the training samples for recognizing the images with sunglasses and scarf. All images are cropped and downsampled to 40×50 pixels according to the eye positions like Fig. 5 (b). Moreover, the occlusion detection results ($r=4$) are utilized for DEF and iDEF, simultaneously, to test the recognition rate of these two algorithms. The experimental results on AR database are expressed as Table 2. From Table 2, the proposed method outperforms the B-LRC with DEF algorithm by a margin of 1.8 percent and 14.6 percent in case of occlusion by sunglasses and scarf, respectively. It is easy to find that recognition accuracy drops rapidly by using DEF when scarf occlusion occurred, the proposed method could handle this issue better by using i-DEF. Because if the occluded portion in the test image y is correctly removed, it can be well approximated as a linear combination of X_i^k of the same subject, and the error between y^k and reconstruction vector $X_i^k \hat{\alpha}_i^k$ should be sparse. This error may affect any part of the image and may be arbitrarily large in magnitude. l_1 -norm is typically used to characterize the sparsity of vectors. On the other hand, l_2 -norm reflects the overall difference between two vectors, and it is sensitive to large errors, as analyzed in the literature [25].

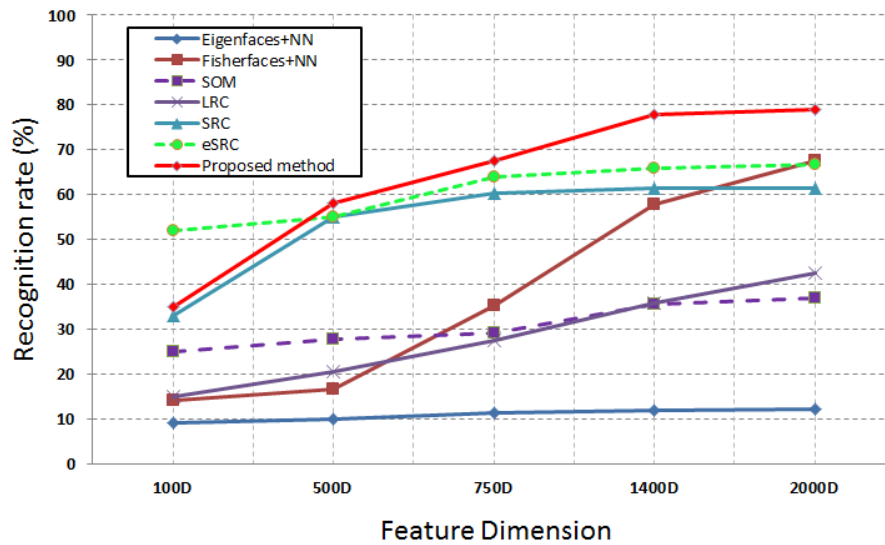


Fig. 7. Recognition rates of different methods on a subset of the AR database with the feature dimensions varying from 100 to 2000 for scarf occlusion

Table 2. Recognition results (%) by using DEF and i-DEF

Algorithm	Sunglasses	Scarf	Average
B-LRC+DEF	96.7	64.3	80.5
B-LRC+iDEF	98.5	78.9	88.7

Comparisons results indicate the effectiveness of removing the contaminated portions and the recognition algorithm could well proceed on the remaining parts of the image. The utilization of l_1 -norm instead of l_2 -norm for classification is reasonable based on analysis above. Combing with the analysis in section 4.1, the proposed method is with low

computational cost and high efficiency. It is reasonable to believe that the proposed method can handle robust face recognition in practical application due to its advantage corresponding to other methods.

5. Conclusion

In this paper, a novel method based on linear regression-based classification (LRC) algorithm is proposed in this paper. First, all images are downsampled and divided into several blocks, and then we exploit the evaluator of each block to determine the clear blocks of the test face image by using linear regression technique. Next, the remained uncontaminated blocks are utilized to partial occluded face recognition issue. Furthermore, an improved Distance-based Evidence Fusion approach is proposed to decide in favor of the class with average value of corresponding minimum distance.

The experimental results based on proposed method that performed on AR face database and extended Yale B database show the effectiveness and robustness. Based on these experimental results, we can conclude that the proposed method has better performance on the issue of partial occluded face recognition than the conventional algorithms such as eigenface+NN, fisherfaces+NN and the state-of-the-art method such as LRC, SRC and eSRC. Moreover, the computational cost of classification by proposed method will decrease correspondingly due to removing the contaminated blocks, and the i-DEF can achieve satisfactory capacity of classification.

Additionally, the issue of frontal face recognition is concerned in this paper; we would like to expand our face recognition system for more comprehensive situations in future work, to enhance its practicability for real world.

References

- [1] M. Turk, A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991. [Article \(CrossRef Link\)](#)
- [2] T. J. Chin, D. Suter, "A Study of the Eigenface Approach for Face Recognition," *Technical Report of Monash University*, Dept. Elect & Comp. Sys Eng, pp. 1-18, 2004. [Article \(CrossRef Link\)](#)
- [3] D. Blackburn, M. Bone, and P. Phillips, "Facial Recognition Vendor Test 2000: Evaluation Report," *National Institute of Science and Technology*, 2000. [Article \(CrossRef Link\)](#)
- [4] J. Lu, N. P. Kostantinos and N. V. Anastasios, "Face recognition using LDA-based algorithm," *IEEE Trans. Neural Networks*, vol. 14, no. 1, pp. 195-200, 2003. [Article \(CrossRef Link\)](#)
- [5] Q. Yang and C. Guo, "Face Recognition Based on Sparse Representation and RNN Optimization," *JDCTA*, vol. 5, no. 9, pp. 314-322, 2011. [Article \(CrossRef Link\)](#)
- [6] S. A. Nazeer, N. Omar and M. Khalid, "Face Recognition System using Artificial Neural Networks Approach," in *Proc. of the IEEE International Conference on Signal Processing, Communications and Networking*, pp. 420-425, 2007. [Article \(CrossRef Link\)](#)
- [7] T. H. Le, "Appling Artificial Neural Networks for Face Recognition," *Advances in Artificial Neural Systems*, vol. 2011, pp. 1-16, 2011. [Article \(CrossRef Link\)](#)
- [8] J. Huang, X. Shao, and H. Wechsler, "Face Pose Discrimination Using Support Vector Machines," *Technical report of George Mason University and University of Minnesota*, pp. 154-156, 1998. [Article \(CrossRef Link\)](#)
- [9] H. K. Ekenel, R. Stiefelhagen, "Why is Facial Occlusion a Challenging Problem," in *Proc. of third international conference of ICB*, pp. 299-308, 2009. [Article \(CrossRef Link\)](#)
- [10] A. Abate, M. Nappi, D. Riccio and M. Tucci, "Occluded Face Recognition by Means of the IFS," in *Proc. of second International Conference of ICIAR*, pp. 1073-1080, 2005. [Article \(CrossRef Link\)](#)

- [11] M. Aleix, "Recognizing Imprecisely Localized, Partially Occluded, and Expression Variant Faces from a Single Sample Per Class," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 748-763, 2002. [Article \(CrossRef Link\)](#)
- [12] N. Amirhosein, A. Esam and A. Majid, "Illumination Invariant Feature Extraction and Mutual-Information-Based Local Matching for Face Recognition under Illumination Variation and Occlusion," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2576-2587, 2011. [Article \(CrossRef Link\)](#)
- [13] G. Baudat, F. Anouar, "Generalized Discriminant Analysis using a Kernel Approach," *Neural Computation*, vol. 12, no. 10, pp. 2385-2404, 2000. [Article \(CrossRef Link\)](#)
- [14] C. Benjamin, B. David, M. Philip and M. Jitendra, "A Real-Time Computer Vision System for Vehicle Tracking and Traffic Surveillance," *Transportation Research Part C: Emerging Technologies*, vol. 6, no. 4, pp. 271-288, 1998. [Article \(CrossRef Link\)](#)
- [15] R. Brunelli, T. Poggio, "Face Recognition: Features Versus Template," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, 1993. [Article \(CrossRef Link\)](#)
- [16] W. Xu, S. Ok and E. Lee, "Intelligent Music Player Based on Human Motion Recognition," *CCIS*, vol. 262, pp. 387-396, 2011. [Article \(CrossRef Link\)](#)
- [17] M. Marsico, M. Nappi and D. Riccio, "FARO: Face Recognition Against Occlusions and Expression Variations," *IEEE Transactions on System, Man, and Cybernetics-Part A: Systems and Humans*, vol. 40, no. 1, pp. 121-132, 2010. [Article \(CrossRef Link\)](#)
- [18] W. Xu, E. Lee, "A hybrid method based on dynamic compensatory fuzzy neural network algorithm for face recognition," *International Journal of Control, Automation and Systems*, vol. 12, issue 3, pp. 688-696. [Article \(CrossRef Link\)](#)
- [19] M. Pontil, A. Verri, "Support Vector machines for 3D object Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 637-646, 1998. [Article \(CrossRef Link\)](#)
- [20] W. Xu, E. Lee, "A Novel Multi-view Face Detection Method Based on Improved Real Adaboost Algorithm," *KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS*, vol. 7, no. 11, pp. 2720-2736, 2013. [Article \(CrossRef Link\)](#)
- [21] K. Hotta, "A View-Invariant Face Detection Method Based on Local PCA Cells," *Journal of Advanced Computational Intelligence and Intelligence Informatics*, vol. 8, no. 11, pp. 1490-1498, 2008. [Article \(CrossRef Link\)](#)
- [22] H. Schneiderman and T. Kanade, "A Statistical Method for 3D Object Detection Applied to Faces and Cars," in *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 746-751, 2000. [Article \(CrossRef Link\)](#)
- [23] J. Kim, J. Choi, J. Yi and M. Turk, "Effective representation using ICA for Face Recognition Robust to Local Distortion and Patiral Occolusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 12, 2005. [Article \(CrossRef Link\)](#)
- [24] J. Wright, M. Yi, J. Mairal, G. Sapiro, T. S. Huang, Y. Shuicheng, "Sparse representation for computer vision and pattern recognition," in *Proc. of IEEE 98*, pp. 1031-1044, 2010. [Article \(CrossRef Link\)](#)
- [25] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210-227, 2009. [Article \(CrossRef Link\)](#)
- [26] I. Naseem, R. Togneri and M. Bennamoun, "Linear regression for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 2106-2112, 2010. [Article \(CrossRef Link\)](#)
- [27] R. Barsi, D. Jacobs, "Lambertian Reflection and Linear Subspaces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 218-233, 2003. [Article \(CrossRef Link\)](#)
- [28] http://www.stanford.edu/~boyd/l1_ls/
- [29] S. J. Kim, K. Koh, M. Lustig, "A interior point method for large scale l1-regularized least squares," *IEEE Journal on Selected Topics in Signal Processing*. vol. 1, no. 4, pp. 606 - 617, 2007. [Article \(CrossRef Link\)](#)
- [30] A. Georghiadis, P. Belhumeur, and D. Kriegman, "From Few to Many Illumination Cone Models

- for Face Recognition under Variable Lighting and Pose,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643-660, 2001. [Article \(CrossRef Link\)](#)
- [31] Martinez and R. Benavente, “The AR Face Database,” *CVC Technical Report 24*, June 1998. [Article \(CrossRef Link\)](#)



Wenkai Xu received his B. S. at Dalian Polytechnic University in China (2006-2010) and Master degree at Tongmyong University in Korea (2010-2012). Currently, he is studying in Department of Information and Communications Engineering Tongmyong University for PH. D. His main research areas are image processing, computer vision, biometrics and pattern recognition.



Suk-Hwan Lee received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from Kyungpook National University, Korea in 1999, 2001, and 2004 respectively. He worked at Electronics and Telecommunications Research Institute in 2005. He is currently an Associate Professor in the Department of Information Security at Tongmyong University, which he started in 2005. He works as an Editor of Korea Multimedia Society Journal, is a member of IEEE, IEEK, IEICE and also is an officer of IEEE R10 Changwon section. His research interests include multimedia signal processing, multimedia security, digital signal processing, bio security, and computer graphics.



Eung-Joo Lee received his B. S., M. S. and Ph. D. in Electronic Engineering from Kyungpook National University, Korea, in 1990, 1992, and Aug. 1996, respectively. Since 1997, he has been with the Department of Information & Communications Engineering, Tongmyong University, Korea, where he is currently a professor. From 2000 to July 2002, he was a president of Digital Net Bank Inc. From 2005 to July 2006, he was a visiting professor in the Department of Computer and Information Engineering, Dalian Polytechnic University, China. His main research interests include biometrics, image processing, and computer vision.