Random projection-based partial feature extraction for robust face recognition

Chunfei Ma, June-Young Jung, Seung-Wook Kim, Sung-Jea Ko*
Department of Electrical Engineering, Korea University, Republic of Korea

A R T I C L E   I N F O

Article history:
Received 12 May 2014
Received in revised form
13 August 2014
Accepted 3 September 2014
Communicated by: Ran He
Available online 16 September 2014

Keywords:
Face recognition
Feature extraction
Robustness
Random projection
Compressed sensing

A B S T R A C T

In this paper, a novel feature extraction method for robust face recognition (FR) is proposed. The proposed method combines a simple yet effective dimensionality increasing (DI) method with an information-preserving dimensionality reduction (DR) method. For the proposed DI method, we employ the rectangle filters which sum the pixel values within a randomized rectangle window on the face image to extract the feature. By convolving the face image with all possible rectangle filters having various locations and scales, the face image in the image space is projected to a very high-dimensional feature space where more discriminative information can be incorporated. In order to significantly reduce the computational complexity while preserving the most informative features, we adopt a random projection method based on the compressed sensing theory for DR. Unlike the traditional holistic-based feature extraction methods requiring the time-consuming data-dependent training procedure, the proposed method has the partial-based and data-independent properties. Extensive experimental results on representative FR databases show that, as compared with conventional feature extraction methods, our proposed method not only achieves the higher recognition accuracy but also shows better robustness to corruption, occlusion, and disguise.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Face recognition (FR) has been a very popular topic in computer vision research for several decades [1–10]. Recently FR has been successfully applied to a wide range of applications, such as biometric identity authentication, human computer interaction, and intelligent surveillance. As an indispensable procedure in FR, the feature extraction method remains to be investigated due to various challenges including illumination change, facial expression variation, and noise, such as corruption, occlusion, and disguise, etc.

The subspace analysis-based dimensionality reduction (DR) methods, such as principal component analysis (PCA) [1], linear discriminant analysis (LDA) [2], and independent component analysis (ICA) [3], have been employed to project the face image in high-dimensional image space into a low-dimensional subspace. In the subspace, the face can be represented more compactly, thereby helping to remove those aforementioned change, variation, and noise. The PCA determines the basis vectors by finding the directions of the large variance in data [1]. Then the extracted basis vectors are used to construct a measurement matrix, namely the projection matrix, for DR. Unlike the PCA that best describe the data, the LDA searches for the basis vectors in the underlying subspace that can best discriminate the data among classes [2]. As a generalization of PCA, ICA aims to find the independent nonorthogonal basis vectors [3] which reconstruct the data better than PCA in the presence of noise.

However, the conventional subspace analysis-based feature extraction methods can only achieve the limited performance in FR task [3–8]. The reasons are: first, the direct projection from the image space to a subspace cannot extract the most discriminative information for classification. Second, the most representative structural information in the image space may not be preserved due to the inherent property of the DR method [1,2]. Third, the illumination change or the expression variation, existing in the original image space, is retained by the conventional DR method. In contrast, recent research work emphasizes the importance of the high-dimensionality feature in FR task. In [9], the dense facial landmark-based multiscale feature sampling has been proposed to extract the high-dimensional feature for high performance FR. Liu et al. [10] and Zhao et al. [11] showed that mapping the face image from the input space to the high-dimensional feature space through Gabor filters helps to incorporate more discriminative information for better FR. Moreover, the works [12–15] on other image classification problems also reveal the significance of the high-dimensional feature. Yang et al. [12] showed that the over-completed high-dimensional representation is more separable, and Sanchez et al. [13] reported on the necessity of high-
dimensional features in large-scale image classification. Pooling in spatial [14] and feature spaces [15] also lead to the higher dimensionality and the better performance.

In addition, the aforementioned subspace analysis-based DR methods require the data-dependent training procedure, which makes the FR system inefficient [16,17]. When face images are added or deleted in the FR system, these methods inevitably need to re-compute the projection matrix, which requires the update of all face images. Random projection (RP) [16–21], as a newly emerged DR method, has attracted more attention due to its efficiency and data-independent property. In RP, no training samples are required to calculate the projection matrix since it can be generated beforehand. Since the subspace estimated by RP is independent of the samples and their dimension, RP does not require the update of the projection matrix when the data changes. The RP has been successfully applied to the feature extraction for texture classification [19]. The RP-based DR method also has been used to obtain the prominent performance on object tracking [20]. Moreover, it has been reported in [21] that RP can achieve the comparable performance as the PCA-based Eigenspace method in FR. However, the research of the more accurate and robust RP-based feature extraction method remains to be an open issue.

Moreover, most adopted features for FR are holistic-based features, such as Eigenfaces [1], Fisherfaces [2], Laplacianfaces [22], and variants [4,5]. It has been shown that these features are sensitive to the various types of noise such as occlusion, corruption and disguise. In contrast, the partial-based facial features [23–28], such as patches around eyes or nose, are more efficient and robust than the holistic-based ones. Savvides et al. [24] showed that the partial feature is more discriminative than the holistic-based one in FR. Seo et al. [25] and Jun et al. [26] illustrated that the local features are very robustness to facial images variations in terms of robust FR. Heisele et al. [27] demonstrated that the facial component-based representation is superior to the holistic-based one. Zhu et al. [28] illustrated the effectiveness of the multiscale patch-based collaborative representation for FR.

In this paper, motivated by these aforementioned works, we propose a novel partial-based feature extraction method which combines a simple yet effective dimensionality increasing (DI) method with an information-preserving DR method. For the proposed DI method, we employ the rectangle filters which sum the pixel values within a randomized rectangle window on the face image to extract the feature. By convolving the face image with all the possible rectangle filters having various locations and scales, the face image in the image space can be projected to a very high-dimensional feature space where more discriminative information can be incorporated. In order to significantly reduce the computational complexity while preserving the most informative features, a random projection method based on the compressed sensing theory is adopted for DR. To further enhance the FR performance, a multi-radius local binary pattern (MLBP)-based image representation method is also proposed. Unlike the traditional holistic-based feature extraction method requiring the time-consuming data-dependent training procedure, the proposed method has the data-independent properties. Furthermore, the proposed feature extraction method can be easily synthesized as a single measurement matrix which fuses the aforementioned DI and DR procedures. Thus it is very efficient to generate the measurement matrix since this matrix is independent of the training dataset and just needs to be computed only once offline.

The rest of the paper is organized as follows. Section 2 presents and analyzes the proposed feature extraction method. Section 3 performs experiments and Section 4 concludes the paper.

2. Proposed method
The proposed feature extraction method primarily consists of three parts as shown in Fig. 1. The original face image is first represented by using the proposed MLBP to incorporate more structural information. Then the MLBP image vector will be mapped to a very high dimensional space by the proposed DI method to further incorporate discriminative information for better classification. Finally, the DR is adopted for facilitating practical application while preserving the salient information of face image in the aforementioned very high dimensional space. Some notations used in this paper are summarized in Table 1.

### Table 1
Summary of some notations used in this paper.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x \in \mathbb{R}^m$</td>
<td>Transformed image vector with dimension $m$</td>
</tr>
<tr>
<td>$y \in \mathbb{R}^n$</td>
<td>Extracted feature vector with dimension $n$</td>
</tr>
<tr>
<td>$p_i \in \mathbb{R}^m$</td>
<td>Rectangle filter with dimension $m$</td>
</tr>
<tr>
<td>$R \in \mathbb{R}^{m \times m}$</td>
<td>Rectangle filter matrix</td>
</tr>
<tr>
<td>$P \in \mathbb{R}^{1\times n}$</td>
<td>Random projection matrix</td>
</tr>
<tr>
<td>$M \in \mathbb{R}^{1\times m}$</td>
<td>Measurement matrix</td>
</tr>
</tbody>
</table>

---

**Fig. 1.** Overview of the proposed feature extraction method.
face image, we represent the face image by using the proposed multi-radius LBP (MLBP) as follows: We first calculate the LBP maps of the face image with different radii (Fig. 2(b) and (c)), namely the $LBP_{P,R}$, where parameter $P$ controls the quantization of the angular space, whereas $R$ determines the spatial resolution of the operator [29]. And then the LBP maps are fused together in a certain manner as shown in Fig. 2(d). In order to speed up the computation, an integral MLBP map in Fig. 2(e) of the fused map is pre-computed.

Note that our proposed MLBP is different from the previous work [30] which concatenates the LBP histograms with different radii together. In this paper, we adopt the average map of $LBP_{8,1}$ and $LBP_{8,2}$ for FR since the extensive experiments indicate that this average map produces the best FR result. Note that the other local descriptors such as LGP, LDN, and LDP can also be employed in the same manner as MLBP.

Most local descriptor-based feature extraction methods extract the feature by concatenating the histograms obtained from sub-regions of the descriptor-filtered face image. However, this holistic-based method loses the discriminative spatial information which is essential for robust FR. To solve this problem and strengthen the robustness of the feature, we propose a method to extract the feature in a partial-based way. To this end, the feature is extracted by weighting the sums of the responses of several rectangle filters of which the locations and scales are randomly generated. An example of a randomly generated feature extraction operator is shown in Fig. 2(f). The description of the rectangle filters is presented in the following section. The detailed weighting method is illustrated in Section 2.4.

2.2. Rectangle filters for DI

In this section, aiming to further incorporate the informative and discriminative information for better classification, we employ the rectangle filters to map the face image in the MLBP space into a very high-dimensional space called multiscale and multi-radius LBP (MMLBP) space.

![Image](https://via.placeholder.com/150)

**Fig. 2.** Example of the proposed MLBP representation and feature extraction. (a) Input image. (b) $LBP_{8,1}$ map. (c) $LBP_{8,2}$ map. (d) The average map of (b) and (c). (e) The integral LBP map of (d). (f) One feature extraction operator consists of several rectangle filters with randomized locations and scales.

![Image](https://via.placeholder.com/150)

**Fig. 3.** Illustration of the multiscale rectangle filters matrix adopted for DI.
Given a face image with width $w$ and height $h$, the rectangle filters are defined as follows:

$$p_{u_0, v_0}(u, v) = \begin{cases} 1, & u_0 \leq u \leq w, v_0 \leq v \leq h; \\ 0, & \text{otherwise} \end{cases}$$

where $u_0$ and $v_0$ represent the offset coordinates of the rectangle filter $p_i$ with width $w_i$ and height $h_i$. For each face image, it can be found that the number of the exhaustive rectangle filters is approximated to $L = m^2 = (wh)^2$ (i.e., the number of the possible locations times that of the possible scales). By convolving with these filters, the face image can be mapped to the MMLBP space with dimension $L$.

Here, for effective computation, we formulate the filter convolutions as simple vector dot-products. In this manner, the $\mathbf{P}$ can be formulated as a row vector having the same dimension as the face image. Then a large multiscale rectangle filter matrix $\mathbf{P}$ can be obtained simply by stacking each $p_i$ together (see Fig. 3).

Since the dimension of the MMLBP space, $L$, is in general on the order of $10^6$–$10^{10}$, it is hard to directly utilize the high dimensional feature for practical application of FR [9]. To break the curse of dimension, we use a very sparse RP matrix, $\mathbf{R} \in \mathbb{R}^n \times L$, to project the face image in the MMLBP space (with dimension $L$) into a very low dimension space (with dimension $n$, $n \ll L$). The detailed description of the DR method we adopted will be given in the following section.

### 2.3. Random projection (RP) for DR

RP independently selects the feature in a random manner while preserving the most salient information of the signal [16–18]. A RP matrix $\mathbf{R} \in \mathbb{R}^{n \times L}$ having unit-length and orthogonal rows, projects the high-dimensional data $\mathbf{z} \in \mathbb{R}_L$ into a low-dimensional subspace $\mathbf{y} \in \mathbb{R}_n$,

$$\mathbf{Rz} = \mathbf{y},$$

Fig. 4. Illustration of the sparse random projection matrix adopted for DR. The white, gray and black elements represent the zero, positive and negative coefficients of RP matrix, respectively.

Fig. 5. Illustration of the feature extraction process by the proposed measurement matrix. (a) The proposed measurement matrix. (b) Every entry in the feature vector is obtained by weighting the summations of several rectangle filters' responses on the integral feature map.

Fig. 6. Recognition rates versus the feature dimensions ($n$) with different numbers of rectangle filters ($c$). The first and second row show, respectively, the results on the extend Yale B and AR database when various classifiers are employed.
where \( n \ll L \). RP is motivated by Johnson–Lindenstrauss (JL) lemma stating that a set of points in a high-dimensional Euclidean space can be mapped down onto a low-dimensional subspace such that the distances between the points are approximately preserved [33], Baraniuk et al. proved that the RP satisfying the JL lemma also holds true for the restricted isometry property (RIP) [34] in the compressive sensing [35]. Therefore, with a very high probability the original data can be reconstructed with the minimum error. In [21], a Gaussian RP matrix has been successfully applied to FR. However, the high dimension of the Gaussian PR matrix leads to computationally inefficient DR. In this paper, instead of the Gaussian PR matrix, we use a very sparse RP matrix [36] satisfying the RIP [34], thereby resulting in a more efficient DR. As shown in Fig. 4, the adopted RP matrix \( R \) with i.i.d entries is defined as
\[
\begin{cases}
1 & \text{with probability } \frac{1}{2} \\
0 & \text{with probability } 1 - \frac{1}{2} \\
-1 & \text{with probability } \frac{1}{2n}\frac{1}{2}
\end{cases}
\] (3)
By setting \( s = L/c \), where \( c \ll 10 \), is a empirically determined constant, we obtain a very sparse matrix whose most entries are equal to zero. For each row of \( R \), only \( c \) non-zero entries are needed to be computed, which requires very low computational complexity of \( O(cn) \).

### 2.4. Feature extraction by the proposed measurement matrix

As stated in the previous subsection, the feature responses from the \( c \) rectangle windows are linearly combined to produce the entry of the final feature vector.

Next we introduce a method of utilizing the entries of each row of \( R \) to calculate the weighted summation. Like most of the subspace analysis-based feature extraction methods, the proposed method can also be operated using a single measurement matrix, \( M \in R^{n \times m} \), which is obtained by combining the proposed multi-scale filters matrix, \( P \in R^{2 \times m} \), with the sparse RP matrix, \( R \in R^{n \times L} \) (see Fig. 5(a))
\[
M = RP.
\] (4)
Thus, the feature extraction is simply given by the following equation:
\[
Mx = y
\] (5)
where \( x \in R^m \) is the MLBP vector of the face image and \( y \in R^n \) is the final extracted feature vector. As shown in see Fig. 5(b), the entry of \( y \) is obtained by weighting the responses of those randomly activated rectangle filters as follows:
\[
y_i = \sum_j m_{ij}x_j
\] (6)
where \( m_{ij} \) and \( x_j \) represent the entries of \( M \) and \( x \). Note that since the non-zero coefficients in \( R \) can be positive or negative, the feature is obtained by simply comparing the differences of the filter responses of the randomized rectangle filters on the MLBP map of the face image.

### 2.5. Analysis of the proposed method

Compared to the subspace-based feature extraction methods, the proposed method has the following advantages: (1) the MLBP representation can incorporate more structural information of the
face image. (2) the high-dimensional feature obtained by using the proposed multi-scale rectangle filters can be considered as an over-completed representation [12] which is able to include more discriminative information, and (3) the multi-scale rectangle filters comprehensively encode the important micro and macro structures of the face.

In addition, due to the multi-scale and spatially localized properties, the extracted feature can capture the critical spatial relationship information among the local facial structures. The robustness of the proposed method is multifold: first, in terms of illumination change, the proposed MLBP is invariant against monotonic gray-scale change and intensity inhomogeneity, etc. [29]. Second, in terms of expression variation, the spatially distributed multi-scale rectangle filters can capture the invariant information from the local regions of the face image [27]. Third, in terms of occlusion, the proposed method can still utilize the rest parts of the image which are not corrupted or occluded [28]. As a result, the proposed method is not only robust to the illumination change but also less sensitive to the local changes of the face image such as expression, occlusion, and inaccurate alignment than the holistic-based method.

Finally, the compressed sensing theory guarantees that the extracted feature is able to preserve the most salient information of the face image in the high-dimensional feature space. This ensures that we can classify the face image in the low-dimensional feature space effectively and efficiently.

2.6. Discussion of the dimensionality

In general, the feature dimension is highly related with the discriminative ability of the feature. In the proposed method, in order to guarantee the effectiveness of the RP-based DR, the theoretical bound [16] for the feature dimension n that satisfies the JL lemma is as follows:

\[ n \geq \left( \frac{4 + 2 \beta}{e^2 / 2 - e^3 / 3} \right) \ln(d), \]

where \( d \) is the number of the face images, \( 0 < \epsilon < 1 \) and \( \beta > 0 \). While Bingham et al. [17] pointed out that this bound is much higher than the one that suffices to give good result. Candes et al. [37] subsequently presented an alternative bound satisfying the RIP in compressed sensing [33]. This bound is much tighter than the aforementioned one

\[ n \geq C S \log (L/S) \]

where \( C \) and \( S \) are the constants. Note that this bound is much tighter than the aforementioned one. Through the extensive experiments, we found that the proposed method can produce the satisfactory performance when \( n \geq 100 \) (\( L=10^6 \), \( C=2 \), and \( S=10 \) [20]).

3. Experimental results

In this section, we perform experiments on representative face databases to demonstrate the performance of the proposed feature. In Section 3.1, we first evaluate the performance of the proposed feature with different numbers of rectangle filters on two representative face databases, extended Yale B [38] and AR [39]. Next in Section 3.2, the proposed MLBP is compared with the conventional LBP feature in terms of various feature extraction methods. In Section 3.3, we compare the proposed feature with the conventional features on four databases, extended Yale B, AR, CMU PIE [40], and LFW [41], in conjunction with various classifiers. Then in Section 3.4,
we demonstrate the robustness of the proposed feature under random block occlusion, random pixel corruption, and real disguise. Finally, we compare the computational cost of the proposed feature with the conventional features in Section 3.5.

In the experiments, as in [23], all face images were cropped and aligned by using the eye locations supplied by face databases. We employed three famous classifiers, nearest neighbor (NN) classifier, nearest subspace (NS) classifier, and collaborative representation based classification (CRC) classifier [42] for experiment. For fair comparison, all the experiments are repeatedly performed for 10 times and the averaged FR results are finally reported. The face recognition rates are computed with the feature dimension, \( n = 100, 300, 500, \) and 700.

3.1. The impact of filter number on FR performance

We first examine the FR performance for the different number of rectangle filters. Fig. 6 shows the face recognition rates for \( c = 2, 4, 6, 8, \) and 10. Several interesting observations can be found in this figure: First, the performance tends to improve as \( n \) increases. Second, the proposed features with different \( c \) values achieve the similar performance when the same classifier is employed. Third, the proposed feature with the CRC classifier achieves the best performance on both two databases. As discussed in Section 2.3, the \( c \) determines the computational complexity of the proposed method. Thus taking the complexity as well as the performance into account, we adopt \( c = 4 \) through all the following experiments.

3.2. Comparison between MLBP and LBP

To investigate the efficacy of the proposed MLBP representation, we compare the performance of MLBP with the conventional LBP representation [29] in terms of two types of feature extraction methods. The first method is the conventional holistic histogram-based feature extraction method [30]. The second one is the proposed partial feature extraction method, namely our proposed DI and DR based method. The histogram-based features were implemented with various block sizes of \( 3 \times 3, 5 \times 5, 7 \times 7, \) and \( 9 \times 9 \). The experiments were conducted on Extend Yale B and AR databases and the CRC classifier was employed. Fig. 7 shows the comparison results of two methods. It can be observed that the MLBP performs slightly worse than the LBP when using the holistic histogram-based representation. However, the MLBP performs much better than the LBP when the proposed partial-based feature extraction method is adopted. It indicates that, compared to the conventional LBP, the MLBP can capture more useful information which is critical to the success of the proposed partial-based feature extraction method. In addition, by comparing
Fig. 7(a) with (c), and Fig. 7(b) with (d), we can also see that the performance of MLBP is better when the proposed feature extraction method is adopted; in the meanwhile, the proposed method uses substantially lower dimensional feature than the histogram-based method (e.g., the histogram dimension of the image with $3 \times 3$ blocks equals to $n = 3 \times 3 \times 256 = 2304$). Therefore it can also be inferred that, for MLBP, the histogram-based method is less effective and efficient than the proposed method.

3.3. Face recognition under illumination changes and expression variations

In this experiment, we evaluate the performance of the proposed feature in FR under illumination changes and expression variations. We compare the proposed feature with the popular holistic-based features; Eigenfaces [1], Randomface [23], Fisherfaces [2], and local histogram-based LBP [29] and LPQ [32] features. As in [29], the LBP and LPQ features were implemented with various block sizes of $3 \times 3, 5 \times 5, 7 \times 7,$ and $9 \times 9$. However, due to the high dimensionality (in the order of $10^3$–$10^4$), it is hard to directly compare this histogram-based LBP and LPQ features with the proposed feature in the same dimension. Moreover, the maximal number (i.e., the dimension) of the valid Fisherfaces is less than the number of class of the database [2]. Therefore only the best performances of Fisherfaces, histogram-based LBP, and LPQ features instead are given for comparison. Furthermore, we also compare the proposed method with and without using the proposed MLBP representation. The feature without using the MLBP representation is called compressed multiscale gray-scale feature (CMG). Thus we consider 12 combinations of features with classifiers as follows: Eigenfaces with NN (E+NN), Randomface with NN (R+NN), CMG with NN (CMG+NN), the proposed feature with NS (Our+NS), Eigenfaces with NS (E+NS), Randomface with NS (R+NS), CMG with NS (CMG+NS), the proposed feature with NS (Our+NS), Eigenfaces with CRC (E+CRC), Randomface with CRC (R+CRC), CMG with CRC (CMG+CRC), and the proposed feature with CRC (Our+CRC).

3.3.1. Extended Yale B database

The extended Yale B database contains 2414 frontal face images from 38 individuals. We used the cropped and normalized $64 \times 56$ pixel face images which were taken under various illumination conditions. Some sample images are shown in Fig. 8. We randomly split the database into two halves. One half was used for training (i.e., about 32 images per subject), and the other half for testing. Randomly splitting the database ensures that our results are independent of the choice of the training and testing data. Fig. 9 shows the recognition rates versus feature dimensions. It can be seen that our proposed feature achieves the best results in all the dimensions when NS and CRC are employed. The recognition rates of Our+NN are similar to those of E+NN and R+NN. Note that Our+CRC produces the best recognition rate of 99.8% at $n=700$, compared to 94.7% of CMG+CRC at $n=700$, 97.7% of R+CRC at $n=500$, and 98.8% of E+CRC at $n=700$. Moreover, the CMG feature achieves very low recognition rates when NN is employed. Although the CMG achieves the good performance when NS and CRC are employed, the performance is lower than the proposed feature. In addition, from Table 2, we can see that the
highest recognition rates of Fisherfaces, histogram-based LPQ and LBP features are, respectively, 96.5%, 98.1%, and 98.4% with the CRC classifier. These rates are lower than the result of the proposed feature with the CRC classifier when $n \geq 300$. The comparison results in this experiment validate that the proposed method is more robust to illumination changes than other features, which is mainly due to the strengths of proposed MLBP representation.

3.3.2. AR database

The AR database consists of over 4000 frontal images from 126 individuals. As in [23], a subset (with only illumination and mainly due to the strengths of proposed MLBP representation. more robust to illumination changes than other features, which is

results in this experiment validate that the proposed method is more robust to illumination changes than other features, which is mainly due to the strengths of proposed MLBP representation.

3.3.2. AR database

The AR database consists of over 4000 frontal images from 126 individuals. As in [23], a subset (with only illumination and mainly due to the strengths of proposed MLBP representation. more robust to illumination changes than other features, which is mainly due to the strengths of proposed MLBP representation.

Table 4

<table>
<thead>
<tr>
<th>Feature</th>
<th>Performance (%)</th>
<th>Dimension</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP-histogram</td>
<td>91.4</td>
<td>$6.40 \times 10^3$</td>
<td>$5 \times 3$ blocks</td>
</tr>
<tr>
<td>LPQ-histogram</td>
<td>92.3</td>
<td>$6.40 \times 10^3$</td>
<td>$5 \times 3$ blocks</td>
</tr>
<tr>
<td>Fisherfaces</td>
<td>91.2</td>
<td>67</td>
<td>100 subjects</td>
</tr>
<tr>
<td>Our</td>
<td>94.4</td>
<td>700</td>
<td>–</td>
</tr>
</tbody>
</table>

Fig. 14. Sample images of six people in LFW database.

Fig. 15. Recognition rates on LFW database, for various features and classifiers. (a) NN. (b) NS. (c) CRC.
recognition rates of Fisherfaces, histogram-based LPQ and LBP features are, respectively, 37.1%, 39.5%, and 37.3% with the CRC classifier. These rates are lower than the result of the proposed feature with the CRC classifier when $n \geq 300$. Though the FR difficulty of the LFW database is comparatively higher than other databases, the proposed method can still outperforms the other features and achieve the favorable performance. It is expected that the performance of proposed method on the LFW database can be further improved by utilizing the landmark-based pose normalization method [9].

3.4. Face recognition with occlusion

In this subsection, we evaluate the robustness of the proposed feature under different types of occlusions and disguise, such as random block occlusion, random pixel corruption, and sunglass disguise [23]. We compare the proposed feature with Randomface [23] and Eigenfaces [1] when the CRC classifier is employed.

3.4.1. FR under random pixel corruption

To test the robustness of proposed feature against random pixel corruption, we use the same experimental setting in Section 3.2. The extended Yale B database was randomly split into two halves, one half was used as training samples, and the other half with random generated pixel corruptions was used for testing. Images were resized to $96 \times 84$ pixel. Some examples can be found in Fig. 16(a). For each testing image, we replaced a certain percentage of its pixels by uniformly distributed random values within $[0, 255]$. The corrupted pixels were randomly chosen in each test image.

Table 6 shows the results of $R+\text{CRC}$, $E+\text{CRC}$, and $\text{Our}+\text{CRC}$ with $n=700$ when the percentage of corrupted pixels is from 10% to 40%. Note that the advantage of $\text{Our}+\text{CRC}$ over $R+\text{CRC}$ and $E+\text{CRC}$ is very clear. Especially, $\text{Our}+\text{CRC}$ produce the recognition rates of 98.6% and 94.5% when 10% and 20% pixels are corrupted, respectively, while $R+\text{CRC}$ has those of 43.5% and 22.5%, and $E+\text{CRC}$ has those of 88.1% and 78.8%, respectively.

3.4.2. FR under random block occlusion

In this part, we test the robustness of the proposed feature to random block occlusion. We also used the same experimental setting as the above section. The difference is that the testing samples were occluded by the random generated block (an image). Fig. 16(b) shows some sample images with random block occlusions.

### Table 5

Best performance on LFW database.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Performance (%)</th>
<th>Dimension</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP-histogram</td>
<td>37.3</td>
<td>$1.25 \times 10^4$</td>
<td>7 x 7 blocks</td>
</tr>
<tr>
<td>LPQ-histogram</td>
<td>39.5</td>
<td>$6.04 \times 10^3$</td>
<td>5 x 5 blocks</td>
</tr>
<tr>
<td>Fisherfaces</td>
<td>37.1</td>
<td>157</td>
<td>158 subjects</td>
</tr>
<tr>
<td>Our</td>
<td>43.9</td>
<td>700</td>
<td>–</td>
</tr>
</tbody>
</table>

### Table 6

Corruption test (random pixel).

<table>
<thead>
<tr>
<th>Corruption rate</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R+\text{CRC}$</td>
<td>43.5%</td>
<td>22.5%</td>
<td>12.1%</td>
<td>8.1%</td>
</tr>
<tr>
<td>$E+\text{CRC}$</td>
<td>88.1%</td>
<td>78.8%</td>
<td>64.6%</td>
<td>47.9%</td>
</tr>
<tr>
<td>$\text{Our}+\text{CRC}$</td>
<td>98.6%</td>
<td>94.5%</td>
<td>83.5%</td>
<td>64.0%</td>
</tr>
</tbody>
</table>

Fig. 16. Recognition under random pixel corruption and random block occlusion. (a) Test images from extended Yale B, with random pixel corruption. The corruption percentages of pixel from left to right are 10%, 20%, 30%, and 40%, respectively. (b) Test images with random block occlusions. The occlusion percentages of pixel from left to right are 10%, 20%, 30%, and 40%, respectively. (c) Original images.
Table 7 lists the results of R+CRC, E+CRC, and Our+CRC under the percentage of occlusion from 10% to 40% at \( n = 300 \). We observe that Our+CRC achieves the higher recognition rates than R+CRC and E+CRC for all four occlusion percentages. Especially, Our+CRC exhibits the recognition rates of 98.3% and 94.9% when 10% and 20% pixels are occluded, respectively, while R+CRC only has those of 87.2% and 78.3%, and E+CRC has those of 90.6% and 84.2%.

### 3.4.3. FR with under face disguise

A subset from the AR database is used in this experiment. This subset consists of 2599 images from 50 males and 50 females (26 samples per class). We did three tests: one follows the experimental setting in [23], while the other two more challenging ones follow the experimental setting in [45]. Images were resized to 165×120. Fig. 17(a) shows the example of the test images with sunglasses. The face recognition rates of the aforementioned methods are computed for the feature dimension \( n = 100, 200, \) and \( 300 \).

For the first disguise test, 799 images (8 samples per subject) of non-occluded frontal views with various facial expressions in Sessions 1 and 2 were used for training, while two separate subsets with sunglasses of 200 images (1 sample per subject per Session, with neutral expression) used for testing. The experimental results are listed in Table 8. We observe that Our+CRC outperforms R+CRC and E+CRC in all dimensions. Especially, Our+CRC achieves the high recognition rates of 90.5% at \( n = 300 \), 51% higher than that of R+CRC, 40% higher than that of E+CRC.

In the second and third disguise tests, we conduct the experiment under more complex disguise with variations of illumination and longer data acquisition interval. 400 images (4 neutral images with different illuminations per subject) of non-occluded frontal views in Session 1 were used for training, while the disguised images (3 images with various illuminations and sunglasses per subject per session) in Sessions 1 and 2 were respectively used for testing. Tables 9 and 10 show that the recognition rates of Our+CRC are much higher than those of both R+CRC and E+CRC in all dimensions.

### 3.5. Computational cost

Apart from recognition rate, computation cost is a critical issue for practical FR systems. As discussed in Section 2.3, the computational complexity of the proposed method is \( \mathcal{O}(cn) \) which results in very light computational burden. As the proposed method is independent of the data, the measurement matrix \( \mathbf{M} \) can be generated before the feature extraction process, which can significantly reduce the online computational cost. Specifically, in actual implementation, instead of mapping each face image first

---

**Table 7**

<table>
<thead>
<tr>
<th>Occlusion rate</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>R+CRC</td>
<td>87.2%</td>
<td>78.3%</td>
<td>66.6%</td>
<td>55.4%</td>
</tr>
<tr>
<td>E+CRC</td>
<td>90.6%</td>
<td>84.2%</td>
<td>77.7%</td>
<td>67.4%</td>
</tr>
<tr>
<td>Our+CRC</td>
<td>98.3%</td>
<td>94.9%</td>
<td>84.7%</td>
<td>69.2%</td>
</tr>
</tbody>
</table>

---

**Fig. 17.** Recognition under sunglass disguises. (a) Test images from AR database, with sunglasses disguises. (b) Images without disguises.
into the MMLLBp space using P and then selecting the feature using \( R \), we only need to store the nonzero entries of the very sparse RP matrix \( R \) and then extract the activated rectangle features (corresponding to those nonzero entries of \( R \) in \( P \)). This can lead to the low computational cost both on running time and memory thereby speeding up the proposed method. To intuitively assess the computational cost, we compare the CPU time of the proposed method with the histogram-based LBP (LPB-hist) and LPQ (LPQ-hist) features, Eigenfaces, Randomfaces, Fisherefaces, and CMG. The CPU time equals to the total recognition time divided by the number of test samples. The extended Yale B database was used in this experiment. For fair comparison, the experiment was conducted under the same feature dimension and classifier. The dimension was set as \( n = 256 \) and the CRC classifier was adopted. For the histogram-based LBP and LPQ features, the 256 bins histograms were extracted for comparison. For Fisherfaces, the dimension was set as \( n = k - 1 \), where \( k \) is the number of class of the database (i.e., \( k = 37 \) for extended Yale B). The programming environment is Matlab version R2012a. The desktop used is Intel (R) Core (TM) Duo 2.80 GHz CPU and with 4 G RAM.

From the Table 11, we can observe that the running time of the proposed method is faster than Fisherfaces and comparable to the Eigenfaces, histogram-based LBP and LPQ features. Compared to the CMG feature, the proposed method requires about 2.7 ms more to recognize one sample on average, which is mainly due to the calculation of the MLBP representation. It is expected that the optimized and parallel implementation can further improve the real-time property of the proposed method.

<table>
<thead>
<tr>
<th>Table 11</th>
<th>Computational costs for extended Yale B database.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LBP-hist</td>
</tr>
<tr>
<td>CPU times (ms)</td>
<td>4.6</td>
</tr>
</tbody>
</table>

4. Conclusion

This paper presents a novel feature extraction method combining the proposed DI with the sparse RP-based DR to extract the discriminative and salient information from facial image. The proposed feature extraction method is very efficient to operate since it does not require any time-consuming training process due to its data-independent property. The extensive experimental results show that the proposed method not only achieves the high recognition rate on four representative FR databases containing both illumination changes and expression variations, but also presents strong robustness to occlusions, corruptions and disguises. Note that it is expected that the proposed feature extraction method can be applied to other biometric recognition tasks such as palm print and iris recognitions, and other image classification tasks such as natural scene categorization and texture classification. We hope our proposed method can inspire more work on extracting more informative features based on the random projection and compressed sensing theory.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government Ministry of Education, Science and Technology (MEST) (No. 2012 R1A2A4A01008384).

References


Chunfei Ma received the B.S. degree in Automation from Southeast University in 2008 and the M.S. degree in Flight Vehicle Design from Harbin Institute of Technology in 2010. He is currently pursuing the Ph.D. degree in Electrical Engineering at Korea University. His research interests are in the areas of pattern recognition, computer vision, and human computer interface.

June-Young Jung received the B.S. degree in Software from Seo-Kyeong University in 2009. He joined the Multimedia Communication Processing Lab in the Department of Electronic Engineering of Korea University in 2009. He is now a Ph.D. candidate at Korea University. His interests are computer vision, image stitching, and sensor fusion for object detection and recognition.

Seung-Jea Ko (M’88-SM’97-F’12) received the Ph.D. degree in 1988 and the M.S. degree in 1986, both in Electrical and Computer Engineering, from State University of New York at Buffalo, and the B.S. degree in Electronic Engineering at Korea University in 1980. In 1992, he joined the Department of Electronic Engineering at Korea University where he is currently a Professor. From 1988 to 1992, he was an Assistant Professor of the Department of Electrical and Computer Engineering at the University of Michigan-Dearborn. He has published over 171 international journal articles and 210 international conference proceedings. He also holds over 60 patents registered on video signal processing, computer vision, and multimedia communications.

He is the 1999 Recipient of the LG Research Award given to the Outstanding Information Technology and Communication Researcher. He received the Hae-Dong best paper award from the Institute of Electronics Engineers of Korea (IEEK) (1997), the best paper award from the IEEE Asia Pacific Conference on Circuits and Systems (1996), the research excellence award from Korea University (2004), and the technical achievement award from the IEEE Consumer Electronics (CE) Society (2012). He received a 15-year service award from the TPC of ICCE in 2014. He has served as the General Chairman of ITC-CSCC 2012 and the General Chairman of IEICE 2013. He is a member of the editorial board of the IEEE Transactions on Consumer Electronics. He is a Past President of the IEEK and the Vice-President of the IEEE CE Society, and an IEEE Fellow.