

# Robust Face Recognition Based on Part-Based Localized Basis Images

Jongsun Kim and Juneho Yi

School of Information & Communication Engineering,  
Sungkyunkwan University, Korea  
Biometrics Engineering Research Center  
{jskim, jhyi}@ece.skku.ac.kr

**Abstract.** In order for a subspace projection based method to be robust to local distortion and partial occlusion, the basis images generated by the method should exhibit a part-based local representation. We propose an effective part-based local representation method using ICA architecture I basis images that is robust to local distortion and partial occlusion. The proposed representation only employs locally salient information from important facial parts in order to maximize the benefit of applying the idea of “recognition by parts.” We have contrasted our representation with other part-based representations such as LNMF (Localized Non-negative Matrix Factorization) and LFA (Local Feature Analysis). Experimental results show that our representation performs better than PCA, ICA architecture I, ICA architecture II, LFA, and LNMF methods, especially in the cases of partial occlusions and local distortions.

## 1 Introduction

Subspace projection techniques represent a face as a linear combination of low rank basis images. They employ feature vectors consisting of coefficients that are obtained by simply projecting facial images onto their basis images. Therefore, the performance of face recognition methods using subspace projection is directly related to the characteristics of their basis images. Among popularly used techniques are PCA, ICA and FLD [1-5]. ICA can be applied to face recognition in two different representations: ICA architecture I and II [2]. The goal of this research is to develop an optimal set of basis images for face recognition robust to partial occlusions and local distortions.

Facial image representations based on different basis images are illustrated in Figure 1. PCA and ICA architecture II basis images, as shown in Figure 1 (a) and (b), respectively, display global properties in the sense that they assign significant weights to potentially all the pixels. This accords with the fact that PCA basis images are just scaled versions of global Fourier filters [12]. In contrast, ICA architecture I basis images are spatially more localized. This local property of ICA architecture I basis images makes the performance of ICA architecture I based recognition methods robust to partial occlusion and local distortion, such as local changes in facial expression, because spatially local features only influence small parts of facial

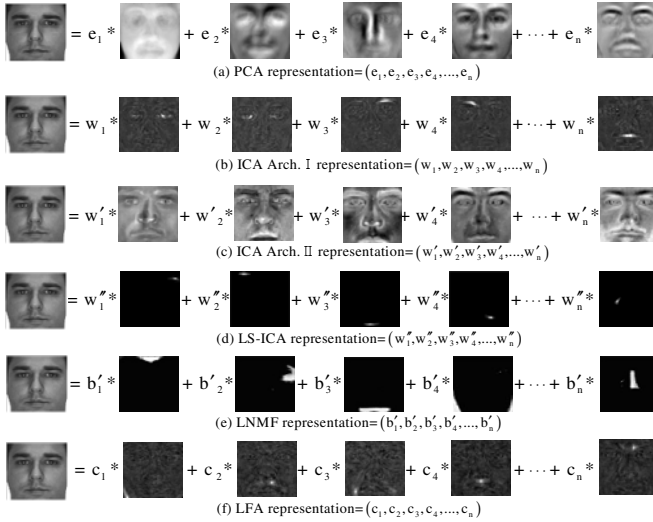
images. However, ICA architecture I basis images do not display perfectly local characteristics, in the sense that pixels that do not belong to locally salient feature regions still have some nonzero weight values. These pixel values in non-salient regions would appear as noise and contribute to the degradation of the recognition.

In order for a subspace projection based method to be robust to partial occlusions and local distortions, its basis images should effectively realize a part-based local representation. Local representation is essential to robustness to partial occlusions and local distortions because successful face recognition can be achieved by representing some important facial parts that correspond to feature regions such as eyes, eye brows, nose and lips. This “recognition by parts” paradigm [11] has been popular in the object recognition research because the approach can be successfully applied to the problem of object recognition with occlusion. Among representative part-based local representations are Local Feature Analysis (LFA) [6] and Local Non-negative Matrix Factorization (LNMF) [13] methods. The LFA method extracts local features based on second-order statistics. However, basis image from the LFA representation are not perfectly localized as shown in Figure 1 (f). Thus, pixels in non-salient regions degrade the recognition performance in the case of partial occlusions and local distortions. Recently, the LNMF method was reported in the literature, which led to an improved version of Non-negative Matrix Factorization (NMF) [11]. In the LNMF method, locality constraints are imposed on the factorized matrices from NMF in terms of sparsity in matrix components. They successfully localized the components in basis images. However, the locality constraints do not guarantee that meaningful facial features should be localized in their basis images. As an example of the LNMF representation in Figure 1 (e) illustrates, some LNMF basis images represent regions such as cheek, forehead and jaw that are not discriminant features for face recognition.

We propose new basis images based on ICA architecture I, called LS-ICA (locally salient ICA) basis images, where only locally salient feature regions are retained. The idea of “recognition by parts” can be effectively realized for face recognition using LS-ICA basis images since each LS-ICA basis image represent only locally salient regions. These regions correspond to important facial feature regions such as eyes, eye brow, nose and lips. Note that ICA architecture I produces basis images that are localized edge filters [12], and they correspond to meaningful facial feature regions.

Our method for face recognition is characterized by two ideas: The first is the creation of the LS-ICA basis images using a modified version of Kurtosis maximization to remove residual nonlocal modulation in ICA architecture I basis images; these are used to represent faces. The second idea is to use LS-ICA basis images in the decreasing order of class separability so as to maximize the recognition performance. Experimental results show that LS-ICA performs better than PCA, ICA architecture I, ICA architecture II, LFA, and LNMF, especially in the cases of partial occlusions and local distortions such as local changes in facial expression.

The rest of this paper is organized as follows. Section 2 briefly describes the ICA, LFA and LNMF methods that are most relevant to our research. We present the proposed LS-ICA method in section 3. Section 4 gives experimental results.



**Fig. 1.** Facial image representations using (a) PCA, (b) ICA architecture I, (c) ICA architecture II, (d) proposed LS-ICA, (e) LNMF and (f) LFA basis images: A face is represented as a linear combination of basis images. The basis images were computed from a set of images randomly selected from the AR database. Using LS-ICA basis images, the concept of “recognition by parts” can be effectively implemented for face recognition.

## 2 Related Work

### 2.1 ICA (Independent Component Analysis)

ICA is a widely used subspace projection technique that projects data from a high-dimensional space to a lower-dimensional space [2-4]. This technique is a generalization of PCA that decorrelates the high-order statistics in addition to the second-order moments. In this research, we compute ICA basis images using the FastICA algorithm [3] while other methods such as InfoMax [2] or Maximum likelihood [4] can also be employed.

The FastICA method computes independent components by maximizing non-Gaussianity of whitened data distribution using a kurtosis maximization process. The kurtosis measures the non-Gaussianity and the sparseness of the face representations [12]. The FastICA algorithm is briefly described as follows. Let  $\mathbf{S}$  be the vectors of unknown source signals and  $\mathbf{X}$  be vectors of observed mixtures. If  $\mathbf{A}$  is an unknown mixing matrix, then the mixing model can be written as  $\mathbf{X} = \mathbf{AS}$ . The task is to estimate the independent source signals  $\mathbf{U}$  by computing the separating matrix  $\mathbf{W}$  that corresponds to the mixing matrix  $\mathbf{A}$  using the following relation

$$\mathbf{U} = \mathbf{WX} = \mathbf{WAS}. \tag{1}$$

First, the observed samples are whitened. Let us denote the whitened samples by  $\mathbf{Z}$ . Then, we search for the matrix such that the linear projection of the whitened samples by the matrix  $\mathbf{W}$  has maximum non-Gaussianity of data distribution. The kurtosis of  $\mathbf{U}_i = \mathbf{W}_i^T \mathbf{Z}$  is computed as in equation (2) and the separating vector  $\mathbf{w}_i$  is obtained by maximizing the kurtosis [3].

$$kurt(\mathbf{U}_i) = \left| E\{(\mathbf{U}_i)^4\} - 3\left(E\{(\mathbf{U}_i)^2\}\right)^2 \right|. \tag{2}$$

ICA can be applied to face recognition in two different architectures [2]. The ICA architecture I considers the input face images,  $\mathbf{X}$ , as a linear combination of statistically independent basis images,  $\mathbf{S}$ , combined by an unknown matrix,  $\mathbf{A}$ . On the other hand, the ICA architecture II finds statistically independent coefficients that represent input images. The coefficients obtained by projecting input images onto the statistically independent basis images are not statistically independent. The ICA architecture II basis images display global properties as shown in Figure 1 (c). Since the kurtosis maximization yields sparseness of basis images, the ICA architecture I basis images are spatially localized edge filters [12]. However, they do not display perfectly local characteristics in the sense that pixels that do not belong to locally salient feature regions still have some nonzero weight values. These pixel values would contribute to the degradation of the recognition performance in the case of local distortion and partial occlusion.

**2.2 LFA (Local Feature Analysis)**

LFA defines a set of local topographic kernels that are derived from the principal component eigenvectors  $\mathbf{E}$  and coefficients  $\mathbf{D}$  according to covariance matrix  $\mathbf{S}$  using the following equation.

$$\mathbf{K} = \mathbf{E}\mathbf{D}^{-\frac{1}{2}}\mathbf{E}^T \text{ where } \mathbf{D}^{-\frac{1}{2}} = \text{diag}\left(\frac{1}{\sqrt{\lambda_i}}\right) \quad i=1,\dots,p \tag{3}$$

$\lambda_i$ 's are eigenvalues of the covariance matrix,  $\mathbf{S}$ . The rows of  $\mathbf{K}$  contain the kernels. The kernels are topographic in that they are indexed spatially. The number of kernels corresponds to the number of pixels in an image. The LFA uses a sparsification algorithm in order to reduce the dimensionality of the representation. The algorithm iteratively selects kernels that have the largest mean reconstruction error. We are concerned with basis images from the LFA method. They not perfectly localized and pixels in non-salient regions contribute to the degradation of the recognition performance.

**2.3 LNMF (Local Non-negative Matrix Factorization)**

The LNMF method [13] is a technique that improves the standard NMF method [11]. It is aimed at learning spatially localized, parts-based subspace representation of basis images by imposing additional constraints on the NMF basis. The following objective function is used to compute LNMF basis image:

$$D(X \parallel BH) = \sum_{i=1}^m \sum_{j=1}^n \left( X_{ij} \log \frac{X_{ij}}{[BH]_{ij}} - X_{ij} + [BH]_{ij} + \alpha U_{ij} \right) - \beta \sum_i V_i \tag{4}$$

where  $\alpha, \beta > 0$  are some constants,  $U = B^T B$ ,  $V = HH^T$  and  $B, H \geq 0$  means that all entries of basis images  $B$  and coefficients  $H$  are non-negative. The minimization of  $U = B^T B$  imposes both maximum sparsity in  $H$  and minimum redundancy between

different bases. On the other hand, by maximizing  $\sum_i V_{ii}$ , basis components that carry much information about the training images are retained. Refer to [13] for further justification of the objective function. The LNMF update rule for  $H$  uses square root as in equation (5) to satisfy the additional constraints. The update for  $B$  is identical to that of NMF [14].

$$H_{aj}^{(t+1)} = \sqrt{H_{aj}^{(t)} \sum_i B_{ai}^{T(t)} \frac{X_{ij}}{(B^{(t)} H^{(t)})_{ij}}} \tag{5}$$

Figure 1 (e) shows an example of the LNMF representation. As described earlier, the additional constraints only focus on locality and it is not necessarily guaranteed that meaningful facial features are localized in their basis images. We can see that some LNMF basis images represent regions such as cheek, forehead and jaw that are not discriminant features for face recognition.

### 3 The LS-ICA (Locally Salient ICA) Method

The computation of LS-ICA basis images consists of two steps: The first step is concerned with the creation of part-based local basis images based on ICA architecture I. The second is to order the basis images obtained in the order of class separability for good recognition performance.

The LS-ICA method creates part-based local basis images by imposing additional localization constraint in the process of the kurtosis maximization. The solution at each iteration step is weighted so that it becomes sparser by only emphasizing large pixel values. Let  $\mathbf{u}$  be a solution vector at an iteration step, we can define a weighted solution as  $\mathbf{b}$  where  $b_i = |u_i|^\alpha u_i$  and  $\mathbf{b} = \mathbf{b} / \|\mathbf{b}\|$ .  $\alpha > 1$  is a small constant. The kurtosis is maximized in terms of  $\mathbf{b}$  instead of  $\mathbf{u}$  as in equation (6).

$$kurt(\mathbf{b}) = \left| E\{(\mathbf{b})^4\} - 3\left(E\{(\mathbf{b})^2\}\right)^2 \right|. \tag{6}$$

A solution to the above function can be found by using the following update rules:

$$\mathbf{w}^{(t+1)} = E \left\{ |u_i|^\alpha \mathbf{Z} \left( |u_i|^\alpha \mathbf{w}^{T(t)} \mathbf{Z} \right)^3 \right\} \tag{7}$$

where  $\mathbf{w}$  is a separating vector and  $\mathbf{Z}$  contains whitened image samples. The resulting basis image is  $b_i = |u_i|^\alpha (\mathbf{w}^T \mathbf{Z})_i$ . As an alternative method, we would like to point out that other simple operations such as morphological operations can be employed to detect salient regions from ICA architecture I basis images.

We then compute a measure of class separability,  $r$ , for each LS-ICA basis vector and sort the LS-ICA basis vectors in the decreasing order of class separability [2]. The class separability,  $r$ , of a basis vector is defined as the ratio

$$r = \frac{\sigma_{between}}{\sigma_{within}} . \quad (8)$$

$\sigma_{between}$  and  $\sigma_{within}$  denote, respectively, the between-class variability and within-class variability of its corresponding projection coefficients of training images. We then create new LS-ICA basis images from the LS-ICA basis images selected in the decreasing order of the class separability. This way, we can achieve both dimensionality reduction and good recognition performance. The LS-ICA representation is based on the idea of localized edge filters that come from ICA basis images. The resulting basis images contain localized facial features that are discriminant for face recognition.

## 4 Experimental Results

We have used facial image databases such as the AT&T [8] and AR [9] and databases in order to compare the recognition performance of LS-ICA with that of PCA, ICA architecture I, ICA architecture II, LFA and LNMF methods. Figure 2 shows example images from these databases. For fair comparisons with the above methods, their basis images were also used in the decreasing order of class separability,  $r$ . We have computed recognition performances for three different distance measures (L1, L2, cosine) to see if there is any performance variation depending on the distance measure used [7].

In order to show the performance comparisons under occlusion, we have used the AT&T database. It consists of 400 images of 40 persons, 10 images per person. The images are taken against a dark homogeneous background and the subjects are in an up-right, frontal position with tolerance for some side movement. A set of 10 images for each person is randomly partitioned into five training images and five testing images. The occlusion is simulated in an image by using a white patch of size  $s \times s$  with  $s \in \{10, 20, 30\}$  at a random location as shown in Figure 3. Figure 4 compares the six representations under varying degrees of occlusion, in terms of the recognition accuracies versus the size  $s \times s$  of occluding patch. The LS-ICA and LNMF methods performed better than the other methods under partial occlusion, especially as the patch size increases. The LS-ICA method that only makes use of locally salient information from important facial parts achieved higher recognition rates than the LNMF method. The LFA and ICA architecture I methods appear influenced by pixels not belonging to salient regions. The ICA architecture I method showed better performance than the LFA method. Bartlett et al. [15] have also reported that the ICA representation performs better than the LFA representation in cases of facial expression analysis and face recognition. This experimentally shows that the ICA architecture I better represents some important facial parts than the LFA method.

The AR database contains 800 frontal facial images from 100 subjects. The number of images used for training and testing are 200 and 600 images, respectively. Test images contain local distortions and occlusions such as changes in facial expression

and sunglasses worn. Figures 5 shows the recognition performances for the AR database. The recognition rate of the LS-ICA method was consistently better than that of PCA, ICA architecture I , ICA architecture II , LNMF and LFA methods regardless of the distance measures used. We can see that the LS-ICA representation is an effective part-based local representation for face recognition robust to local distortion and partial occlusion.

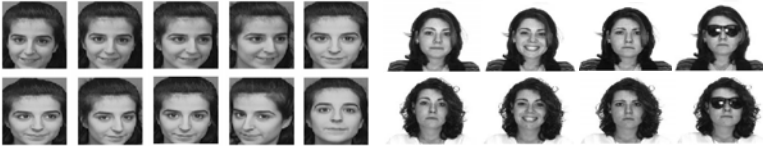


Fig. 2. Example images from AT&T (left) and AR (right) facial databases



Fig. 3. Example AT&T images having random occluding patches of sizes (from left to right) 10x10, 20x20, and 30x30

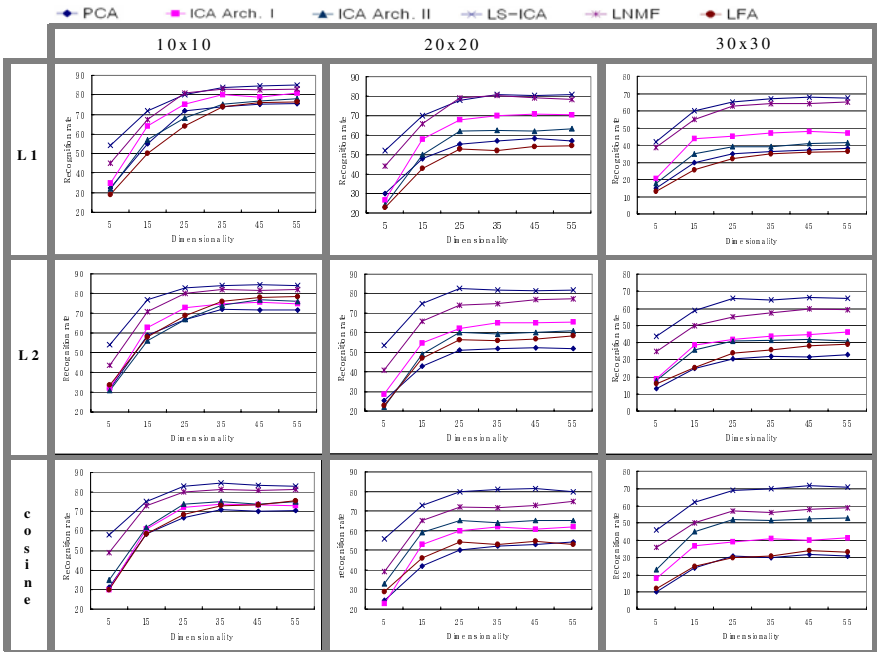
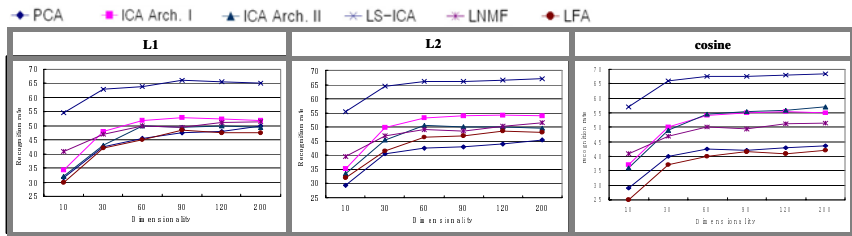


Fig. 4. Recognition performance versus the size (in 10x10, 20x20, 30x30) of occluding patches for PCA, ICA1, ICA2, LNMF, LFA, and the LS-ICA method for the AT&T database



**Fig. 5.** The recognition performance of PCA, ICA1, ICA2, LNMF, LFA and LS-ICA methods for the AR facial databases

## 5 Conclusions

Application of the idea of “recognition by parts” is necessary for the problem of face recognition under partial occlusion and local distortion. To maximize the benefit of “recognition by parts”, we have proposed an effective part-based local representation based on ICA architecture I basis images. The basis images are created by imposing additional localization constraint in the process of computing ICA architecture I basis images. The performance of the LS-ICA method was consistently better than other representative local representation based methods, especially in the cases of facial images that have partial occlusions and local distortions such as changes in facial expression.

## Acknowledgement

This work was supported by the Korea Science Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University and BK21.

## References

- [1] M. A. Turk and A. P. Pentland, “Eigenfaces for recognition,” *Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.
- [2] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, “Face Recognition by Independent Component Analysis,” *IEEE Trans. Neural Networks*, vol. 13, no. 6, pp. 1450-1464, 2002.
- [3] Aapo Hyvarinen and Erki Oja, “Independent component analysis: a tutorial,” [http://www.cis.hut.fi/~aapo/papers/IJCNN99\\_tutorialweb/](http://www.cis.hut.fi/~aapo/papers/IJCNN99_tutorialweb/), 1999.
- [4] A. Hyvärinen, “The Fixed-point Algorithm and Maximum Likelihood Estimation for Independent Component Analysis,” *Neural Processing Letters*, vol. 10, pp. 1-5, 1999.
- [5] P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection,” *IEEE PAMI*, vol. 19, no. 7, pp. 711-720, 1999.



- [6] P. Penev and J. Atick, "Local Feature Analysis: A general statistical theory for object representation," *Network: Computation in Neural Systems*, vol. 7, no. 3, pp. 477-500, 1996.
- [7] B. A. Draper, K. Baek, M. S. Bartlett, and J. R. Beveridge, "Recognizing faces with PCA and ICA," *Computer Vision and Image Understanding*, vol. 91, no. 1, pp. 115-137, 2003.
- [8] <http://www.uk.research.att.com/facedatabase.html>.
- [9] A. M. Martinez and R. Benavente, "The AR face database," CVC Tech, 1998.
- [10] A. P. Pentland, "Recognition by parts," *IEEE Proceedings of the First International Conference on Computer Vision*, pp. 612-620, 1987.
- [11] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788-791, 1999.
- [12] A. J. Bell and T. J. Sejnowski, "The Independent Components of Natural Scenes are Edge Filters," *Vision Research*, vol. 37, no. 23, pp. 3327-3338, 1997.
- [13] S. Z. Li, X. W. Hou, H. J. Zhang, "Learning Spatially Localized, Parts-Based Representation," *Computer Vision and Pattern Recognition*, vol. 1, pp. 207-212, 2001.
- [14] S. Wild, J. Curry, A. Dougherty, "Motivating Non-Negative Matrix Factorizations," *In Proceedings of the Eighth SIAM Conference on Applied Linear Algebra*, July 2003.
- [15] M. S. Bartlett, *Face Image Analysis by Unsupervised Learning*, Kluwer Academic Publishers, 2001.