An Efficient Face Recognition Algorithm Based on Robust Principal Component Analysis

Ziheng Wang, Xudong Xie
TNLIST and Department of Automation
Tsinghua University
Beijing, China
zihengwang.thu@gmail.com, xdxie@tsinghua.edu.cn

Abstract—**In this paper, an efficient face recognition algorithm is proposed, which is robust to illumination, expression and occlusion. In our method, a human face image is considered as a multiplication of a reflectance image and an illumination image. Then, this illumination model is used to transfer input images. After the transformation, the robust principal component analysis is employed to recover the intrinsic information of a sequence of images of one person. Finally, a new similarity metric is defined for face recognition. Experiments based on different databases illustrate that our method can achieve consistent and promising results.**

Keywords—**Face recognition, illumination model, robust principal component analysis (RPCA).**

I. INTRODUCTION

Compared with other biometric features such as fingerprint and iris, human faces have clear advantages of being natural and non-intrusive. Therefore, face recognition has attracted significantly increased attention from both academic and industrial communities in last decade. It has been used in a wide variety of applications, such as credit card verification, criminal identification, scene surveillance, etc.

In order to be robust in practical use, a face recognition system must achieve acceptable performance in uncontrolled environmental conditions, such as varying illumination, different expressions, ages, poses, etc. However, it is still a great challenge to most of the existing algorithms [1, 2]. A solution is to use more images captured under various conditions per subject for training. Linear discriminant analysis (LDA) [3] seeks to find a linear transformation that maximizes the between-class scatter and minimizes the within-class scatter of the training set. Locality preserving projection (LPP) [4] obtains a face subspace that best describes the essential face manifold structure, and preserves the local structure of the image space. Sparse representation-based classification (SR) [5] formulates the recognition problem as one of classifying among multiple linear regression models obtained from the sparse signal representation.

In this paper, an efficient method is proposed for robust face recognition under various conditions. In our method, an input image is considered as a multiplication of a reflectance image and an illumination image [6]. Base on robust principal component analysis (RPCA) [7], we can eliminate the effects of not only illumination but also occlusions. In other words, the

intrinsic information of a human face can be recovered based on a sequence of his images. We define a new similarity metric, which can be used for measuring the similarity between a query image and a sequence of images of one person. Then, this similarity metric can be used for face recognition. Experiments are performed based on different databases, i.e. the AR database [8], the Yale database [9], the YaleB database [10], and the PIE database [11]. Experimental results show that our proposed method outperforms other methods, i.e. PCA [12], LDA, LPP and SR, in all cases.

II. FACE RECOGNITION BASED ON RPCA

A. Illumination Model

In [6], a face image is supposed to be a multiplication of a reflectance image and an illumination image:

$$I(x, y) = L(x, y) \cdot R(x, y) \tag{1}$$

where $I(x, y)$ is an input image, and $R(x, y)$ and $L(x, y)$ denotes the reflectance image and the illumination image, respectively. [6] proposed a method to separate the reflectance image and the illumination image. However, this method cannot handle local shape variations, e.g. expression changes. Furthermore, for an occluded image, e.g. a man with sunglasses, its performance degrades greatly. Therefore, we cannot apply the method proposed in [6] for face recognition directly in real applications. In our method, we just apply this illumination model to perform the logarithmic transformation for an input image.

B. Robust Principal Components Analysis

The algorithm RPCA is aimed to recover the low-rank matrix A from the corrupted observations D = A + E, where corrupted entries E are unknown and the errors can be arbitrarily large but assumed to be sparse. One of its modeling with face images is proposed in [13].

This model directly adds errors to the original images. However, as discussed in Section 2.1, the illumination images (i.e. the error images) are more reasonable and practical to be multiplied to the reflectance image (i.e. the original image) to get the observed corrupted image. Therefore, we convert them into the log domain. Denote I_i^j, A_i^j, E_i^j by the logarithms of the corrupted observed face image, the original face image and the

error of the jth image of the ith subject, respectively. We can obtain that $I_i^j = A_i^j + E_i^j$. Define

$$vec: \mathcal{R}^{w \times h} \to \mathcal{R}^{(w \times h) \times 1}$$

as the function that transforms a $w \times h$ image matrix into a $(w \times h) \times 1$ vector by stacking its columns, then we have $vec(I_i^j) = vec(A_i^j) + vec(E_i^j)$.

We consider all images in the log domain. Assume that we are given m subjects and each subject has n images, we can define that:

$$D_i := [vec(I_i^1)| \cdots |vec(I_i^n)] = A_i + E_i \quad (2)$$
$$A_i := [vec(A_i^1)| \cdots |vec(A_i^n)] \in \mathcal{R}^{m \times n} \quad (3)$$
$$E_i := [vec(E_i^1)| \cdots |vec(E_i^n)] \in \mathcal{R}^{m \times n} \quad (4)$$

where $i = 1, \cdots, m$, D_i is formed by stacking the n image vectors of the ith subject, A_i and E_i are the corresponding original images matrix and the error matrix, respectively. Since all images of the same subject are approximately linearly correlated, A_i can be regarded as a low-rank matrix and E_i is large but sparse [13]. In this case, as proven in [7], A_i and E_i can be efficiently recovered from D_i. Figure 1 shows the recovered original images from the set of images of one subject in the AR database.

C. Face Recognition Based on RPCA

In this section, we will define a new similarity metric, which can be used for measuring the similarity between a query image and a sequence of images. Then, this similarity metric can be used for face recognition.

Given m persons and each subject has n training images $\{I_i^j | i = 1, \cdots, m; j = 1, \cdots, n\}$, we should classify a query image I. The basic idea of our algorithm is to insert the test image into the observation matrix D of each subject in the training set, and then to recover its error E. Consider the ith subject in the training set, we have

$$D_i = [vec(I_i^1)| \cdots |vec(I_i^n)|vec(I)].$$

Then the error matrix can be generated by RPCA:

$$E_i = [vec(E_i^1)| \cdots |vec(E_i^n)|vec(E_i)],$$

where E_i stands for the error of a test image. Here we define a similarity metric:

$$F(I, I_i) := |vec(E_i)| \quad (5)$$

where $F(I, I_i)$ denotes the similarity between the input image I and a class of images which belong to the ith subject. And $|vec(E_i)|$ is the Euclidean distance of $vec(E_i)$ in our algorithm.

If the query image I belongs to the ith subject, D_i contains the images of the same subject so that the assumption that A_i is linearly correlated and low-rank can be satisfied. In this case, the parameters of E_i should be 'small'. In other words, $F(I, I_i)$ should be small. Otherwise, the value of $F(I, I_i)$ is relatively large. Figure 2 illustrates the recovered original image of a test sample when it is combined with different training subjects. For face recognition, the test image I is recognized as the subject which has the smallest value of $F(I, I_i)$.
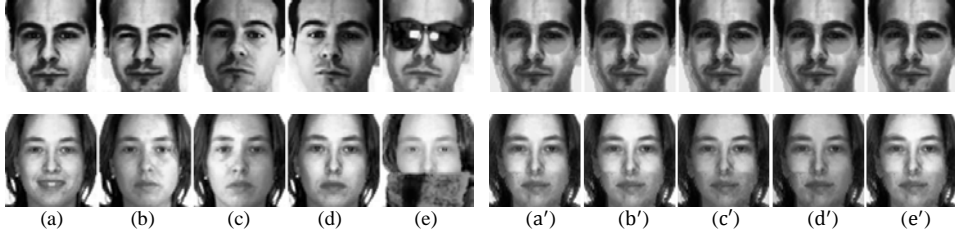

Figure 1. Images recovered by RPCA. (a), (b), (c), (d), (e) are five images of the same subject. (a′), (b′), (c′), (d′), (e′) are the corresponding recovered images by RPCA combined with illumination model. The sunglasses and scarf in (e) are successfully removed.
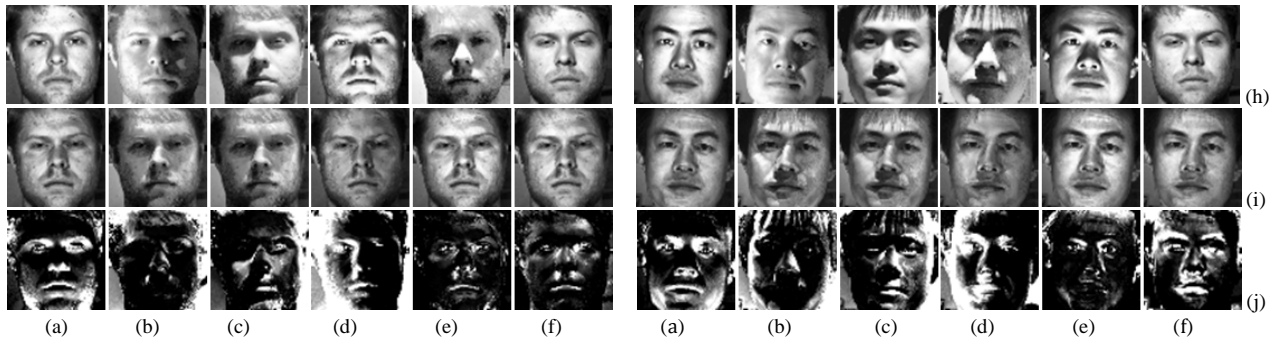

Figure 2. Recovered images when combined with different training subjects. (a) – (e) are the training images of the subject, (f) is the query image. (h) The corrupted images, (i) the recovered images and (j) the error images. The F values (see Eq. (5)) between the query image and the two sets of images of different subjects are 11.4 (left column) and 22.1(right column), respectively.

III. EXPERIMENTAL RESULTS

In this section, we will evaluate the performances of the proposed algorithm, namely FRPCA, for face recognition based on different face databases. All testing images are normalized to a resolution of 64×64 pixels based on the eye locations, and the color ones are converted to grayscale. To enhance the global contrast of the images and reduce the effect of uneven illumination, histogram equalization is applied to all the images.

A. Face Recognition with Different Methods

In this section, our proposed algorithm is tested on four standard face image databases, i.e. the Yale, YaleB, AR, and PIE face database. Face images in different databases are captured under different conditions, such as variable lighting conditions, facial expressions, perspectives and with/without glasses. The number of distinct subjects and the total number of the respective databases are tabulated in Table 1. In each database, five images of each subject are randomly chosen as training images. Others constitute the testing set. Besides, images with occlusions, i.e. sunglass and scarf, in the AR database are not included for evaluation in this section and Section *D*.

TABLE I.   FACE DATABASES USED IN THESE EXPERIMENTS

Face database	Yale	YaleB	AR	PIE
Number of subjects	15	10	121	68
Number of total images	165	640	847	1768

Our method is compared with other algorithms, i.e. PCA, LDA, LPP and SR. The accuracy rates of different algorithms are shown in Table 2.

TABLE II.   ACCURACY RATE OF DIFFERENT ALGORITHMS.

%	Yale	YaleB	AR	PIE
PCA	86.67	84.75	59.92	92.91
LDA	90.00	96.44	85.12	97.62
LPP	85.56	89.83	73.97	92.98
SR	90.00	90.17	55.37	96.92
FRPCA	**93.33**	**98.47**	**89.26**	**97.86**

These results indicate that our method can always achieve the best performance in all cases. Specifically, our algorithm outperforms others greatly on the databases which are under large illumination and expression variations, such as the YaleB database and AR database. This is because on one hand, the illumination model explains face images in a more practical way, and on the other hand, RPCA can eliminate the shape variations better than other methods.

B. Face Recognition with Different Number of Training Images per Subject

In Section *A*, the number of training images for each subject is 5. In this section, we will evaluate the performance when the number of training images per subject varies. Based on the Yale database, k images are randomly selected as the training set, and the others form the testing set, where $k = 1, \cdots, 10$.

Figure 3 shows the variation tendency of different face recognition algorithms.

We can see that all algorithms achieve excellent performance when enough training samples are obtained. On the contrary, when the number of training samples is small, the recognition rate of our method is slightly less than LPP, LDA and SR. This is because in this case, the matrix D itself in RPCA tends to be low-rank. Thus we cannot obtain the exact value of E. The similarity metric won't work normally in this case. However, our algorithm still maintains better performance over other algorithms in most cases，and as the number of training samples increases, it shows a faster growth of the accuracy rate.
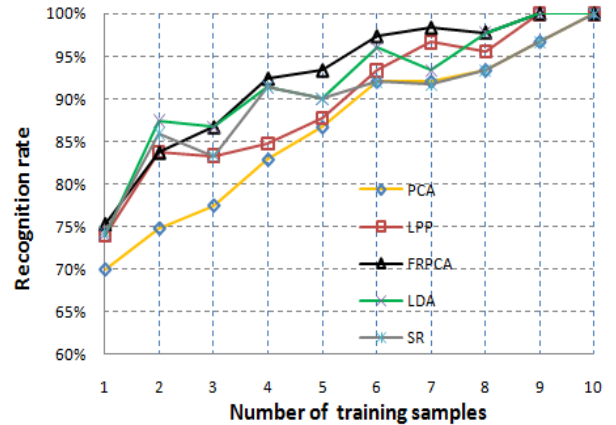

Figure 3.   Accuracy rate variation of different algorithms with different number of training samples per subject.

C. Performance on Images with Occlusions

Occlusions such as sunglasses and scarf on human faces may not be avoided in many practical cases. Robust and efficient face recognition algorithms must also achieve good performance when faces are partially occluded. We use the occluded face images in the AR database to test the performance of different algorithms. Some cropped images are shown in Figure 4. These occluded images of the database form the testing set and five images without occlusion of each subject are selected as the training set. Since not all the subjects in the AR database have all these occluded face images, only 104 subjects are selected for the evaluation in this section. Experimental results are shown in Table 3.

From these results, we can see that occluded images greatly reduce the performances of PCA, LDA, LPP and SR, where the best recognition rate is only about 27%. At the same time, a very high recognition rate can still be maintained by our algorithm. This is because the former four algorithms cannot remove the disturbance of the occlusion, while our algorithm is very robust to the occlusion because RPCA can approximately totally remove the occlusion, as shown in Figure 1.

TABLE III.   ACCURACY RATE OF DIFFERENT ALGORITHMS WITH OCCLUDED TESTING IMAGES.

%	PCA	LDA	LPP	SR	FRPCA
AR	18.27	27.40	17.15	6.3	**28.81**

Figure 4. Some cropped images with occlusions in the AR Database. Each subject used in our algorithm has three images with sunglasses and three images with scarf. These six images of each subject form the testing set.

TABLE IV. COMPARISON OF FRPCA WITH AND WITHOUT ILLUMINATION MODEL.

%	Yale	YaleB	AR	PIE
Without illumination model	92.22	86.07	74.38	97.46
With illumination model	**93.33**	**98.47**	**89.26**	**97.86**

D. Comparison of FRPCA with/without Illumination model

One of the key improvements of our algorithm is combining the illumination model with RPCA. In this section we will evaluate the performance of our algorithm with and without the preprocessing, which converts images into the log domain. Table 4 illustrates results on different databases.

These results suggest large improvement of our algorithm FRPCA when combined with the illumination model. As discussed in Section 2, errors are more practical to be multiplied by the original image. In the log domain, the errors are transferred into the addition of the origin, and therefore RPCA can achieve a better performance with this better model. From Table 4, we can observe that more obvious improvements of recognition rate can be seen on the YaleB database and AR database, of which the training samples are captured under larger varying illumination and expressions than those of other databases.

IV. CONCLUSION

In this paper, we have proposed a novel method, namely FRPCA, for face recognition. In our method, an input image is considered as a multiplication of a reflectance image and an illumination image. Then based on this illumination model, images are transferred into log domain. Considering the characteristics of RPCA, we define a new similarity metric, which can effectively measure the similarity between a query image and a class of images belonging to one person. Good performances can be achieved even in cases of variable conditions (such as poses, illumination and facial expression). Especially, for the occluded images, our method can achieve a better performance than other methods. In our future research, this method can be further improved to tackle more facial variations and to be used in real face recognition applications.

REFERENCES

[1] X. Xie and K. M. Lam, "Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image", *IEEE Trans. Image Processing*, vol. 15(7), pp.2481-2492, 2006.

[2] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, "Face recognition: a literature survey", *ACM Comput. Survey*, vol. 35(4). pp. 399-458, 2003.

[3] P. N. Belhumer, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, no. 7, pp. 711-720, 1997.

[4] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27(3), pp. 328-340, 2005.

[5] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31(2), pp. 210-227, 2009.

[6] Y. Weiss, "Deriving intrinsic images from image sequences," *Proc. Ninth IEEE Int'l Conf. Computer Vision*, pp. 68-75, July 2001.

[7] J. Wright, A. Ganesh, S. Rao, and Y. Ma. "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," *Journal of the ACM*, submitted for publication.

[8] A.M. Martinez, R. Benavente, The AR face database, CVC Technical Report no. 24, June 1998.

[9] Yale University, http://cvc.yale.edu/projects/yalefaces/yalefaces.html, 1997

[10] Yale University, http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html, 2001.

[11] T. Sim, S. Baker, M. Bsat, "The CMU pose, illumination, and expression (PIE) database," *Proceedings of Fifth International Conference on Automatic Face and Gesture Recognition*, May 2002.

[12] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Machine Intell.* vol. 19(7), pp. 711-720, 1997.

[13] Y. Peng, A. Ganesh, J. Wright, and Y. Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2010.