

Robust Face Recognition via Occlusion Detection and Masking

Tan Guo^{1, a}, Xiao Heng Tan¹ and Chao Chen Xie¹

¹College of Communication Engineering, Chongqing University, 400044, Chongqing, China

Abstract. Sparse representation-based classification (SRC) method has demonstrated promising results in face recognition (FR). In this paper, we consider the problem of face recognition with occlusion. In sparse representation-based classification method, the reconstruction residual of test sample over the training set is usually heterogeneous with the training samples, highlighting the occlusion part in test sample. We detect the occlusion part by extracting a mask from the reconstruction residual through threshold operation. The mask will be applied in the representation-based classification framework to eliminate the impact of occlusion in FR. The method does not assume any prior knowledge about the occlusion, and extensive experiments on publicly available databases show the efficacy of the method.

1 Introduction

Face recognition (FR) is one of the most intensively investigated topics in biometrics. It has a wide range of applications, e.g., information security, access control, smart payment and law enforcement [1]. A common face recognition system mainly consists of two modules, namely feature extraction and classifier design. In the past decades, many algorithms have been proposed for the two aspects. For feature extraction, classical dimensionality reduction methods include Principle Component Analysis (PCA) [2] and Linear Discriminative Analysis (LDA) [3]. These methods aim to construct projections that can map the high-dimensional face image into a lower-dimensional space to get a compact representation of face image. The most used classifiers are Nearest Neighbor (NN) and Nearest Feature Subspace (NFS) [4].

Generally speaking, occlusion poses a significant obstacle in a robust, real-world face recognition system [5-6]. This difficulty is mainly due to the error in query sample incurred by occlusion. In recent years, representation based methods for classification [6, 11, 14] has attracted much attention due to their powerful performance in FR. The most representative approach is sparse representation-based classification (SRC). Wright et al. [6] exploited the discriminative nature of sparse representation via l_1 -norm regularization. The method has shown great potential in handling various occlusions, such as real disguise, continuous occlusion, random pixel corruption, etc. SRC triggers many following works, including the l_1 -graph for image classification [7], kernel based SRC [8], Gabor feature based SRC [9], and robust sparse coding [10].

Meanwhile, Naseem et al. [11] proposed a linear regression classification (LRC) algorithm using the concept that patterns from a single-object class lie on a linear subspace. Luan et al. [12] indicated that error image derived from the LRC was a better choice than original image to identify occluded regions, and presented an appearance-based method for face recognition based on LRC and level set method to deal with contiguous occlusion. Sparse error of Robust PCA was also extracted for FR under varying illumination and occlusion [13].

Zhang et al. [14] indicated that it is the collaborative representation but not the l_1 -norm sparsity that makes SRC powerful for face classification, and present a collaborative representation based classification (CRC) method. CRC aims to represent the query sample using all the training samples with a weaker l_2 -norm regularization, and then perform classification according to the contribution of each class in representing the query sample. Nevertheless, the estimated coefficients of an occluded test sample may be corrupted and unstable.

In this paper, we present a novel method for robust face recognition in the presence of occlusion, such as block occlusion, pixel occlusion and real disguise. As introduced above, CRC actually performs a separation of query sample into the parts similar and dissimilar with the training set. The occlusion part in query samples is considered to be heterogeneous with the training samples, and thus can not be well reconstructed by the training samples, and is left in the reconstruction residual. We exploit the reconstruction residual to detect occlusion, and get a mask for the test sample. The mask will be used in the CRC scheme to

^a Corresponding author: tanguo@cqu.edu.cn

isolate the side effect of occlusion in FR.

The remainder of this paper is organized as follows: In section 2, we introduce the CRC and the proposed method. Section 3 presents the experimental results. The conclusion is drawn in Section 4.

2 CRC and Proposed method

2.1. CRC

CRC was proposed by Zhang et al. [14]. The work argued the function of collaborative representation in SRC [6], and empirically showed that the non-sparse l_2 -norm regularization could outperform the state-of-the-art SRC based on l_1 -norm regularization, both in computational complexity and classification accuracy. We first review the typical CRC algorithm.

Assume there are C distinguished training classes with n_i samples in each class. Let $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2 \dots \mathbf{D}_C]$, where $\mathbf{D}_i \in \mathbb{R}^{m \times n_i}$ is the data matrix of the i th class. Each column of \mathbf{D}_i denotes a training sample of the i th class. For a test sample $\mathbf{y} \in \mathbb{R}^m$, CRC uses all the samples in the training set to collaborative represent it as

$$\mathbf{y} = \mathbf{D}\mathbf{X} \quad (1)$$

Considering that the dictionary \mathbf{D} can be under-determined, they use l_2 -norm to regularize the solution. The CRC aims to solve the following problem.

$$\mathbf{X} = \arg \min_{\mathbf{X}} \{ \|\mathbf{y} - \mathbf{D}\mathbf{X}\|_2^2 + \mu \|\mathbf{X}\|_2^2 \} \quad (2)$$

The solution of (2) can be analytically derived as $\mathbf{X} = (\mathbf{D}^T \mathbf{D} + \mu \cdot \mathbf{I})^{-1} \mathbf{D}^T \mathbf{y}$. The class-specified representation residual $\|\mathbf{y} - \mathbf{D}_i \mathbf{X}_i\|_2$ together with the coefficient \mathbf{X}_i are used for classification.

$$\text{identity}(\mathbf{y}) = \arg \min_i \{ \|\mathbf{y} - \mathbf{D}_i \mathbf{X}_i\|_2 / \|\mathbf{X}_i\|_2 \} \quad (3)$$

Where \mathbf{X}_i is the representation coefficient vector in \mathbf{X} associated with training class \mathbf{D}_i .

2.2 Proposed method

As mentioned earlier, the goal of the proposed method is to detect and eliminate the occlusion part in test sample. Firstly, we need to detect the occlusion part in test sample. As discussed in [6, 12, 14], in representation based classification methods, such as LRC, the reconstruction residual was a better choice than original image to identify occluded regions. Similarity, the reconstruction residual in CRC also has such property. The occluded parts in query sample are always heterogeneous with the training samples, highlighting the occlusion part in query sample. Figure 1 shows an example for distribution of intensity values of reconstruction residual by CRC. We can find most of the values are near 0, in the interval $[-0.005, 0.005]$. The large values in the reconstruction residual are known to be outliers in query sample that can not be well fitted by the training samples. Based on this, we propose to extract the occluded parts in query sample from the reconstruction residual of CRC model.

Firstly, we use all the training samples to represent the test sample, and solve the following problem

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \{ \|\mathbf{y} - \mathbf{D}\mathbf{X}\|_2^2 + \eta \|\mathbf{X}\|_2^2 \} \quad (4)$$

Notably, we set η a relatively large value to relax the regularization on the reconstruction fidelity, which helps for the extraction of occlusion part. Then we get the reconstruction residual $(\mathbf{y} - \mathbf{D}\hat{\mathbf{X}})$ of \mathbf{y} . The pixels whose value meet the condition in (5) is assigned as 1, otherwise 0.

$$m = \begin{cases} 1 & |y - \mathbf{D}\hat{\mathbf{X}}| < \sigma \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

σ in (5) is a threshold needed to be specified. This way, we can get a mask of the test sample. The zero entities in m indicate the outliers in test sample. Figure 2 shows an example of this case. Figure 2(b) is a mask extracted from the reconstruction residual of Figure 2(a), and Figure 2(c) shows the test sample with mask, from which we can see that the occlusion part is well eliminated.

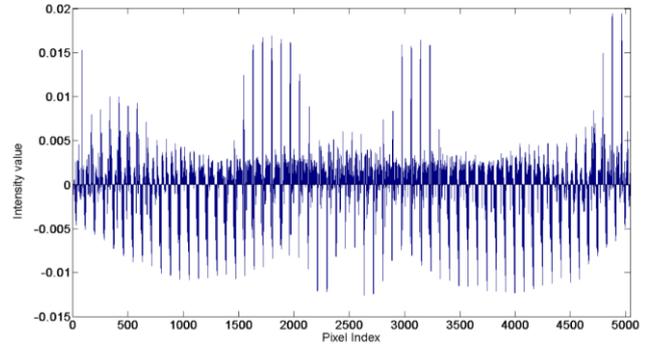


Figure 1. An example for distribution of intensity values of reconstruction residual in CRC.

Once the binary mask of test sample obtained, we solve the following regularized least square problem with mask.

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \left\{ \frac{1}{2} \left(\|\text{diag}(\mathbf{m})(\mathbf{y} - \mathbf{D}\mathbf{X})\|_2^2 + \lambda \|\mathbf{X}\|_2^2 \right) \right\} \quad (6)$$

Let $\mathbf{M} = \text{diag}(\mathbf{m})$, we have

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \left\{ \frac{1}{2} \left(\|\mathbf{M}(\mathbf{y} - \mathbf{D}\mathbf{X})\|_2^2 + \lambda \|\mathbf{X}\|_2^2 \right) \right\}$$

The closed form solution of above equation can be derived by taking partial derivations and setting them equal to $\mathbf{0}$ as

$$-(\mathbf{M}\mathbf{D})^T (\mathbf{M}\mathbf{y} - \mathbf{M}\mathbf{D}\hat{\mathbf{X}}) + \lambda \hat{\mathbf{X}} = \mathbf{0} \quad (7)$$

The solution is

$$\hat{\mathbf{X}} = (\mathbf{D}^T \mathbf{M}^T \mathbf{M} \mathbf{D} + \lambda \cdot \mathbf{I})^{-1} \mathbf{D}^T \mathbf{M}^T \mathbf{M} \mathbf{y} \quad (8)$$

Decision is made in favor of the class with the minimum regularized residual.

$$\text{identity}(\mathbf{y}) = \arg \min_i \left\{ \left\| \mathbf{M}(\mathbf{y} - \mathbf{D}_i \hat{\mathbf{X}}_i) \right\|_2 / \left\| \hat{\mathbf{X}}_i \right\|_2 \right\} \quad (9)$$

In summary, the main steps of our method are as follows.

Step1. Input: matrix of training samples $\mathbf{D} \in \mathbb{R}^{m \times n}$, a test sample with occlusion $\mathbf{y} \in \mathbb{R}^m$.

Step2. Normalize the columns of \mathbf{D} and \mathbf{y} to have unit l_2 -norm.

Step3. Use Eq. (4) and Eq. (5) to estimate the pixels shaded in \mathbf{y} , and get a mask for it.

Step4. Solve the regularized least square problem with binary mask in Eq. (6).

Setp5. Give the identity of the test sample y on the basis of Eq. (9).

The essence of our method is to detect and eliminate the impact of occlusion in query sample, and use the clear part to perform classification.

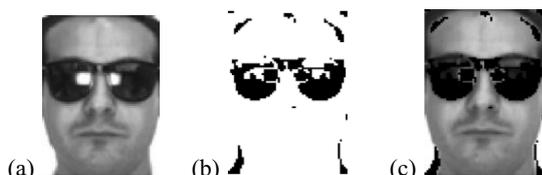


Figure 2. (a) A test sample with sunglasses. (b) the binary mask extracted from (a). (c) test sample with mask.

3 Experiments

To verify the performance of our algorithm in dealing with occlusion, we use four state-of-the-art algorithms for comparison. They are the Nearest Neighbor (NN), LRC [11] and CRC [14] and SRC [6]. All the Matlab implementations are on a PC with Intel core i5 3.3GHz CPU, 8 GB RAM, and Windows 7 operating system.

3.1 Recognition with random pixel occlusion

In this experiment, we test the robustness of the proposed method in the presence of pixel corruption. We use the Extended Yale B database for this experiment. The Extended Yale B database [15] consists of 2414 frontal face images of 38 individuals under various laboratory-controlled lighting conditions. Subset 1 consists of 266 images (seven images per subject) under normal lighting conditions. Subsets 2 includes 12 images per subject, characterize slight illumination variations (as shown in Figure 3). For Subset 2, we replace 10%~60% of its pixels by uniformly distributed random values within $[0,255]$ (as shown in Figure 4). The corrupted pixels are randomly chosen for each query image and the location are unknown to the algorithm. All images are cropped and resized to 64×56 . Table 1 reports the recognition results of each algorithm. From this table, one can see our method classifies all the test samples correctly even in the case of 60% of the pixels are occluded.



Figure 3. Two Subsets of Extended Yale B database. First row illustrates images from Subset 1, and second row Subset 2.

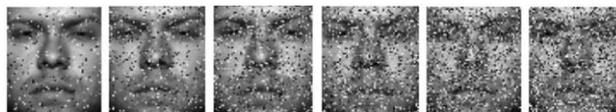


Figure 4. Test samples under varying level of random pixel occlusion, from left 10% to right 60%.

Table 1. Recognition results on the Extended Yale B database under varying level of random pixel occlusion using Subset 1 as training set.

	0%	10%	20%	30%	40%	50%	60%
NN	99.3	99.3	99.1	98.6	95.6	84.6	79.2
LRC [11]	100	100	100	100	99.1	94.3	87.7
CRC [14]	100	100	100	100	100	99.6	99.1
SRC [6]	100						
Ours	100						

3.2 Recognition with contiguous occlusion

The Extended Yale B database is used in this experiment. We use Subset 1 for training and Subsets 2 for testing. Various levels of contiguous occlusion are simulated by replacing a square block of each query image with an unrelated image (as shown in Figure 5.). The location of occlusion is randomly chosen and is unknown to the algorithms. Compared to random pixel corruption, contiguous occlusion is certainly a worse type of errors. As reported in Table 2, our algorithm outperforms other competing methods in the presence of contiguous occlusion.

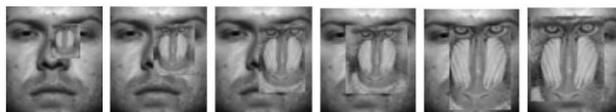


Figure 5. Test samples under varying level of contiguous occlusion, from left 10% to right 60%.

Table 2. Recognition results on the Extended Yale B database under varying level of contiguous occlusion using Subset 1 as training set.

	0%	10%	20%	30%	40%	50%	60%
NN	99.3	99.1	96.3	96.3	84.9	73.9	45.6
LRC [11]	100	100	99.8	99.1	89.5	77.2	49.1
CRC [14]	100	100	99.6	98.9	93.0	77.6	28.5
SRC [6]	100	100	100	100	100	90.1	67.1
Ours	100	100	100	100	100	92.3	68.4

3.3 Recognition with real disguise occlusion

In this section, experiments are carried out using a subset from the AR database [16] consisting of 100 individuals, 50 male and 50 female. 799 images of non-occluded frontal views with various facial expressions are used for training, while two separate sets of 200 images are selected for testing [12]. The first test set contains images of the subjects wearing sunglasses, which occlude roughly 25% of the face image. The second test set considered includes images of the subjects wearing a

scarf, which occludes roughly 40% of the face image. All images are cropped and resized to 84×60 (as shown in Figure 6).

We also present the recognition accuracy of SRC and LRC using partition scheme [6,11]. Each test sample was partitioned into 8 regions. From the recognition results in Table 3, It can be seen that our algorithm has a higher recognition rate comparing with other contrastive methods.

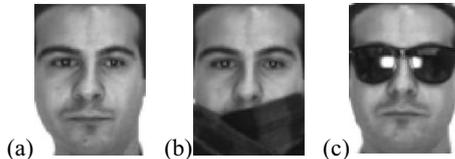


Figure 6. The training and test samples with sunglasses and scarf in AR database. (a) is the training sample, and (b)-(c) for test.

Table 3. Recognition results on the AR database with real disguise occlusion

	Scarf	Sunglasses
NN	12.0	69.0
CRC	90.0	66.5
LRC	11.5	68.0
LRC(partitioned)	73.0	93.0
SRC	60.5	87.0
SRC(partitioned)	92.0	94.0
Ours	92.0	99.5

4 Conclusions

In this paper, we present a novel method for robust face recognition in the presence of occlusion. We exploit the reconstruction residual of CRC model to detect the occlusion in test sample, and get a binary mask of the occlusion. Then the mask is used in the CRC scheme to isolate the side effect of occlusion in FR. Experimental results on Extended Yale B and AR database demonstrate that our algorithm performs well in dealing pixel occlusion, block occlusion and real disguise. The method does not assume any prior knowledge about the occluded regions, and can be analytically solved, and thus can be applied in a robust, real-time face recognition system.

Acknowledgments. This work is supported by National Nature Science Foundation of China (No.61571069), the open research fund of Chongqing key Laboratory of Emergency Communications and Chongqing University Postgraduates' Innovation Project (No.CYB15030).

References

1. Jafri, R., Arabnia, H. R. A survey of face recognition techniques. *Jips*, 11(19) (2009) 29-33
2. Turk, M., Pentland, A, Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3 (1) (2007) 71-86

3. Belhumeur, P. N., Hespanha, J. P., Kriegman, D. J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19 (7) (1997) 711-720
4. Chien, J. T., Wu, C. C, Discriminant waveletfaces and nearest feature classifiers for face recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24 (12) (2002) 1644-1649
5. Fidler, S., Sko aj, D., Leonardis, A. Combining reconstructive and discriminative subspace methods for robust classification and regression by subsampling. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28 (3) (2006) 337-350
6. J.Wright, A.Y.Yang, A.Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 31 (2) (2009) 210–227
7. B. Cheng, J. Yang, S. Yan, Y. Fu, and T. Huang. Learning with 11-graph for image analysis. *IEEE Trans. Image Processing*, 19 (4) (2010) 858-866
8. S. Gao, I. Tsang, L. Chia, Kernel sparse representation for image classification and face recognition. In *ECCV*, 6314 (2010) 1-14
9. M. Yang, L. Zhang. Gabor feature based sparse representation for face recognition with Gabor occlusion dictionary. In *ECCV*, 6316 (2010) 448-461.
10. M. Yang, L. Zhang, J. Yang, D. Zhang. Robust sparse coding for face recognition. *IEEE*, 42(7) (2011) 625-632
11. Naseem, I., Togneri, R., Bennamoun, M. Linear regression for face recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32 (11) (2010) 2106-2112.
12. X. Luan, B. Fang, Face recognition with contiguous occlusion using linear regression and level set method. *Neurocomputing*, 122 (2013) 386–397
13. X. Luan, B. Fang, Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion. *Pattern Recognition*, 47 (2) (2014) 495-508
14. L. Zhang, M. Yang, X. Feng, Sparse representation or collaborative representation: Which helps face recognition? *Proceedings of the 2011 International Conference on Computer Vision IEEE Computer Society*, (2011) 471-478
15. A. Georghiades, P. Belhumeur, D. Kriegman, From few to many: Illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23 (2001) 643-660
16. A. Martinez, R. Benavente, The AR face database, *CVC Tech. Report*. 24, 1998.