Chapter 5
Face Recognition Based on Maximum Sparse Coefficients of Object Region

Zineng Xu, Hongjun Li, Xiangyu Jin and Ching Y. Suen

Abstract Face recognition is an active topic in recognition systems, while face occlusion is one of the most challenging problems for recognition. Recently, robust sparse coding achieved the state-of-the-art performance, especially when dealing with occluded images. However, robust sparse coding is known that only guarantees the coefficient is global sparse when solving sparse coefficients. In this paper, we enable the elements in the object region to approximate global maximum by fitting the distribution of elements in the object region with successful recognition. The efficacy of the proposed approach is verified on publicly available databases. Furthermore, our method can achieve much better performance when the training samples are limited.

Keywords Face recognition · Maximum sparse coefficient · Occlusion

5.1 Introduction

Recently, face recognition (FR) approaches based on linear representations have led to state-of-the-art performance [1]. The most representative approach is sparse representation-based classification (SRC) [1]. However, SRC is not robust to contiguous occlusion such as sunglasses and scarf. Therefore, many related works [2–4] have been developed. In particular, the robust sparse coding (RSC) [2] method has achieved very good performance in FR with various occlusions. In order to construct a more robust model for sparse coding of face images, RSC finds

Z. Xu · H. Li (✉) · X. Jin
School of Electronic Information Engineering, Nantong University,
Nantong 226019, China
e-mail: lihongjun103@126.com

H. Li · C.Y. Suen
Centre for Pattern Recognition and Machine Intelligence,
Concordia University, Montreal, QC H3G 1M8, Canada

a maximum likelihood estimation solution of the coding coefficients and obtains high recognition accuracy; however, it is also very time-consuming, like SRC. Recently, Yang et al. [5] proposed a fast and robust FR method, which is faster than SRC and RSC.

Motivated by the recent success of RSC, we proposed a novel approach to improve the recognition rates by approximating the ideal sparse coefficients.

5.2 Related Works

Wright et al. [1] proposed a sparse representation-based classification scheme for FR. Let $D = [D_1, D_2, \ldots, D_s]$ be the set of training samples from all the s classes, where D_i is the subset of the training samples from class i. Denote y a testing sample. In SRC, y is sparsely coded over D via l_1-minimization:

$$\underset{\alpha}{\arg\min}\{||y - D\alpha||_2^2 + \lambda||\alpha||_1\} \tag{5.1}$$

where λ is a scalar constant, α is the coding vector of y over D.

Yang et al. [2] proposed RSC to solve this problem by LASSO:

$$\underset{\alpha}{\min} ||y - D\alpha||_2^2 \quad \text{s.t.} \quad ||\alpha||_1 \leq \sigma \tag{5.2}$$

where $\sigma > 0$ is a constant. They used a weighted regression function to measure the representation residual. The final equation can be described as:

$$\underset{\alpha}{\min}\{||W^{1/2}(y - D\alpha)||_2^2 + \lambda||\alpha||_1\} \tag{5.3}$$

where W is a diagonal matrix, each element is the weight assigned to each pixel of the test image y.

5.3 Proposed Method

5.3.1 Motivation

In sparse coding, the ideal sparse coefficients α can be described as follows:

$$\alpha = [0, \ldots, 0, \alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,k_i}, 0, \ldots, 0]^T \tag{5.4}$$

where k_i is the number of class i. We can know that RSC uses l_1-regularized least squares [6] to solve the sparse coefficients. Although this method can achieve good sparse representation, it only ensures the coefficients global sparse. In practice, we

hope the nonzero elements of sparse coefficients α can be concentrated in the object region as far as possible, and the remaining values are as small as possible. RSC is supervised, which means we know the labels of testing images. Thus we can easily identify the location of object region, and it is possible to control the distribution of sparse coefficients. We analyze how RSC method solves sparse coefficients. In RSC, when the testing image y belongs to class i, the solution of sparse coefficients can be performed as follows:

$$
\alpha^{y\in i} = [\overbrace{0_{1,1}, \alpha_{1,2}\ldots, \alpha_{1,k_1}}^{\text{region } 1}, \ldots, \overbrace{\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,k_i}}^{\text{object region } i}, \ldots
$$
$$
\underbrace{\alpha_{m,1}, 0_{m,2}, \ldots, \alpha_{m,k_m}}_{\text{region } m}, \ldots, \underbrace{0_{s,1}, \alpha_{s,2}, \ldots, \alpha_{s,k_s}}_{\text{region } s}]^T \tag{5.5}
$$

where $\alpha_{1,k_1}, \ldots, \alpha_{i,k_i}, \ldots, \alpha_{m,k_m}, \ldots, \alpha_{s,k_s}$ are nonzero elements of coefficients vector $\alpha^{y\in i}$, s is the number of classes. Denote $\text{SOR}_x = \alpha_{x,1} + \alpha_{x,2} + \cdots + \alpha_{x,k}$ as the sum of elements of coefficient vector in the region x, $x = 1, 2, \ldots, s$.

5.3.2 Maximum Sparse Coefficients of Object Region

We analyze the distribution of coefficients α in the object region with successful recognition. Figure 5.1 shows some distributions of elements in object regions. We can see that the value of nonzero elements in each object region is exponential growth. Considering that $\alpha < 1$, we choose logistic function to fit the distribution of elements in the object region. The basic formula of logistic function is

$$
L(r) = \frac{1}{1 + p^{-r}} \tag{5.6}
$$

where r is independent variable, p is a constant. Here, we modify the model in Eq. (5.6) to solve our problem.

The new formula can be described as follows:

$$
L(\alpha_{\text{or}}) = \begin{cases} \frac{c}{1 + p^{(a - b\alpha_{\text{or}})}}, & \max(\alpha_{\text{or}}) < z \\ \frac{c_1}{1 + p_1^{(a_1 - b_1\alpha_{\text{or}})}}, & z \le \max(\alpha_{\text{or}}) < z_1 \\ \frac{c_2}{1 + p_2^{(a_2 - b_2\alpha_{\text{or}})}}, & z_1 \le \max(\alpha_{\text{or}}) < 1 \end{cases} \tag{5.7}
$$

where $a, a_1, a_2, b, b_1, b_2, c, c_1, c_2, p, p_1, p_2, z, z_1$ and z_2 are constants, α_{or} is coefficient in the object region. From the result in Fig. 5.2, we can identify these parameters as follows: $a = 2.02$, $b = 20.01$, $c = 0.40$, $p = 8.01$; $a_1 = 4.00$, $b_1 = 20.00$, $c_1 = 0.50$, $p_1 = 5.99$; $a_2 = 3.31$, $b_2 = 10.52$, $c_2 = 0.74$, $p_2 = 4.05$.
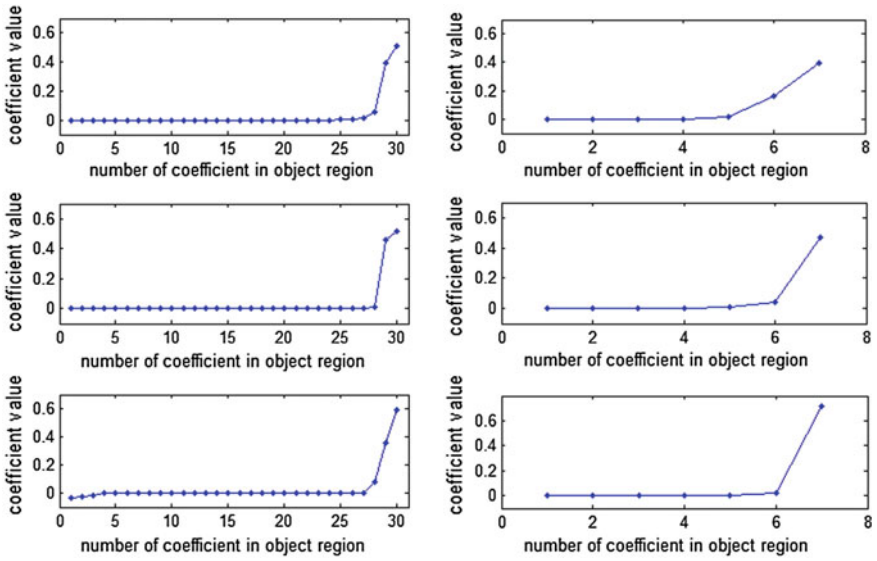
Fig. 5.1 Distribution of coefficients in object regions of different testing images
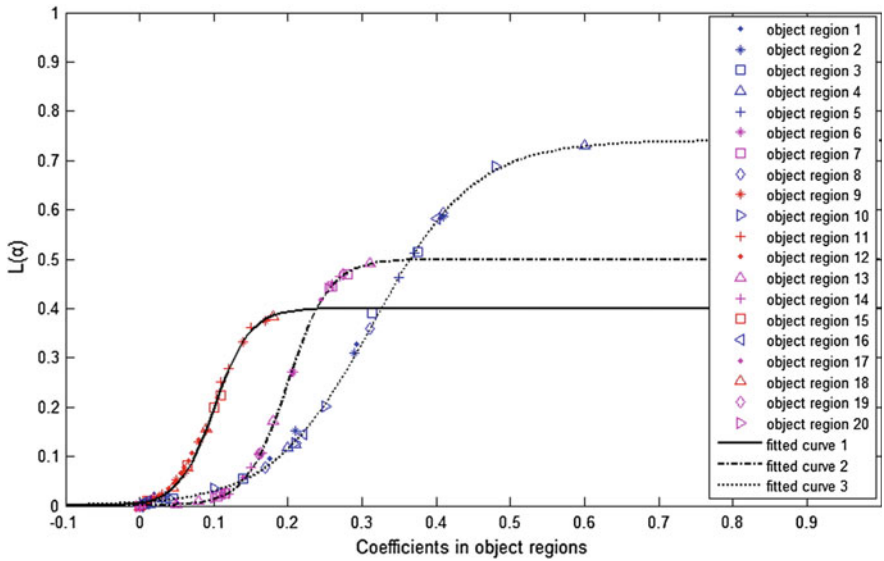

Fig. 5.2 Distribution of object regions with successful recognition and its fitted *curves*

For simplicity, we let $z = 0.2$, $z_1 = 0.35$ according to the range of independent variable. In order to globally maximize SOR_i, we need to find the maximum SOR_x. If $x \neq i$, we use the following formula to decrease its value:

$$M(\alpha_{mr}) = \frac{\text{SOR}_{max} - \Delta\text{SOR}_i}{\text{SOR}_{max}} \alpha_{mr} \tag{5.8}$$

where α_{mr} is the coefficient in the maximum region, SOR_{max} is the maximum region, ΔSOR_i is the increment in the object region. Using Eqs. (5.7) and (5.8), we can ensure SOR_i to be SOR_{max}. Finally, the sparse coefficients can be described as follows:

$$\alpha = [\alpha_{r1}, \alpha_{r2}, \ldots, L(\alpha_{ri}), \ldots, M(\alpha_{rs}), \ldots, \alpha_{rk-1}, \alpha_{rk}]^T \tag{5.9}$$

where α_{rj} is the coefficient in region j, $j = 1, 2, 3, \ldots, k$. $\alpha_{ri} = \alpha_{or}$, $\alpha_{rs} = \alpha_{mr}$.

5.3.3 Algorithm

We summarize the overall procedure of our algorithm as follows:

Step 1. Input: Normalized test sample y, dictionary $D = [D_1, D_2, \ldots, D_s]$
Step 2. Solve α: $\min_{\alpha}\{||W^{1/2}(y - D\alpha)||_2^2 + \lambda||\alpha||_1\}$

    a. $j = 1$, Initialize W_j and residual e_j
    b. Initialize α
    c. Identify SOR_{max} and α_{max}
    d. If $(\text{SOR}_i < \text{SOR}_{max})$ and $(\alpha_{i-max} < \alpha_{max})$

$$\alpha = [\alpha_{r1}, \alpha_{r2}, \ldots, L(\alpha_{ri}), \ldots, M(\alpha_{rs}), \ldots, \alpha_{rk-1}, \alpha_{rk}]^T \tag{5.10}$$

    Else go to next step. (α_{i-max} is the maximal element in the object region)
    e. Sparse coding by l_1-regularized least squares: α, then go back to c.
    f. Compute residual: $e_j(y) = ||y - D\alpha||_2^2$
    g. Update weights: W_j
    h. $j = j + 1$, go back to b. until the maximal number of iterations is reached.

Step 3. Output: the identity of y as $\text{Identity}(y) = \text{argmin}\{e_j\}$

5.4 Experimental Results

In this section, we carry out experiments on benchmark face databases (AR [7] and Extended Yale B [8]) to demonstrate the performance of our algorithm. In all experiments, we compare our method with some popular methods such as CRC_RLS [9], FDDL [10], and RSC [2].

5.4.1 Recognition Without Occlusion

Extended Yale B database: The Extended Yale B database consists of 2414 frontal-face images of 38 individuals. We used the cropped and normalized 45×40 face images. Figure 5.3 shows some facial images from the Extended Yale B database. Table 5.1 shows the recognition rates versus feature dimension by CRC_RLS, FDDL, RSC and our method.

Fig. 5.3 Samples of different illuminations from the Extended Yale B database

Table 5.1 Face recognition rates on the Extended Yale B database

Algorithm	Feature dimension		
	84	120	240
CRC_RLS (%)	96.6	97.2	98.4
FDDL (%)	93.1	96.4	98.0
RSC (%)	**98.8**	99.3	99.3
Our method (%)	98.7	**99.6**	**99.6**

Fig. 5.4 Samples from the AR database

AR database: The AR database consists of over 4,000 frontal images from 126 individuals. Figure 5.4 shows some facial images for example. The comparison of our method with its competing methods is given in Table 5.2. Table 5.3 shows that our method achieves much higher recognition rates than other three methods in different numbers of training samples. Our method performs better than RSC when there are few training samples.

Table 5.2 Face recognition rates on the AR database

Algorithm	Feature dimension			
	30	54	120	300
CRC_RLS (%)	64.6	80.6	90.1	93.7
FDDL (%)	53.3	77.4	86.4	91.4
RSC (%)	70.3	86.7	94.6	96.1
Our method (%)	**71.0**	**91.1**	**98.0**	**98.9**

Table 5.3 Recognition rates of different numbers of training samples on the AR database

Num	Algorithm			
	CRC_RLS (%)	FDDL (%)	RSC (%)	Our method (%)
1	59.4	–	65.0	**66.9**
2	62.0	65.4	68.0	**73.1**
3	64.3	68.0	72.6	**77.4**
4	76.1	71.0	77.6	**84.3**
5	86.6	81.0	86.9	**91.9**
6	93.9	91.3	95.4	**98.3**
7	93.7	91.4	96.1	**98.9**

5.4.2 Recognition with Real Disguise

In this section, we evaluate the robustness of our method to real disguise. We select a subset of the AR database consisting of both neutral and corrupted images in this experiment. The following three scenarios are considered for performance evaluation: Sunglasses, Scarf, and Sunglasses + Scarf. Figure 5.5 shows one person's images with different disguises from training and testing images. Table 5.4 shows that our recognition rates are higher than other methods. The experimental results also demonstrate the good robustness of our method to face occlusion.
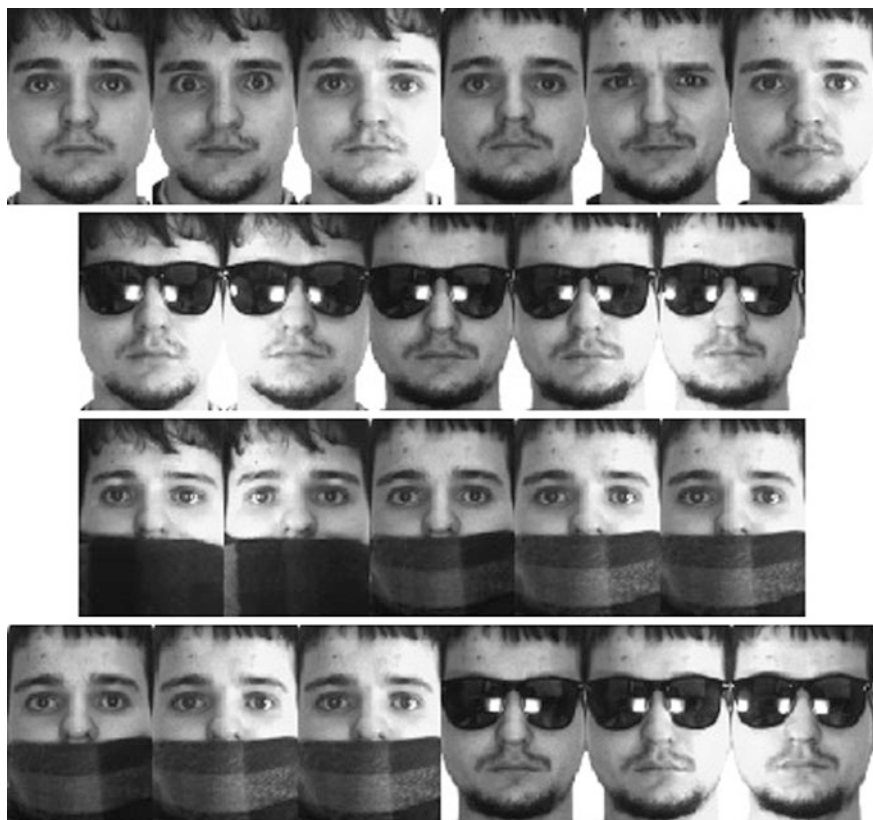

Fig. 5.5 The first row is training images, the rest are testing images

Table 5.4 Face recognition rates on the AR database with disguise occlusion

Algorithm	Sunglasses	Scarf	Sunglasses + scarf
CRC_RLS (%)	47.8	46.8	51.3
RSC (%)	91.6	88.8	92.2
Our method (%)	**91.8**	**90.0**	**92.7**

5.5 Conclusion

In this paper, motivated by the recent success of RSC, we present a novel method to recognize face image with illumination variations and occlusion. We add constraints and use logistic function to optimize the sparse coefficients. Our extensive experimental results on benchmark face databases show that the proposed method is effective and robust. Compared with RSC, our method achieves higher recognition rates, even when the training samples are limited.

References

1. Wright J, Yang AY, Ganesh A et al (2009) Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell 31(2):210–227
2. Yang M, Zhang L, Yang J et al (2013) Regularized robust coding for face recognition. IEEE Trans Image Process 22(5):1753–1766
3. Ou WH, You XG, Tao DC et al (2014) Robust face recognition via occlusion dictionary learning. Pattern Recogn 47(4):1559–1572
4. Li HJ, Nobile N, Suen CY (2014) Face recognition based on discriminative dictionary with multilevel feature fusion. Lect Notes Artif Intell 8774:252–263
5. Yang M, Feng ZZ, Shiu SCK et al (2014) Fast and robust face recognition via coding residual map learning based adaptive masking. Pattern Recogn 47(2):535–543
6. Goldstein T, Osher S (2009) The split Bregman method for L1-regularized problems. SIAM J Imaging Sci 2(2):323–343
7. Martinez A, Benavente R (1998) The AR face database. Technical report 24, CVC
8. Georghiades A, Belhumeur P, Kriegman D (2001) From few to many: illumination cone models for face recognition under variable lighting and pose. IEEE Trans Pattern Anal Machine Intell 23(6):643–660
9. Zhang L, Yang M, Feng X (2011) Sparse representation or collaborative representation: which helps face recognition? In: Proceedings of ICCV, pp 471–478
10. Yang M, Zhang L, Feng X et al (2014) Sparse representation based fisher discrimination dictionary learning for image classification. Int J Comput Vision 109(3):209–232