Postprint of Sparse Representation Face Recognition Method Based on L2-Norm Sample Reconstruction Constraints

Authors: Mi Jianxun, Zhikai Lin

Date: 2019-01-28T00:00:00+00:00

Abstract

Sparse representation classification methods achieve favorable classification performance with large-scale training sample spaces, but suffer from high computational costs. To address this issue in sparse representation methods, this work proposes constructing an L_2-norm constraint on the reconstructed samples, which enhances competition among category-specific components within the reconstructed samples to achieve a group sparsity effect, ultimately improving classification accuracy. Since the method can directly obtain a closed-form solution, the computational cost of solving is significantly reduced, while the sparsity of the resulting coefficients is comparable to that of conventional methods. Comparative experiments with similar methods on publicly available face and object image datasets demonstrate that the method achieves excellent image recognition performance under complex conditions.

Full Text

Sparse Representation Classification via L2-Norm Based Reconstruction Sample Constraint for Face Recognition

Mi Jianxuna,b†, Lin Zhikaia,b

aSchool of Computer Science & Technology, bChongqing Key Laboratory of Computational Intelligence,
Chongqing University of Posts & Telecommunications, Chongqing 400065, China

Abstract

Sparse representation classification (SRC) demonstrates strong performance when the training sample space is sufficiently large, but suffers from high computational cost. To address this limitation, we propose constructing an L2-norm constraint on reconstruction samples to intensify competition among different class components during reconstruction, thereby achieving group sparsity effects and improving classification accuracy. Since our method admits a closed-form solution, computational costs are substantially reduced while producing coefficient sparsity comparable to traditional approaches. Comparative experiments on public face and object image datasets demonstrate that the proposed method achieves excellent recognition performance under complex conditions.

Key words: sparse representation; face recognition; collaborative representation; reconstruction sample

0 Introduction

Face recognition represents a prominent research direction in computer vision and pattern recognition. As a crucial biometric identification method, it plays vital roles in numerous real-world applications including airport security systems, mobile face unlocking, and facial payment authentication. Existing face recognition techniques primarily fall into three categories: geometry-based methods, template-based methods, and mathematical statistical model-based methods [1]. As deployment deepens, algorithms must handle increasingly complex scenarios where variations in illumination, expression, pose, and occlusion significantly impact performance, creating urgent demand for more robust and efficient solutions.

Since Wright et al. [2] introduced Sparse Representation-based Classification (SRC) for face recognition in 2009, the approach has attracted widespread attention. SRC performs sparse linear regression of test faces over a dictionary composed of training samples to obtain reconstructed faces, then compares similarity between reconstruction components across different classes and the test sample to determine identity. SRC employs L1-norm regularization on representation coefficients to encourage sparsity. Wright argued that with sufficient test samples and adequately sparse coefficients, SRC can perfectly reconstruct test samples. The L1-regularized minimization problem is solved through iterative optimization, with relevant algorithms surveyed in [3]. Experiments demonstrate that SRC achieves excellent classification results in face recognition when training samples are abundant. Subsequent research [4–9] has further validated its effectiveness in image processing and face recognition.

However, Zhang et al. contend that SRC' s success primarily stems from collaborative representation mechanisms, leading them to propose CRC_RLS (Collaborative Representation Classification with Regularized Least Square) [10]. CRC_RLS replaces L1-norm with L2-norm regularization to achieve similar sparse constraints. Due to its closed-form solution, CRC_RLS offers significantly faster computation than SRC, though with lower coefficient sparsity. To achieve comparable sparsity while maintaining computational efficiency, Xu et al. developed a series of supervised sparse representation methods [11-15]. Among them, TPTSR [11] (Two-Phase Test Sample Representation) fits samples in two stages: first eliminating training samples with small contributions, then refitting with remaining samples to achieve SRC-like sparsity. Since both fitting stages apply L2-norm constraints, TPTSR incurs minimal computational cost. Both CRC_RLS and TPTSR aim to achieve SRC-like sparsity with faster solving speeds, a research direction that also includes weighted approaches.

AWCRC (Adaptive and Weighted Collaborative Representation Classification) [16] represents a typical weighted collaborative representation method that applies weighted constraints to coefficients based on distances between test and training samples to obtain sparse coefficients. ProCRC [17] achieves fast sparse coefficients by constraining the norm of reconstruction samples from multiple classes, while DSRC [18] similarly constructs constraints in the sample reconstruction space. Their strong performance demonstrates that appropriate constraints on reconstruction samples can effectively substitute for L1-norm coefficient regularization.

To develop faster and more robust face recognition algorithms, this paper proposes a sparse representation classification method based on L2-norm constraints on reconstruction samples. By constraining the joint modulus length of every two class components in the reconstructed sample, our approach simultaneously controls both the magnitude of different class reconstruction components and their angular relationships, thereby increasing the probability that test samples are represented by single-class samples and improving overall classification performance. Comparative experiments on mainstream face and object image databases demonstrate the proposed method' s strong performance in image recognition.

1 Sparse Representation Classification Methods

Sparse representation classification performs linear regression fitting of test samples over a dictionary composed of training samples, then determines test sample identity based on reconstruction effectiveness across different class components. Let $X = [X_1, X_2, \dots, X_c]$ denote the training sample dictionary containing c classes, where X_i represents samples from class i. Let y denote the test sample. Linear fitting yields $y = Xa + \epsilon$, where $a = [a_1, a_2, \dots, a_c]$ represents the overall coefficient vector and $y_i = X_i a_i$ denotes the reconstruction component for class i.

Unlike direct linear regression (least squares), SRC imposes L1-norm regularization on coefficients, with the objective function:

$$\min_a \|y - Xa\|_2^2 + \lambda \|a\|_1 \tag{1}$$

where the first term is the fidelity term and the second is the regularization term. Traditional least squares includes only the fidelity term, while Equation (1) resembles Lasso [19] in regression, serving both regression and variable selection purposes.

Although L1-norm implicitly selects training samples during regression, iterative solving incurs high computational cost. CRC_RLS replaces the regularization norm with L2-norm:

$$\min_{a} \|y - Xa\|_2^2 + \lambda \|a\|_2^2 \tag{2}$$

CRC_RLS resembles ridge regression, introducing relaxed sparse constraints to maximize fitting capability when dictionary elements are insufficient.

However, when dictionary elements are adequate, L2-norm constraints yield less sparse coefficients, hindering classification. Recent work addresses this by constraining reconstruction samples to induce sparsity [17,18,20,21]. DSRC adds a regularization term constraining reconstruction sample modulus length:

$$\min_{a} \|y - Xa\|_2^2 + \lambda \|a\|_2^2 + \gamma \sum_{i=1}^{c} \|X_i a_i\|_2^2 \tag{3}$$

Due to L1-norm' s properties, most training samples weakly correlated with test samples lose representation opportunity. With limited training samples, test samples may be misclassified due to loss of representation from weakly correlated but correct-class samples. Conversely, L2-norm encourages more training samples to represent test samples, ideally positioning reconstructed samples at training sample centers where each sample contributes. Both L1-norm and L2-norm share a common limitation: classification performance heavily depends on training sample distribution. When test samples seek optimal representation by minimizing regularization terms, all samples combine with equal probability since training samples carry no label information. When inter-class differences are small, test samples likely receive representation from multiple classes, increasing misclassification risk. Incorporating label information into training samples would enhance representation consistency and improve classification.

In regression, Group Lasso [22] applies different constraints to coefficients—L2-norm within classes and L1-norm between classes—to increase the likelihood of combining same-class samples. This paper enhances competition among different class reconstruction samples to increase same-class combination probability, achieving Group Lasso-like sparsity. Our method adds a new regularization term to CRC_RLS, constraining the sum of pairwise modulus lengths among different class components to eliminate weakly expressive classes while strengthening dominant classes in the overall reconstruction, thereby improving recognition capability.

2.1 Algorithm Description

Let $X = [X_1, X_2, \dots, X_c]$ denote a dictionary containing c class training samples, y the test sample, and $a = [a_1, a_2, \dots, a_c]$ the overall coefficient vector. We propose RCSRC (Reconstruction Constraint Sparse Representation Classification), which adds a constraint term on different class components in reconstruction samples to CRC_RLS. The objective function is:

$$\min_a \frac{1}{2}\|y - Xa\|_2^2 + \frac{\lambda}{2}\|a\|_2^2 + \frac{\gamma}{2}\sum_{i=1}^{c}\sum_{j=1,j\neq i}^{c}\|X_i a_i + X_j a_j\|_2^2 \tag{5}$$

where $y_i = X_i a_i$ represents class i' s reconstruction component, and λ and γ balance the fidelity and regularization terms. Direct differentiation of Equation (5) and setting to zero yields the closed-form solution:

$$a = [(1 + \gamma(c-1))X^T X + \lambda I + \gamma\Phi]^{-1} X^T y \tag{6}$$

where I is the identity matrix and Φ is defined as:

$$\Phi = \begin{bmatrix} 0 & X_1^T X_2 & \cdots & X_1^T X_c \\ X_2^T X_1 & 0 & \cdots & X_2^T X_c \\ \vdots & \vdots & \ddots & \vdots \\ X_c^T X_1 & X_c^T X_2 & \cdots & 0 \end{bmatrix}$$

Classification prediction is made by computing distances between reconstruction samples from different class components and the test sample:

$$\text{label}(y) = \arg\min_i \|y - X_i a_i\|_2 \tag{7}$$

The complete algorithm proceeds as follows: First, normalize column vectors in training set X. Then compute coefficient a using Equation (6). Finally, determine test sample label using Equation (7). Input parameters include training set X, test sample y, regularization parameters λ and γ, and total number of classes c.

Additionally, ProCRC constrains the modulus length of reconstruction samples formed by multiple classes on top of CRC_RLS:

$$\min_a \|y - Xa\|_2^2 + \lambda\|a\|_2^2 + \gamma\sum_{i=1}^{c}\|X_i a_i\|_2^2 \tag{8}$$

Experiments demonstrate that constraints on reconstruction sample modulus length effectively enhance coefficient sparsity and improve classification accuracy.

2.2 Algorithm Analysis

This research extends traditional sparse representation to develop computationally efficient methods with adequate coefficient sparsity. Both theoretical and experimental evidence [2] confirms sparse representation' s advantages in face recognition, though sparse constraints only benefit classification when dictionary elements are sufficiently numerous.

RCSRC' s objective function (5) comprises three components: a fidelity term, a regularization term, and a reconstruction sample modulus constraint. The fidelity term ensures reconstruction samples vary within reasonable bounds of test samples. The L2-norm regularization on coefficients maintains coefficients within reasonable ranges. The third term ensures proper competition among different class reconstruction components, promoting same-class sample combination while suppressing weakly expressive classes. Expanding the third term reveals:

$$\sum_{i=1}^{c}\sum_{j=1,j\neq i}^{c}\|X_i a_i + X_j a_j\|_2^2 = \sum_{i=1}^{c}(c-1)\|X_i a_i\|_2^2 + \sum_{i=1}^{c}\sum_{j=1,j\neq i}^{c} y_i^T y_j$$

where $\|X_i a_i\|_2^2$ represents the modulus length of reconstruction component for class i, and $y_i^T y_j$ indicates correlation between class i and j components. The modulus constraint suppresses weakly correlated components to small values, while correlation constraints minimize inter-class correlations. This benefits scenarios with numerous training sample classes by enabling final reconstruction samples to be dominated by only a few class components, eliminating weakly correlated categories' influence on classification. Since SRC and CRC_RLS lack label information, they blindly select samples based solely on correlation. As training sample classes increase, distribution spaces from different classes increasingly overlap. When test samples reside within overlapping regions of multiple class distribution spaces, selected training samples inevitably span multiple spaces. RCSRC, however, leverages known class information to holistically eliminate weakly correlated class samples, constructing reconstruction samples spanning fewer spaces.

To demonstrate RCSRC' s coefficient sparsity, we conducted experiments on the Extended YaleB [23] face database using 31 classes with 20 faces per class. Figure 1 [Figure 1: see original paper]-3 show coefficient distributions for a test face image from class 1 computed by RCSRC, CRC_RLS, and SRC respectively. To quantitatively compare sparsity, we computed the Sparsity Concentration Index (SCI) as defined in [2]:

$$\mathrm{SCI}(a) = \frac{c \cdot \max_i \|\delta_i(a)\|_1 / \|a\|_1 - 1}{c-1} \in [0,1]$$

where $\delta_i(a)$ extracts coefficients belonging to class i. Larger SCI values indicate higher concentration—SCI = 1 means coefficients concentrate in a single class while others are zero; SCI = 0 means absolute coefficient sums are equal across all classes. Computed SCI values are: RCSRC = 0.2045, CRC_RLS = 0.1393, and SRC = 0.9308. Computation times are: RCSRC = 0.040482 s, CRC_RLS = 0.014769 s, SRC = 0.294314 s (experimental environment: Intel i9-7980XE CPU, 16 GB RAM, MATLAB 2018a). RCSRC achieves coefficient sparsity between SRC and CRC_RLS while offering significant speed advantages over SRC.

3 Experimental Analysis

We evaluated RCSRC' s performance across mainstream face and object image databases: Georgia Tech (GT) [24], Extended YaleB [23], LFW [25], and COIL-20 [26]. Compared algorithms include SRC [2], CRC [10], ProCRC [17], and DSRC [18], with SRC solved via L1LS. All experiments ran on a PC with MATLAB 2018a, Intel i9-7980XE CPU, and 16 GB RAM.

a) Georgia Tech Face Database. GT contains 750 face images from 50 individuals (15 images per person), featuring expression, illumination, and scale variations. All images were grayscale-converted, manually cropped, and resized to 25×20 pixels, then stretched into column vectors to form the dictionary. Figure 4 [Figure 4: see original paper] shows cropped sample images. We randomly selected 8 and 10 images per class for training, using the remainder for testing, repeating each experiment 10 times. Table 1 presents average classification accuracy and variance.

b) Extended YaleB Face Database. Extended YaleB comprises 2,414 face images from 31 subjects captured under controlled laboratory lighting conditions. We used 64 images per person, cropped and resized to 20×20 pixels (Figure 5 [Figure 5: see original paper]). Randomly selecting 10, 15, 20, 25, and 30 images per class for training (remainder for testing), we performed 10 trials per configuration. Results appear in Table 2 .

RCSRC outperforms other methods by 1%–2%. Accuracy increases with training sample size up to 20 samples per class, then stabilizes or slightly declines. This occurs because the linear space spanned by training samples becomes saturated (dictionary columns exceed rows), enabling perfect linear representation of test samples. Additionally, similar classes experience expanded representation spaces with more samples, increasing misclassification risk.

c) LFW Face Database. LFW contains over 13,000 web-collected face images with pose, illumination, expression, and occlusion variations. Figure 6 [Figure 6: see original paper] shows aligned and cropped examples, all resized to 50×40 pixels. We employed two protocols: (1) select classes with >16 samples, using 15 for training; (2) select classes with >11 samples, using 10 for training. Table 3 shows results. LFW' s uncontrolled collection conditions make experiments challenging due to large intra-class variation and overlapping inter-class distribution spaces. RCSRC' s use of label information clarifies class boundaries, yielding superior results.

d) COIL-20 Object Database. COIL-20 contains 20 object classes with 72 multi-view images each, all resized to 15×15 pixels (Figure 7 [Figure 7: see original paper]). We randomly selected 10, 15, and 20 images per class for training, conducting 10 trials per configuration. Table 4 shows results. Object recognition exhibits larger inter-individual differences than face recognition, resulting in smaller performance gaps across methods, with RCSRC leading by 1%-3%.

4 Conclusion

Traditional SRC achieves excellent classification under sparse constraints but suffers high computational cost, while relaxed CRC_RLS reduces cost but yields insufficient sparsity. Addressing these limitations, we propose RCSRC, an L2-norm reconstruction sample constraint method that rapidly obtains sparse coefficients by constraining reconstruction sample modulus lengths and inter-class component correlations. Comparative experiments on mainstream face and object databases demonstrate the proposed method' s strong performance in complex face recognition scenarios.

References

[1] Wu Jinyi, Zhou Delong. Survey of face recognition [J]. Application Research of Computers, 2009, 26 (9): 3205-3209.

[2] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31 (2): 210-227.

[3] Yang A Y, Zhou Zihan, Balasubramanian A G, et al. Fast l1-Minimization Algorithms for Robust Face Recognition [J]. IEEE Trans on Image Processing, 2013, 22 (8): 3234-3246.

[4] Zhang Yao, Xu Bin, Zhou Shangbo, et al. Image super-resolution with adaptive regularization sparse representation [J]. Application Research of Computers, 2013, 30 (3): 938-941.

[5] Liu Tao, Mi Jianxun, Liu Ying, et al. Robust face recognition via sparse boosting representation [J]. Neurocomputing, 2016, 214: 944-957.

[6] Kang Lipan, Chen Fangfu, Fan Zizhu. New extended sparse representation for face recognition [J]. Application Research of Computers, 2016, 33 (3): 937-939.

[7] Tang Zhenyu, Meng Fangrong, Wang Zhixiao. Fast face recognition with regularized least square via sparse representation [J]. Application Research of Computers, 2016, 33 (9): 2831-2836.

[8] Mi Jianxun, Fu Qiankun, Li Weisheng. Adaptive class preserving representation for image classification [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2017: 7427-7435.

[9] Shen Tao, Li Keqing, Xia Yu. Single sample recognition based on robust sparse representation [J]. Application Research of Computers, 2018, 35(11): 3491-3496.

[10] Zhang Lei, Yang Meng, Feng Xiangchu. Sparse representation or collaborative representation: Which helps face recognition? [C]//Proc of IEEE International Conference on Computer vision. 2011: 471-478.

[11] Xu Yong, Zhang David, Yang Jian, et al. A two-phase test sample sparse representation method for use with face recognition [J]. IEEE Trans on Circuits and Systems for Video Technology, 2011, 21 (9): 1255-1262.

[12] Xu Yong, Zuo Wangmeng, Fan Zizhu. Supervised sparse representation method with a heuristic strategy and face recognition experiments [J]. Neurocomputing, 2012, 79: 125-131.

[13] Mi Jianxun. Face image recognition via collaborative representation on selected training samples [J]. Optik, 2013, 124 (18): 3310-3313.

[14] Xu Yong, Zhu Qi, Fan Zizhu, et al. Using the idea of the sparse representation to perform coarse-to-fine recognition [J]. Information Sciences, 2013, 238(7): 138-148.

[15] Mi Jianxun, Sun Yueru, Lu Jia. Robust face recognition based on supervised sparse representation [C] // Proc of International Conference on Intelligent Computing. Cham: Springer, 2018: 253-259.

[16] Timofte R, Van Gool L. Adaptive and weighted collaborative representations for image classification [J]. Pattern Recognition Letters, 2014, 43(7): 127-135.

[17] Cai Sijia, Zhang Lei, Zuo Wangmeng, et al. A probabilistic collaborative representation based approach for pattern classification [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2016: 2950-2959.

[18] Xu Yong, Zhong Zuofeng, Yang Jian, et al. A new discriminative sparse representation method for robust face recognition via L2 regularization [J]. IEEE Trans on Neural Networks and Learning Systems, 2017, 28(10): 2233-2242.

[19] Tibshirani R. Regression shrinkage and selection via the lasso [J]. Journal of the Royal Statistical Society. Series B (Methodological), 1996: 267-288.

[20] Chi Hongmei, Xia Haifeng, Zhang Lifang, et al. Competitive and collaborative representation for classification [J]. Pattern Recognition Letters, 2018. https://doi.org/10.1016/j.patrec.2018.06.019.

[21] Yuan Haoliang, Li Xuecong, Xu Fangyuan, et al. A collaborative-competitive representation based classifier model [J]. Neurocomputing, 2018, 275(1): 627-635.

[22] Yuan Ming, Lin Yi. Model selection and estimation in regression with grouped variables [J]. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2006, 68 (1): 49-67.

[23] Georghiades A S, Belhumeur P N, Kriegman D J. From few to many: Illumination cone models for face recognition under variable lighting and pose [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2001, 23 (6): 643-660.

[24] Goel N, Bebis G, Nefian A. Face recognition experiments with random projection [C]//Biometric Technology for Human Identification II. 2005.

[25] Huang G B, Mattar M, Berg T, et al. Labeled faces in the wild: a database for studying face recognition in unconstrained environments [C]//Proc of Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition. 2008.

[26] Nene S A, Nayar S K, Murase H. Columbia object image library (coil-20) [R/OL]. 1996. https://doc.uments.com/g-columbia-object-image-library-coil-20.pdf.

Note: Figure translations are in progress. See original paper for figures.