

Weighted Sparse Image Classification Based on Low Rank Representation

Qidi Wu¹, Yibing Li¹, Yun Lin^{1,*} and Ruolin Zhou²

Abstract: The conventional sparse representation-based image classification usually codes the samples independently, which will ignore the correlation information existed in the data. Hence, if we can explore the correlation information hidden in the data, the classification result will be improved significantly. To this end, in this paper, a novel weighted supervised sparse coding method is proposed to address the image classification problem. The proposed method firstly explores the structural information sufficiently hidden in the data based on the low rank representation. And then, it introduced the extracted structural information to a novel weighted sparse representation model to code the samples in a supervised way. Experimental results show that the proposed method is superiority to many conventional image classification methods.

Keywords: Image classification, sparse representation, low-rank representation, numerical optimization.

1 Introduction

Image classification is a fundamental issue in computer vision, which aims to classify an image into an accurate category. Many literatures about image classification are emerged in the past two decades [Liu, Fu and He (2017)], and several classification frameworks are formed to address the classification problem of vision task. During the mentioned frameworks, sparse-based classification attracted many attentions and are widely studied in recent years. Sparse signal representation has proven to be an extremely powerful tool for acquiring, representing, and compressing high-dimensional signals [Wright, Ma, Mairal et al. (2010)]. In the past few years, sparse representation has been applied to many vision tasks, including face recognition [Yang, Zhang, Yang et al. (2011); Liu, Tran and Sang (2016)], image super-resolution [Yang, Chu and Wang (2010); Yang, Wright, Huang et al. (2010)], denoising and inpainting [Fadili, Starck and Murtagh (2013); Guleryuz (2006)], and image classification [Kulkarni and Li (2011); Thiagarajan and Spanias (2012)]. Due to its wide applicability, people conducted detailed research on image classification based on sparse representation and achieved remarkable results [Tang, Huang and Xue (2016); Gao, Tsang and Chia (2010); Liu, Fu and He (2017)].

¹ College of Information and Communication Engineering, Harbin Engineering University, Harbin, 150001, China.

² Department of Electrical and Computer Engineering, Western New England University, USA.

* Corresponding Author: Yun Lin. Email: linyun@hrbeu.edu.cn.

According to the picture type, image classification is mainly applied in certain specific fields such as face recognition and hyperspectral image classification and remote sensing image classification. In face recognition, based on a sparse representation computed by l_1 -minimization, Wright et al. [Wright, Ma, Mairal et al. (2010)] proposed a general classification algorithm for (image-based) object recognition [Wright, Ganesh, Zhou et al. (2009)]. This new framework provided new insights into two crucial issues in face recognition: feature extraction and robustness to occlusion. The theory of sparse representation helped predict how much occlusion the recognition algorithm could handle and how to choose the training images to maximize robustness to occlusion. Recent research has shown the speed advantage of extreme learning machine (ELM) and the accuracy advantage of sparse representation classification (SRC) in the area of image classification. In order to unify such mutual complementarity and thus further enhance the classification performance, Cao et al. [Cao, Zhang, Luo et al. (2016)] proposed an efficient hybrid classifier to exploit the advantages of ELM and SRC in Cao et al. [Cao, Zhang, Luo et al. (2016)]. While the importance of sparsity is much emphasized in SRC and many related works, the use of collaborative representation (CR) in SRC is ignored by most literature. This paper [Zhang, Yang and Feng (2011)] devoted to analyze the working mechanism of SRC, and indicated that it was the CR but not the l_1 -norm sparsity that made SRC powerful for face classification. The authors proposed a very simple yet much more efficient face classification scheme, namely CR based classification with regularized least square (CRC_RLS), which had very competitive classification results, many classic and contemporary face recognition algorithms work well on public data sets, but degrade sharply when they are used in a real recognition system. This is mostly due to the difficulty of simultaneously handling variations in illumination, image misalignment, and occlusion in the test image. Considering this problem, Wagner et al. [Wagner, Wright, Ganesh et al. (2012)] proposed a conceptually simple face recognition system that achieved a high degree of robustness and stability to illumination variation, image misalignment, and partial occlusion. Besides, Zhang et al. [Zhang, Zhang, Huang et al. (2013)] proposed a pose-robust face recognition method to handle the challenging task of face recognition in the presence of large pose difference between gallery and probe faces. The proposed method exploited the sparse property of the representation coefficients of a face image over its corresponding view-dictionary. In hyperspectral image classification, Tang et al. [Tang, Chen, Liu et al. (2015)] present a hyperspectral image classification method based on sparse representation and superpixel segmentation. By refining the spectral classification results with the spatial constraints, the accuracy of classification is improved by a substantial margin in Tang et al. [Tang, Chen, Liu et al. (2015)]. Noting that the structural information can earn some extra discrimination for the classification framework [Liu, Fu and He (2017)], some methods considering the structural and contextual information are presented. A new sparsity-based algorithm for the classification of hyperspectral imagery was proposed in Chen et al. [Chen, Nasrabadi and Tran (2011)], which relied on the observation that a hyperspectral pixel could be sparsely represented by a linear combination of a few training samples from a structured dictionary. They mainly used Laplacian constraint and joint sparsity model to incorporate the contextual information into the sparse recovery optimization problem in order to improve the classification performance. To overcome this problem

that the sparse representation-based classification (SRC) methods ignore the sparse representation residuals (i.e. outliers), Li et al. [Li, Ma, Mei et al. (2017)] proposed a robust SRC (RSRC) method which could handle outliers. They extended the RSRC to the joint robust sparsity model named JRSRC, where pixels in a small neighborhood around the test pixel were simultaneously represented by linear combinations of a few training samples and outliers. A superpixel tensor sparse coding (STSC) based hyperspectral image classification (HIC) method is proposed, by exploring the high-order structure of hyperspectral image and utilizing information along all dimensions to better understand data in Feng et al. [Feng, Wang, Yang et al. (2017)]. Article [Sun, Qu, Nasrabadi et al. (2014)] proposes a new structured prior called the low-rank (LR) group prior, which can be considered as a modification of the LR prior. Structured Priors for Sparse-Representation-Based Hyperspectral Image Classification. The paper [Zhang, Song, Gao et al. (2016)] presents a new spectral-spatial feature learning method for hyperspectral image classification, which integrates spectral and Zeng et al. [Zeng, Li, Liang et al. (2010)] spatial information into group sparse coding (GSC) via clusters and propose a novel kernelized classification framework based on sparse representation considering the image classification problem based on the similarities between images [Zhang, Zhang, Liu et al. (2015)]. Article [Zeng, Li, Liang et al. (2010)] proposes a fast joint sparse representation classification method with multi-feature combination learning for hyperspectral imagery and incorporate contextual neighborhood information of the image into each kind of feature to further improve the classification performance. The paper [Wang, Xie, Li et al. (2015)] investigates STDA (Sparse Tensor Discriminant Analysis) for feature extraction and it is the first time that STDA is applied for HIS (Hyperspectral imagery) and attempts to adopt STDA to preserve useful structural information in the original data and obtain multiple interrelated sparse discriminant subspaces. In remote sensing image classification, The paper [Shivakumar, Natarajan and Murthy (2015)] presents a novel multi-kernel based sparse representation for the classification of Remotely sensed images. The sparse representation based feature extraction are in a run which is a signal dependent feature extraction and thus more accurate. The paper [Song, Li, Dalla et al. (2014)] proposes to exploit sparse representations of morphological attribute profiles for remotely sensed image classification. By using the sparse representation classification framework to exploit this characteristic of the EMAPs.

With more and more sources of image data on the Internet, images are becoming more and more cluttered, making more and more images lack complete category information. Based on this, the researchers proposed semi-supervised and unsupervised image classification methods. To address the problem of face recognition when there is only few, or even only a single, labeled examples of the face that we wish to recognize. The main idea was that: they used the variation dictionary to characterize the linear nuisance variables via the sparsity framework and prototype face images were estimated as a gallery dictionary via a Gaussian mixture model, with mixed labeled and unlabeled samples in a semi-supervised manner, to deal with the non-linear nuisance variations between labeled and unlabeled samples. Unsupervised learning methods for building feature extractors have a long and successful history in pattern recognition and computer vision. Document [Huang, Boureau and Le (2007)] proposes an unsupervised method to learn the sparse feature detector hierarchy, which is invariant for small shifts and

distortions. Article [Zeiler, Krishnan, Taylor et al. (2010)] present a learning framework where features that capture these mid-level cues spontaneously emerge from image data and the approach is based on the convolutional decomposition of images under a sparsity constraint and is totally unsupervised. The paper [Poultney, Chopra and Cun (2007)] describes a novel unsupervised method for learning sparse, overcomplete features. The model uses a linear encoder, and a linear decoder preceded by a sparsifying non-linearity that turns a code vector into a quasi-binary sparse code vector. The major contribution of paper [Hassairi, Ejbali and Zaied (2016)] is to show how to extract features and train an image classification system on large-scale datasets.

In this paper, we proposed novel sparse representation-based image classification with weighted supervision sparse coding. It firstly constructed the supervision coefficient with Low rank representation, which would help to mine the structural information hidden in the data. And then, the supervision coefficient is incorporated into a weighted sparse representation framework to preserve the coding structure of similar samples. It is verified that the proposed method shows some advantage on the public test dataset.

The remaining of the paper is arranged as follows. In section II, several similar previous works are analyzed. In section III, the proposed method is described in detail. In section IV, the experimental results are shown to demonstrate the superiority of the proposed method. Section V concludes the paper.

2 Related work

Although the traditional spatial pyramid matching (SPM) approach works well for image classification, people empirically found that, to achieve good performance, traditional SPM has to use classifiers with nonlinear Mercer kernels. Accordingly, the nonlinear classifier has to afford additional computational complexity, bearing $O(n^3)$ in training and $O(n)$ for testing in SVM, where n is the number of support vectors. This implies a poor scalability of the SPM approach for real applications.

To improve the scalability, researchers aim at obtaining nonlinear feature representations that work better with linear classifiers

Let X be a set of D -dimensional local descriptors extracted from an image, *i.e.* $X = [x_1, x_2, \dots, x_N] \in R^{D \times N}$. Given a codebook with M entries, $B = [b_1, b_2, \dots, b_M] \in R^{D \times M}$, different coding schemes convert each descriptor into adimensional code to generate the final image representation. This section reviews three existing coding schemes.

2.1 Coding descriptors in VQ

Traditional SPM uses VQ coding which solves the following constrained least square fitting problem:

$$\arg \min_C \sum_{i=1}^N \|x_i - Bc_i\|^2 \quad (1)$$

$$s.t. \|c_i\|_{\ell^1} = 1, c_i \geq 0, \forall i$$

where $C = [c_1, c_2, \dots, c_N]$ is the set of codes for X . The cardinality constraint $\|c_i\|_{\ell^0} = 1$

means that there will be only one non-zero element in each code c_i , corresponding to the quantization id of x_i . The non-negative, ℓ^1 constraint $\|c_i\|_{\ell^1} = 1, c_i \geq 0$ means that the coding weight for x is 1. In practice, the single non-zero element is found by searching the nearest neighbor.

2.2 Coding descriptors in ScSPM

To ameliorate the quantization loss of VQ, the restrictive cardinality constraint $\|c_i\|_{\ell^0} = 1$ in Eq. (1) can be relaxed by using a sparsity regularization term. In ScSPM [Yang, Yu, Gong et al. (2009)], such a sparsity regularization term is selected to be the ℓ^1 norm of c_i , and coding each local descriptor x_i thus becomes a standard sparse coding (SC) [Lee, Battle, Raina et al. (2007)] problem:

$$\arg \min_c \sum_{i=1}^N \|x_i - Bc_i\|^2 + \lambda \|c_i\|_{\ell^1} \quad (2)$$

The sparsity regularization term plays several important roles: First, the codebook B is usually over-complete, *i.e.* $M > D$, and hence ℓ^1 regularization is necessary to ensure that the under-determined system has a unique solution; Second, the sparsity prior allows the learned representation to capture salient patterns of local descriptors; Third, the sparse coding can achieve much less quantization error than VQ. Accordingly, even with linear SVM classifier, ScSPM can outperform the nonlinear SPM approach by a large margin on benchmarks like Caltech-101 [Yang, Yu, Gong et al. (2009)].

2.3 Coding descriptors in LLC

Locality-constrained Linear Coding (LLC) was presented by Wang et al. [Wang, Yang, Yu et al. (2010)]. LLC incorporates locality constraint instead of the sparsity constraint in Eq. 2, which leads to several favorable properties, including better reconstruction, local smooth sparsity and analytical solution. Specifically, the LLC code uses the following criteria [Yuan, Ma and Yuille (2017)]:

$$\min_c \sum_{i=1}^N \|x_i - Bc_i\|^2 + \lambda \|d_i \odot c_i\|^2 \quad (3)$$

$$s.t. \mathbf{1}^T c_i = 1, \forall i \quad (4)$$

where \odot denotes the element-wise multiplication, and $d_i \in \mathbb{R}^M$ is the locality adaptor that gives different freedom for each basis vector proportional to its similarity to the input descriptor x_i . Specifically,

$$d_i = \exp\left(\frac{\text{dist}(x_i, B)}{\sigma}\right) \quad (5)$$

where $\text{dist}(x_i, B) = [\text{dist}(x_i, b_1), \dots, \text{dist}(x_i, b_M)]^T$, and $\text{dist}(x_i, b_j)$ is the Euclidean distance between x_i and b_j . σ is used for adjusting the weight decay speed for the

locality adaptor. Usually we further normalize d_i to be between (0, 1) by subtracting $\max(\text{dist}(x_i, B))$ from $\text{dist}(x_i, B)$. The constraint $1^T c_i = 1$ follows the shift-invariant requirements of the LLC code. Note that the LLC code in Eq. (5) is not sparse in the sense of l^0 norm, but is sparse in the sense that the solution only has few significant values. In practice, we simply threshold those small coefficients to be zero.

With the above description of LLC coding, we can see that the locality constrain will earn extra discrimination for the coding coefficients. Because, some structure information is introduced the coding by enforcing some coefficients with weak sample correlation to be zero, which will help to select more samples within the same subspace. Next, we will proposed a weighted sparse classification framework to improve the coding structure. And then, to explore the supervision information hidden in the data, a novel supervision constraint item about coding approximation are also proposed as the part of the weighted sparse classification framework.

3 Sparse image classification model based on low-rank supervision

3.1 Low-rank representation model

Low-rank matrix recovery is the development and promotion of the compressive sensing theory, using the rank as a measure of the matrix sparseness. The low-rank representation model has well capable of handling high-dimensional signal data. The model can be used to recover damaged observational samples and can find low-dimensional feature space from noisy samples. A matrix representation of the image data, the original observation data matrix $X \in R^{m \times n}$ and the model that uses the rank of matrix for sparse representation is as follows:

$$\min_x \text{rank}(X) \quad \text{s.t.} \quad \Phi(X) = b \quad (6)$$

where $\text{rank}(\bullet)$ is the rank of the matrix \bullet , $\Phi(\bullet)$ is a linear mapping function from original data matrix to observed variable, namely $R^{m \times n} \rightarrow R^p$ and $b \in R^p$. In the model 1, we get a low-rank constraint on matrix X . However, since the function $\text{rank}()$ in the model equation is discrete and non-convex, solving model 1 is an NP-Hard problem. In order to solve such problem, it is a common practice to convert this NP-Hard combinatorial optimization problem to corresponding convex optimization problems by transforming some objective function to some convex function. $\text{rank}()$ can also represent the number of non-zero singular values in the matrix, so we can consider $\min \text{rank}(X)$ is a l_0 -norm constraint problem of matrix singular value. That is to say, minimizing the l_0 -norm of vectors is to minimize the rank of the matrix. Therefore, the solution of minimizing matrix rank can refer to the relevant ideas and methods for solving the minimum l_0 -norm.

If the sample matrix X exists in several low-dimensional spaces, the sample matrix X can be linearly represented by a set of low-dimensional space bases A , that is, the sample itself is selected as the dictionary, the lower rank represents the coefficient matrix Z , and the Z represents each data vector The weight of a linear combination. A low-rank

constraint on the coefficient Z results in a low-rank representation of the sample matrix X as follows:

$$\min_Z \text{rank}(Z) \quad s.t. \quad X = XZ \quad (7)$$

which is a non-convex optimization problem. In the solution, we usually apply convex-relaxation method and replace $\text{rank}(Z)$ with the matrix kernel function $\|Z\|_*$. The following formula:

$$\min_Z \|Z\|_* \quad s.t. \quad X = XZ \quad (8)$$

In practical applications, the signal is likely to be polluted by noise, so the above equation can be further expressed as:

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1} \quad s.t. \quad X = XZ + E \quad (9)$$

where $\|\bullet\|_{2,1}$ is called $l_{2,1}$ -norm, λ is the regularization parameter, E is the residual. For different types of noise, the $l_{2,1}$ -norm has strong processing power.

In recent years, researchers have proposed many efficient optimization algorithms to solve the above problems, such as Iterative Thresholding, Exact Augmented Lagrange Multiplier, Inexact Augmented Lagrange Multiplier and so on. These algorithms can learn from l_1 -norm minimization, the following describes the Exact and Inexact Augmented Lagrange Multiplier algorithm. Before that, we need Augmented Lagrange Multiplier (ALM) algorithm.

The objective function for a constrained optimization problem can be described by:

$$\min f(X) \quad s.t. \quad h(X) = 0 \quad (10)$$

where f is a convex function, which represents mapping from R^n to R . h represents mapping from R^n to R^m . Then the definition of Augmented Lagrange Multiplier function is as follows:

$$L(X, Y, \mu) = f(X) + \langle Y, h(X) \rangle + \frac{\mu}{2} \|h(X)\|_F^2 \quad (11)$$

where Y is Lagrange multiplier. μ is known as the penalty coefficient that is a positive number. $\langle \bullet \rangle$ is inner product operator. When the values of Y and μ are appropriate, the objective function can be solved by Augmented Lagrange algorithm. Assume that $X = (A, E)$, $f(X) = \|A\|_* + \lambda \|E\|_1$, $h(X) = D - A - E$. The Eq. (11) can be expressed by this:

$$L(A, E, Y, \mu) = \|A\|_* + \lambda \|E\|_1 + \langle Y, D - A - E \rangle + \frac{\mu}{2} \|D - A - E\|_F^2 \quad (12)$$

which can be solved by using alternate ways to optimize the update. Let Y and E be fixed values solving minimized L to obtain A . Then fix Y and A , and solve for minimized L to obtain E . It is done iteratively until the convergent solution is eventually obtained that is called the Exact Augmented Lagrange Multiplier method. Since A and E are coupled to each other in Eq. (12), each iteration will be placed in the

same sub-problem during the iteration, which makes the solution feasible. In practice the Exact Augmented Lagrange Multiplier method requires multiple iterations, so the algorithm runs at a lower speed. On the basis of the former, an Inexact Augmented Lagrange Multiplier method is proposed, in which the A and E are updated once at each iteration to obtain the approximate solution of the sub-problem.

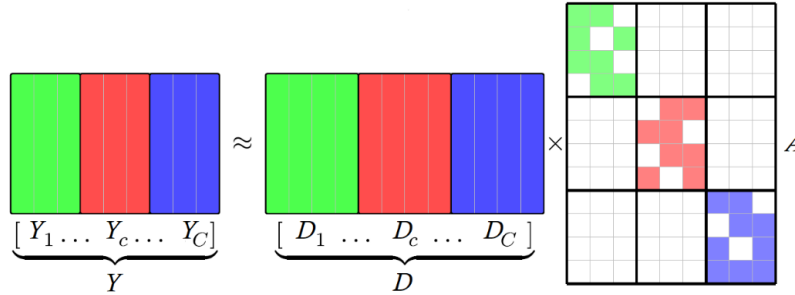


Figure 1: Ideal structure of the coefficient matrix in SRC

The core idea based on sparse representation is that the test sample is expressed by a linear combination of training samples. Because most non-zero elements have the same base corresponding to the test sample. In ideal case, we expect each sample to be in its own category subspace and all the category subspaces do not overlap. Specifically, given a category of images recorded as C and dictionaries $D=[D_1, \dots, D_C]$ which D_c contains training samples from the category $c=1, \dots, C$. That is meaning that a new sample y belonging to the category c can be represented by $y \approx D_c a^c$, a^c is a representation coefficient under a sub-dictionary D_c . So we can use the dictionary D to represent y : $y \approx Da = D_1 a^1 + \dots + D_c a^c + \dots + D_C a^C$. Most of the valid elements should be located in a^c , so the coefficient vector a^c should be sparse. From vector to matrix form, $Y=[Y_1, \dots, Y_c, \dots, Y_C]$ is the entire sample set matrix containing the category samples, the coefficient matrix A will be also sparse, and in the ideal case the matrix A is a block diagonal matrix as shown in Fig. 1.

3.2 Correlation matrix construction

Assume that the test sample matrix $Y=[y_1, y_2, \dots, y_n]$, y_n denotes the no. n test sample vector. Based on above analysis, we can use the sample as a dictionary and use the matrix Z to explore the correlation between samples to obtain the low-rank representation model of the test sample as follows:

$$\min_{Z, E} \|Z\|_* + \lambda \|E\|_1 \quad s.t. \quad Y = YZ + E \tag{13}$$

where $\|\cdot\|_*$ is kernel function, E is residual and λ is regularization parameters. Eq. 13 can be solved using the Augmented Lagrange Multiplier method which is introduced earlier to obtain a low-rank representation matrix Z for Y . Z contains some sample features.

3.3 Sparse image classification model based on low-rank supervision

To make better use of correlations between samples and obtain feature representations with better discrimination capability, we use the low-rank representation matrix Z to perform low-rank supervision constraints on the sparse coding process and propose a new low-rank supervised sparse coding model.

First, we can obtain a supervisory matrix from a low-rank representation matrix, as shown in the following equation:

$$W = \frac{|Z| + |Z^T|}{2} \quad (14)$$

where $|\cdot|$ represents the absolute value of each element in the matrix \cdot . Let the category-known sample matrix $X = [X_1, X_2, \dots, X_k]$ and $X_k = [x_k^1, x_k^2, \dots, x_k^m]$. x_k^m represents the m th sample vector belonging to the category number k . Using a self-describing category-known sample as a dictionary, we can get a new sparse coding model as follows:

$$\min_A \frac{1}{2} \|Y - XA\|_F^2 + \lambda \|P \odot A\|_1 + \eta \sum_{ij} \|A_i - A_j\|_2^2 W_{ij} \quad (15)$$

where A is feature representation matrix, λ and η are regularization parameters. In the above model, we have added the supervision constraint item $\sum_{ij} \|A_i - A_j\|_2^2 W_{ij}$ to increase the difference in coding between different categories, where i and j are the column numbers of the feature representation matrix A . P is the correlated weighted for the sparse coefficient. $P_{ij} = \exp\{-\|y_i - D_j\|_2^2\}$. \odot denotes the element-wise multiplication operator. W_{ij} is used to represent the elements of the row i and column j in the supervising matrix W . By solving the model 10, we can get the representation matrix A with stronger discrimination ability.

After obtaining the feature representation matrix A , the reconstruction error can be used to classify and the classification model as follows:

$$i_k = \arg \min \|y_k - X_i A_k^i\|_2^2 \quad (16)$$

where $k \in \{1, 2, \dots, n\}$, A_k^i is the feature code that represents the k th column in the matrix A corresponding to the i th sample. Find the Eq. (16) to minimize the reconstruction error label is the classification result, that is the formula for the minimum value of k .

4 Solving the sparse image classification model based on low-rank supervision

This section describes the process of solving the above low-rank supervised coding model in detail. The model 10 can be transformed as follows:

$$P(A) = \frac{1}{2} \|Y - XA\|_F^2 + \lambda \|P \odot A\|_1 + \eta \sum_{ij} \|A_i - A_j\|_2^2 W_{ij} \quad (17)$$

Eq. 17 is a non-convex problem. We can make an equidistant transformation to get the following based on $\sum_{ij} \|A_i - A_j\|_2^2 W_{ij} = \text{trace}(ALA^T)$:

$$P(A) = \frac{1}{2} \|Y - XA\|_F^2 + \lambda \|P \odot A\|_1 + \eta \text{trace}(ALA^T) \quad (18)$$

where $\text{trace}(\bullet)$ is trace of matrix, L is lagrange operator and $L = D - W$. The matrix D is a diagonal matrix whose elements on the diagonal correspond to the sum of all elements of the column vector in the matrix W , namely $D_{ii} = \sum_j W_{ij}$, i and j are the row and column numbers of the matrix. This problem is transformed into a convex optimization problem.

Let $G(A) = \frac{1}{2} \|Y - XA\|_F^2 + \eta \text{trace}(ALA^T)$ and $H(A) = \lambda \|P \odot A\|_1$, then Eq. (18) can be replaced with the following form:

$$P(A) = G(A) + H(A) \quad (19)$$

For the optimization solution, we can get the gradient solution formula of $G(A)$ as follows:

$$\nabla G(A) = -X^T(Y - XA) + \eta AL^T \quad (20)$$

where ∇ is a gradient indicator. According to Eq. (19) and Eq. (20), the iterative shrinking algorithm can be used to solve the original model (17), which iterative solution is as follows:

$$A^{k+1} = \arg \min \left\{ \|A - \nabla G(A^k)\|_F^2 + H(A) \right\} \quad (21)$$

As for the Eq. (21), it is a typical weighted l_1 -norm minimization problem, which can be solved with the numerical method in Guleryuz [Guleryuz (2006)]. And then, we can obtain the feature representation matrix A and use the classification model shown in model (16) to classify images.

5 Experimental results and analysis

5.1 Experimental data

For the weighted sparse image classification method based on low-rank representation proposed in this paper, we take the ORL and COIL-20 image datasets as the experimental datasets.

The ORL face dataset contains 400 face images, which were shot by 10 people under different lighting conditions, varying view and expressions. Each person has about 10 face images. Fig. 2 shows some face samples of the ORL dataset.



Figure 2: Two groups of face samples of ORL dataset

The COIL-20 image dataset is collected by Columbia Object Image Library, which contains images of 20 items. During shooting the image, the object is placed on a rotating plate with black background. Through the 360 degree rotation of the plate to change the attitude of the item relative to the camera. Each object in the dataset has 72 images with different angles, and some examples are shown in Fig. 3.



Figure 3: Example images of COIL-20 database

In the experiment, we implemented the sparse representation-based image classification methods SRC, CRC, LRSRC on the above two public datasets. The algorithm proposed in this paper is compared to verify.

5.2 Comparison and analysis of experiments

First of all, in order to verify the validity of the method proposed in this chapter, we use the ORL face image data set to display, and visualize the obtained low-rank representation matrix and its corresponding feature representation matrix. The low-rank representation matrix is shown in Fig. 4, and the characteristic representation matrix is shown in Fig. 5. From these two figures, it can be observed that the representation matrix has a distinct diagonal block structure, indicating that the representation matrix should theoretically have the discriminative ability of classification, the experimental result of the latter also confirmed this prediction. From the comparison of these two graphs, we can also find that compared with the low rank representation matrix, the elements in the diagonal matrix of the representation matrix are enhanced after being coded from the supervision limit to make it have a more obvious diagonal block structure, indicating better class resolution.

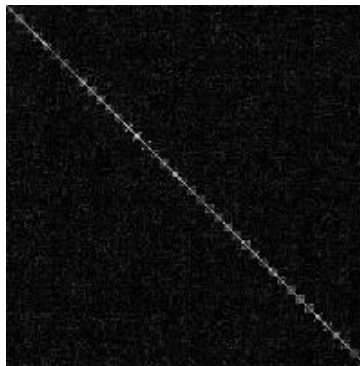


Figure 4: Visualization of low-rank representation matrix for ORL

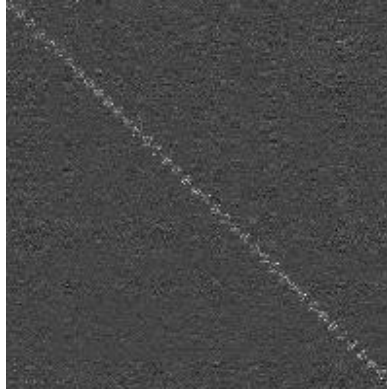


Figure 5: Visualization of sparse representation matrix for ORL

We performed simulation experiments on the above methods in the Extended YaleB face image dataset and COIL-20 image dataset and obtained the corresponding experimental results. We test five times for each method, and take the average accuracy as the final result, which is shown in Tab. 1. The number shown in the parentheses denotes the number of training samples.

Table 1: Average accuracy of each classification method

	ORL(5)	COIL-20(30)
SRC	$88.35 \pm 2.31\%$	$95.83 \pm 1.71\%$
CRC	$90.02 \pm 1.57\%$	$96.21 \pm 2.56\%$
LRSRC	$91.79 \pm 1.89\%$	$97.03 \pm 2.72\%$
Proposed Method	$93.81 \pm 1.96\%$	$97.93 \pm 2.06\%$

From Tab. 1, we can see that our proposed method shows higher accuracy classification result on all the experimental datasets over the comparison methods. It is verified that the weighted sparse coding with the low rank supervision can greatly improve the structure of coding to be more discriminative.

6 Summary

On the basis of the SRC method, this paper proposes a weighted sparse image classification method based on low-rank supervision. The proposed one strengthens the sample discrimination ability of the feature representation matrix, which makes full use of the correlation between samples to weighted code the image sparsely. Different types of sparse represent the degree of coding discrimination, which is believed to improve the accuracy of image classification. We replace the non-convex terms in the model equivalently, use the iterative shrinking algorithm to solve and verify the correctness through some experiments. Compared with related conventional classification methods, the method proposed in this paper is better than them.

Acknowledgement: This research is funded by the National Natural Science Foundation of China (61771154).

References

- Cao, J.; Zhang, K.; Luo, M.; Yin, C.; Lai, X.** (2016): Extreme learning machine and adaptive sparse representation for image classification. *Neural Networks*, vol. 81, no. C, pp. 91.
- Chen, Y.; Nasrabadi, N. M.; Tran, T. D.** (2011): Hyperspectral image classification using dictionary-Based sparse representation. *IEEE Transactions on Geoscience & Remote Sensing*, vol. 49, no. 10, pp. 3973-3985.
- Fadili, M. J.; Starck, J. L.; Murtagh, F.** (2013): Inpainting and zooming using sparse representations. *Computer Journal*, vol. 52, no. 1, pp. 64-79.
- Feng, Z.; Wang, M.; Yang, S.; Liu, Z.; Liu, L. Z. et al.** (2017): Superpixel tensor sparse coding for structural hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 4, pp. 1632-1639.
- Guleryuz, O. G.** (2006): Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-Part I: Theory. *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 55-71.
- Gao, S.; Tsang, W. H.; Chia, L. T.** (2010): Kernel sparse representation for image classification and face recognition. *Computer Vision-ECCV*, pp. 1-14.
- Hassairi, S.; Ejbali, R.; Zaied, M.** (2016): Sparse wavelet auto-encoders for image classification. *Digital Image Computing: Techniques and Application*, pp. 1-6.
- Huang, F. J.; Boureau, Y. L.; LeCun, Y.** (2007): Unsupervised learning of invariant feature hierarchies with applications to object recognition. *Computer Vision and Pattern Recognition*, pp. 1-8.
- Kulkarni, N.; Li, B.** (2011): Discriminative affine sparse codes for image classification. *Computer Vision and Pattern Recognition*, vol. 32, no. 14, pp. 1609-1616.
- Lee, H.; Battle, A.; Raina, R.; Ng, A. Y.** (2007): Efficient sparse coding algorithms. *Advances in Neural Information Processing Systems*, vol. 19, pp. 721-728.
- Li, C.; Ma, Y.; Mei, X.; Liu, C.; Ma, J.** (2017): Hyperspectral image classification with robust sparse representation. *IEEE Geoscience & Remote Sensing Letters*, vol. 13, no. 5, pp. 641-645.
- Liu, S.; Li, L.; Peng, Y.; Qiu, G.; Lei, T.** (2017): Improved sparse representation method for image classification. *IET Computer Vision*, vol. 11, no. 4, pp. 319-330.
- Liu, L.; Tran, T. D.; Sang, P. C.** (2016): Partial face recognition: A sparse representation-based approach. *IEEE International Conference on Acoustics*, no. 5, pp. 2389-2393.
- Liu, S.; Fu, W.; He, L.** (2017): Distribution of primary additional errors in fractal encoding method. *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 5787-5802.
- Liu, S.; Pan, Z.; Fu, W.; Cheng, X.** (2017): Fractal generation method based on asymptote family of generalized mandelbrot set and its application. *Journal of Nonlinear Sciences and Applications*, vol. 10, no. 3, pp. 1148-1161.

Poultney, C.; Chopra, S.; Cun, Y. L. (2007): Efficient learning of sparse representations with an energy-based model. *Advances in Neural Information Processing Systems*, pp. 1137-1144.

Song, B.; Li, J.; Dalla, M. M.; Li, P.; Plaza, A. et al. (2014): Remotely sensed image classification using sparse representations of morphological attribute profiles. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 5122-5136.

Shivakumar, G. S.; Natarajan, S.; Murthy, K. S. (2015): Implementation of multikernel sparse representation on remote sensing image classification. *Applied and Theoretical Computing and Communication Technology*, pp. 331-337.

Sun, X.; Qu, Q.; Nasrabadi, N. M.; Tran, T. D. (2014): Structured priors for sparse-representation-based hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 7, pp. 1235-1239.

Tang, X.; Chen, J.; Liu, Y.; Sun, Q. (2015): Hyperspectral image classification by fusing sparse representation and simple linear iterative clustering. *Journal of Applied Remote Sensing*, vol. 9, no. 1, pp. 769-780.

Tang, Y.; Huang, S.; Xue, A. (2016): Sparse representation based binary hypothesis model for hyperspectral image classification. *Mathematical Problems in Engineering*, vol. 2016, no. 11, pp. 1-10.

Thiagarajan, J. J.; Spanias, A. (2012): Learning dictionaries for local sparse coding in image classification. *Signals, Systems and Computers*, pp. 2014-2018.

Wright, J.; Ganesh, A.; Zhou, Z.; Wagner, A.; Ma, Y. (2009): Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227.

Wright, J.; Ma, Y.; Mairal, J.; Sapiro, G.; Huang, T. S. et al. (2010): Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1031-1044.

Wagner, A.; Wright, J.; Ganesh, A.; Zhou, Z. H.; Mobahi, H. et al. (2012): Toward a practical face recognition system: Robust alignment and illumination by sparse representation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 34, no. 2, pp. 372-386.

Wang, L.; Xie, X.; Li, W.; Du, Q.; Li, G. (2015): Sparse feature extraction for hyperspectral image classification. *2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)*, pp. 1067-1070.

Wang, J.; Yang, J.; Yu, K.; Lv, F. J.; Huang, T. et al. (2010): Locality-constrained linear coding for image classification. *Computer Vision and Pattern Recognition*, vol. 119, no. 5, pp. 3360-3367.

Yang, M. C.; Chu, C. T.; Wang, Y. C. F. (2010): Learning sparse image representation with support vector regression for single-image super-resolution. *IEEE International Conference on Image Processing*, pp. 1973-1976.

Yuan, G.; Ma, J.; Yuille, A. L. (2017): Semi-supervised sparse representation based classification for face recognition with insufficient labeled samples. *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2545.

- Yang, J.; Wright, J.; Huang, T. S.; Ma, Y.** (2010): Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, vol.19, no. 11, pp. 2861-2873.
- Yang, M.; Zhang, L.; Yang, J.; Zhang, D.** (2011): Robust sparse coding for face recognition. *Computer Vision and Pattern Recognition*, vol. 42, no. 7, pp. 625-632.
- Yang, J.; Yu, K.; Gong, Y.; Huang, T.** (2009): Linear spatial pyramid matching using sparse coding for image classification. *Computer Vision and Pattern Recognition*, pp. 1794-1801.
- Zeiler, M. D.; Krishnan, D.; Taylor, G. W.; Fergus, R.** (2010): Deconvolutional networks. *Computer Vision and Pattern Recognition*, vol. 238, no. 6, pp. 2528-2535.
- Zeng, Z.; Li, H.; Liang, W.; Zhang, S.** (2010): Similarity-based image classification via kernelized sparse representation. *2010 17th IEEE International Conference on Image Processing*, vol. 119, no. 5, pp. 277-280.
- Zhang, X.; Song, Q.; Gao, Z. Y.; Zheng, Y. G.; Weng, P. et al.** (2016): Spectral-spatial feature learning using cluster-based group sparse coding for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4142-4159.
- Zhang, E.; Zhang, X.; Liu, H.; Jiao, L.** (2015): Fast multifeature joint sparse representation for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 7, pp. 1397-1401.
- Zhang, L.; Yang, M.; Feng, X.** (2011): Sparse representation or collaborative representation: Which helps face recognition. *International Conference on Computer Visio*, vol. 2011, no. 5, pp. 471-478.
- Zhang, H.; Zhang, Y.; Huang, T. S.** (2013): Pose-robust face recognition via sparse representation. *Pattern Recognition*, vol. 46, no. 5, pp. 1511-1521.