International Journal of Multimedia and Ubiquitous Engineering
Vol. 9, No. 12 (2014), pp. 269-280
http://dx.doi.org/10.14257/ijmue.2014.9.12.25

Research on Different Representation Methods for Classification

Jiangshu Wei, Xiangjun Qi and Mantao Wang

College of Information and Engineering Technology, Sichuan Agricultural University, Ya'an, 625014, China
sicauwjs@163.com, qxj@sicau.edu.cn, mantaowang@163.com

Abstract

Under today's big data environment, with the rapid development of computer network technology and information technology, data mining is becoming more and more important in computer science. Classification is one of the most important aspects in data mining research Field. Recently, representation methods, such as sparse representation and low rank representation, have been much concerned. They both have wide applications in scientific and engineering fields. However, sparse representation and low rank representation include many methods, although these methods have their own characteristics, they are all effective for handling classification problems. This paper focuses on the performance comparison of different representation methods currently used in handling classification problems and views other conventional methods that can be applied in this field.

Keywords: *Sparse representation, Low rank representation, Classification*

1. Introduction

Data mining is a hot topic in the field of artificial intelligence and machine learning. Data mining is a process, which reveals the implicit, previously unknown and potentially useful information from the large amount of data in the database. With the rapid development of information technology and computer network, Internet data and resources show massive features. In order to manage these massive information effectively, data mining is becoming a hot research field increasingly. However, data mining techniques include many aspects. Classification is one of the most important aspects in data mining.

Given some training samples from multiple classes, the aim of classification task is to assign one of the class labels to a test sample. Classification has been widely used in scientific and engineering fields, such as pattern recognition, data mining, computer vision, etc.

There are also many conventional methods for handling classification problems, such as Nearest Neighbor, Nearest subspace classifier, Linear SVM, *etc.*, Recently, there has been an increasing interest in representation theory. Representation includes many methods, such as sparse representation, low rank representation and collaborative representation, *etc.*, although these methods have their own characteristics, they are all effective for handling classification problems.

However, representation methods also include many algorithms. These algorithms also have advantages and disadvantages. The aim of this paper is to compare the state of the art algorithms in representation theory. We try to clarify the similarities among different representation algorithms and reveal the differences of them.

Wright, *et al.,* [1] proposed the sparse representation classifier (SRC) method for classification. SRC is a classical method, the solution of SRC can be obtained by using l_1

ISSN: 1975-0080 IJMUE

norm minimization. SRC boosts the research of sparsity. Many application problems are also solved by sparse representation methods.

Based on SRC, some paper proposed other methods. Elhamifar and Vidal [2, 3] proposed a Block-Sparse representation for face recognition. Chi and Porikli [4] proposed a Collaborative Representation Optimized Classifier (CROC). Zhang, *et al.*, [5] argued that not the sparse representation, but the usage of collaborative representation is more important with the success of the SRC. They proposed a kind of Collaborative representation classification (CRC_RLS) [5] method, by using l_2 norm minimization.

The low rank representation (LRR) was proposed by Liu [6], which is different from the sparse representation. The aim of sparse representation is to obtain the sparsest solution of each test sample respectively. However, unlike sparse representation, the aim of low rank representation is to find the lowest rank representation of all the test samples jointly. Low rank representation can be also used to handle the classification problems.

This paper is organized as follows. In section 2, sparse representation methods are reviewed. In section 3, collaborative representation methods are reviewed. In section 4, some conventional methods used for classification are reviewed. In section 5, low rank representation methods are reviewed. In section 6, we review the advantages and disadvantages of these methods. In section 7, some experiments on digit recognition and face recognition have been done in order to compare these representation methods. Section 8 concludes this paper.

2. Sparse Representation Methods for Classification

Sparse representation is based on the concept of mathematical norm, which has a long history. With the rapid development of norm minimization methods, such as OMP [7], BP [8] and many other algorithms Sparse representation has been much concerned for many years. Recently, sparse representation has obtained many applications in signal processing, image feature extraction, pattern recognition, image denoising, *etc.*, [9-25].

Pati proposed Orthogonal Matching Pursuit (OMP) method [7], which is an optimization method by using l_0 norm minimization. Chen proposed Basis Pursuit (BP) method [8], which is an optimization method by using l_1 norm minimization. Wright, *et al.*, [1] proposed SRC for classification. SRC is a classical method, it is an optimization method by using l_1 norm minimization. Based on SRC, some papers proposed many other methods. Elhamifar *et al*. [2, 3] proposed a Block-Sparse representation. Chi, *et al*., [4] proposed a Collaborative Representation Optimized Classifier (CROC). Yang and Zhang [5] proposed a kind of Collaborative representation classification (CRC RLS) [5] method, by using l_2 norm minimization.

2.1. Multi-class Classification

If there are K classes, and there are n_i training data from the ith class formed a matrix as $A_i \square [a_{i1,} a_{i2,} ... a_{in_i}] \in R^{m \times n_i}$. A is denoted by the collection of all training samples: $A = [A_1, A_2, ... A_k]$. If given a test sample $y \in R^m$, the aim of multi class classification is to identify y belongs to which class [1, 4, 5, 11].

2.2. Orthogonal Matching Pursuit

Orthogonal Matching Pursuit (OMP) is a well known algorithm, which was proposed by Pati in 1993 [7]. It is a sparse representation method to find the approximate solution of l_0 norm minimization. The steps of OMP algorithm are as follows [11]:

Task: Find the approximate solution of $P_{l_0}: \min_x \|x\|_0$ s.t. $y = Ax$

1) Input:

A matrix concatenated by training samples $A = [A_1, A_2, ...A_k] \in R^{m \times n}$ for *k* classes, a test sample $y \in R^m$, a error threshold ε_0.

2) Initiation:

Initialize *k*=0, and set: The initial solution $x^0 = 0$. The initial residual $r^0 = y - Ax^0 = y$. The initial solution support $S^0 = Support\{x^0\} = \Phi$.

3) Iteration:

a) Compute the errors $\varepsilon(j) = \min_{z_j} = \|a_j z_j - r^{k-1}\|_2^2$ for all *j* using the optimal choice $z_j^* = a_j^T r^{k-1} / \|a_j\|_2^2$.

b) Find a minimizer j_0 of $\varepsilon(j)$: $\forall j \notin S^{k-1}, \varepsilon(j_0) \le \varepsilon(j)$,update $S^k = S^{k-1} \cup \{j_0\}$.

c) Compute x^k, the minimizer of $\|Ax - y\|_2^2$ subject to Support $\{x\} = S^k$.

d) Compute $r^k = y - Ax^k$.

e) If $\|r^k\|_2 < \varepsilon_0$, stop. Otherwise, perform another iteration.

4) Output: The solution is x^k after *k* iterations.

The aim of OMP is to obtain the approximate solution of l_0 norm minimization. However, OMP is a greedy algorithm, so its computational complexity is high. Furthermore, it is sensitive to noise.

2.3. Sparse Representation-based Classification

The l_0 *norm* optimization is NP-hard, a convex relaxation of it can be obtained by replacing them l_0 with l_1 norm. Sparse representation-based classification (SRC) is just a classical method by using l_1 norm minimization. SRC was proposed by Wright, *et al.*, [1]. It is a classical method for classification. Based on SRC, some papers proposed many another methods. The steps of SRC algorithm are as follows [1]:

Task: Find the solution of $P_{l_1}: \min_x \|x\|_1$ s.t. $y = Ax$

1) Input:

A matrix concatenated by training samples $A = [A_1, A_2, ...A_k] \in R^{m \times n}$ for *k* classes, a test sample $y \in R^m$.

2) Solve the l_1 norm minimization problem:

$x_1 = \arg\min_x \|x\|_1$ s.t. $y = Ax$

3) Compute the residuals:

$r_i(y) = \|y - A\delta_i(x_1)\|_2^2$, for i=1,2,…,k.

4) Output:

$identity(y) = \arg\min_i r_i(y)$.

The SRC method looks for the sparsest representation of a test sample by using l_1 norm minimization. The classification results of SRC are good. However, the sparsest representation does not mean obtaining the best classification results. Furthermore, from l_1 norm minimization, the SRC cannot obtain closed form solution, so its computational complexity is high.

2.4. Structured Sparse Representation

The dictionary of the training samples has a structure; it means data from each class forming a few blocks of the dictionary. However, the SRC method only looks for the sparsest representation of a test sample; it does not take into account of the similarity of these samples. The solution of l_1 norm minimization does not indicate the space distribution feature of the samples. Thus, there still remain some problems about multi-class classification using SRC method. Elhamifar proposed structured sparse representation method (SSR) [2, 3], the structured sparse representation looks for a representation, which the test sample involves the minimum number of blocks from the dictionary. The non-convex optimization programs are as follows [2, 3]:

$$P_{l_q/l_0} : min \sum_{i=1}^{n} J(\|x[i]\|_q > 0) \text{ s.t. } y = Ax ,$$

and

$$P'_{l_q/l_0} : min \sum_{i=1}^{n} J(\|A[i]x[i]\|_q > 0) \text{ s.t. } y = Ax ,$$

$J(\cdot)$ is the indicator function, $q \geq 1$, $x[i] \in R^{m_i}$ are the entries of x corresponding to the i-th block of the dictionary. However, the optimization program P_{l_q/l_0} is NP-hard, a l_1 relaxation of it can be given as follows:

$$P_{l_q/l_1} : min \sum_{i=1}^{n} \|x[i]\|_q \text{ s.t. } y = Ax ,$$

and

$$P'_{l_q/l_1} : min \sum_{i=1}^{n} \|A[i]x[i]\|_q \text{ s.t. } y = Ax .$$

Output: $identity(y) = \arg\min_i \|y - B[i]c^*[i]\|_2$

3. Collaborative Representation Methods for Classification

3.1. Collaborative Representation based on Methods for Classification

The solution of sparse representation methods can be only obtained by the l_1 norm minimization. However, it cannot obtain the closed form solution from l_1 norm minimization. The computational complexity of l_1 norm minimization is a little high. Thus, some authors proposed regularized least square method using l_2 norm minimization. Collaborative representation classification (CRC_RLS) is proposed by Zhang and Yang [5], which is a typical method by l_2 norm minimization. The steps of CRC_RLS algorithm are as follows [5]:

Task: Find the solution of $P_{l_1}: \min_x \|x\|_1$ s.t. $y = Ax$

1) Input:

A matrix concatenated by training samples $A = [A_1, A_2, ...A_k] \in R^{m\times n}$ for k classes, a test sample $y \in R^m$.

2) Solve the l_2 norm minimization problem:

$x_2 = \arg\min_x \|x\|_2$ s.t. $y = Ax$

3) Compute the residuals:

$r_i(y) = \|y - A\delta_i(x_1)\|_2^2$, for i=1,2,…,k.

4) Output:

$identity(y) = \arg\min_i r_i(y)$.

From l_2 norm minimization, a closed form solution can be obtained, which gives $x = (A^T A + \lambda I)^{-1} A^T y$. The classification results of CRC_RLS are good. Furthermore, from the l_2 norm minimization, its computational complexity is low. However, the obtained solution is not sparse.

3.2. Collaborative Representation Optimized Classifier

Combined the Nearest Subspace Classifier (NSC) [26] and the Collaborative Representation based Classifier (CRC), Chi and Porikli proposed a collaborative representation optimized classifier (CROC), which depends on the trade-off between the NSC and CRC. The residual of CROC for each class is calculated as:

$$r_i(\lambda) = r_i^{NS} + \lambda r_i^{CR}, \quad \text{for} \quad i = 1, ..., k,$$

where $\lambda \geq 0$. If the ith residual is the minimal, then CROC assigns the test sample to the ith class.

4. Conventional Methods used for Classification

4.1. Nearest Neighbors

Nearest neighbors (NN) was first proposed by Cover and Hart for classification [27]. It was developed to be K nearest neighbor classifier subsequently [28]. This method is a conventional method, which is familiar to us. It can be used in digital image processing, pattern recognition, data mining *etc.*, Zhang and Yang combined the idea of NN and SRC, they presented KNN-SRC method [28]. The idea of NN is very simple, the steps of NN algorithm are as follows [28]:

With a test sample y, for i = 1, 2, ..., k, j = 1, 2, ..., n_i, compute the residuals $r_{ij}(y) = \|y - a_{ij}\|_2^2$. If a residual r_{ij} is the smallest, the NN will judge the test sample y belongs to the i-th class.

NN is a very simple method used for classification; its computational complexity is low. However, the classification results of NN are poor.

4.2. Nearest Subspace Classifier

Nearest subspace classifier (NSC) was proposed by Lee [26]. Chi and Porikli also utilized this method in their paper. The steps of NSC algorithm are as follows [26]:
For *i*= 1, 2, ..., *k*, there are *K* classes, there are n_i training data from the *i*-th class formed a matrix as $A_i \triangleq [a_{i1}, a_{i2}, ...a_{in_i}] \in R^{m \times n_i}$. A_i span a subspace. Compute the residual $r_i^{NS} = \min \|y - A_i x_i\|_2^2$. If the *i*-th residual is the smallest, NSC assigns the test sample y to the *i*-th class.

5. Low Rank Representation Methods

Liu, *et al.*, [6] proposed the low rank representation method, which is different from sparse representation. The aim of the sparse representation is to obtain the sparsest solution of each test sample respectively. However, unlike sparse representation, the aim of low rank representation is to find the lowest rank representation of all the test samples jointly. Low rank representation can be also used to handle the classification problems.

With a set of sample vectors $X = [x_1, x_2, ..., x_n]$, every column is a sample. Each sample can be represented by the linear combination of the basis in a dictionary $A = [a_1, a_2, ..., a_m]$. That is, $X = AZ$. $Z = [z_1, z_2, ..., z_n]$ is the coefficient matrix, each z_i is the representation of x_i. Our aim is to capture the global structure of X. However, sparse representation cannot capture the global structure of X, low rank representation is a more appropriate criterion for capturing the global structure of X. In other words, our aim is to look for a representation Z by solving the following problem:

$$\min_Z \quad rank(Z) \text{ s.t. } X = AZ, \qquad (1)$$

however, due to the discrete property of the rank function, the problem (1) is hard to solve. Thus, the problem (1) can be substituted for solving the problem (2):

$$\min_Z \|Z\|_* \text{ s.t. } X = AZ. \qquad (2)$$

For problem (2), $\|\bullet\|_*$ is the nuclear norm of a matrix, the definition of the nuclear norm is the sum of the singular values of the matrix.

The problem (2) can be solved using Augmented Lagrange Multiplier (ALM) algorithm [29-33], which is a classical method for solving the low rank representation (LRR) problem. LRR can also handle supervised classification problems as SRC. However, there are some problems with LRR. First, it cannot obtain closed form solution from the ALM algorithm. Second, there are too many parameters with the ALM algorithm. Third, the convergence property of ALM cannot be analyzed in detail.

6. Experiments

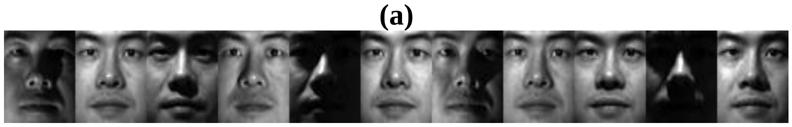
In this section, some experiments on face recognition and digit recognition are presented to show the accuracy of classification. We focus on the comparison of different representation methods mentioned above. Three databases, including Extended-YaleB [1, 4, 5], AR [1, 4, 5] and MNIST Handwritten Digits database [4], are used to test the performance of some methods, including SRC, CRC_RLS and NN. Our experiments focus on the performance comparison of different methods.

6.1. Face Recognition

These methods are tested for comparing the recognition rate. Recognition rate is a percentage, which denotes how many test samples can be classified correctly for all the test samples. Higher recognition rate means the performance of this method is better. In our experiments, the Eigenface is used as preprocessing in feature extraction.

1) Extended Yale-B database: The Extended Yale-B database contains 2414 frontal face images of 38 individuals [1, 4, 5]. The images were cropped and normalized to 54×48. A few images of Extended Yale-B database are shown in Figure 1. Table 1 shows the recognition rates versus feature dimension by SRC, CRC_RLS and NN.

(a)



(b)

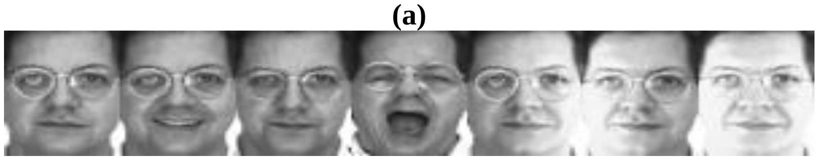
Figure 1. (a) Some Training Samples from the Extended Yale-B Database; (b) Some Test Samples from the Extended Yale-B Database

Table 1. The Recognition Results of Different Methods on the Extended Yale-B Database

Dimension	80	100	120	150	200
NN	69.24%	71.78%	72.96%	74.05%	75.41%
SRC	96.01%	96.37%	96.41%	96.5%	97.19%
CRC-RLS	94.74%	95.64%	95.92%	96.28%	97.01%

2) AR database: The AR database contains about 4000 frontal images for 126 individuals [1, 4, 5]. These images are captured under different illuminations, expressions and facial disguises. The images are cropped to size 60 × 43. A few images of AR database are shown in Figure 2. Table 2 shows the recognition rates versus feature dimension by SRC, CRC_RLS and NN.

(a)



(b)

Figure 2. (a) Some Training Samples from the AR Database; (b) Some Test Samples from the AR Databas

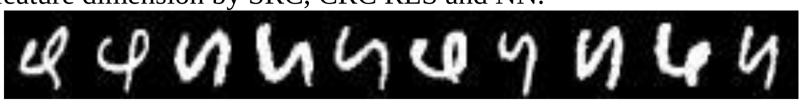
Table 2. The Recognition Results of Different Methods on the AR Database

Dimension	10	15	20	25	30
NN	49.07%	54.08%	58.23%	60.52%	61.80%
SRC	38.05%	50.36%	58.08%	64.95%	70.96%
CRC-RLS	19.46%	36.34%	45.92%	56.51%	64.38%

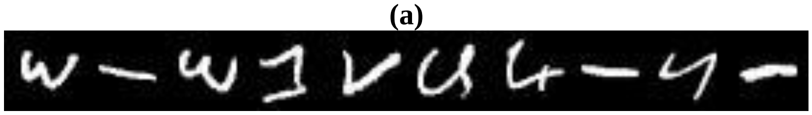
6.2. Digit Recognition

The MNIST handwritten digits database is used to test the property of these methods. The dimension of each image is 28 × 28. Every image, which is a 8 bit gray scale image from 0 to 9 [4].

For the MNIST handwritten digits database, which has a training set of 60,000 samples, and a test set of 10,000 samples of each class? For our experiment, 10 training samples are randomly selected from each class, 10 test samples are also randomly selected from each class. A few images of MNIST database are shown in Figure 3. Table 3 shows the recognition rates versus feature dimension by SRC, CRC RLS and NN.



(a)



(b)

Figure 3. (a) Some Training Samples from the MNIST Database; (b) Some Test Samples from the MNIST Database

Table 3. The Recognition Results of Different Methods on the MNIST Database

Dimension	50	60	70	80	90
SRC	61%	60%	55%	61%	62%
CRC-RLS	59%	58%	58%	59%	57%

7. The Comparison of Different Methods

As mentioned above, there are many representation methods used for handling the classification problems. However, these methods all have their own characteristics, they also have advantages and disadvantages. The reviews of these representation methods are listed in Table 4.

Table 4. The Reviews of Different Representation Methods used for Classification

Algorithm	Advantages	Disadvantages
OMP	Approximate solution of l_0 norm minimization	Greedy algorithm, computational complexity is high, sensitive to noise
SRC	Classification results are good	Cannot obtain closed form solution, computational

		complexity is high
SSR	Classification results are good	Cannot obtain closed form solution, computational complexity is high
CRC_RLS	Classification results are good, computational complexity is low	The obtained solution is not sparse
NN	Computational complexity is low	Classification results are poor
NSC	Classification results are good, computational complexity is low	Sensitive to noise
LRR	Can capture the global structure of samples	Cannot obtain closed form solution, too many parameters, the convergence property of ALM cannot be analyzed in detail

8. Conclusions

Under big data environment, data mining technologies are increasingly becoming a hot research field. However, data mining technologies include many aspects; classification is one of the most important aspects of data mining. Representation methods are very effective for handling classification problems. However, these methods all have their own characteristics; they also have advantages and disadvantages. In the future work, we should combine the advantages of currently used methods. It means that, the improved representation methods, which the classification results should be good and the computational complexity should be low. These methods are expected to apply in classification area.

Acknowledgements

This work was supported by the key fund project of Sichuan Educational department (NO.14ZA0005).

References

[1]. J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma, “Robust face recognition via sparse representation”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 31, no. 2, **(2009),** pp. 210–227.

[2]. E. Elhamifar and R. Vidal, “Robust classification using structured sparse representation”, In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, **(2011),** pp. 1873–1879.

[3]. E. Elhamifar and R. Vidal, “Block-sparse recovery via convex optimization”, Signal Processing, IEEE Transactions on, vol. 60, no. 8, **(2012),** pp. 4094–4107.

[4]. Y. Chi and F. Porikli, “Connecting the dots in multiclass classification: From nearest subspace to collaborative representation”, In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, **(2012),** pp. 3602–3609.

[5]. L. Zhang, M. Yang and X. Feng, “Sparse representation or collaborative representation: Which helps face recognition? In Computer Vision (ICCV)”, 2011 IEEE International Conference on, **(2011),** pp. 471–478.

[6]. G. Liu, Z. Lin and Y. Yu, “Robust subspace segmentation by low-rank representation,” in Proceedings of the 27th International Conference on Machine Learning (ICML-10), **(2010)**, pp. 663-670.

[7]. Y. C. Pati, R. Rezaiifar and P. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition”, In Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on, **(1993),** pp. 40–44.

[8]. S. Chen and D. Donoho, “Basis pursuit”, In Signals, Systems and Computers, 1994. 1994 Conference Record of the Twenty-Eighth Asilomar Conference on IEEE, vol. 1, **(1994),** pp. 41–44.
[9]. A. Adler, M. Elad and Y. Hel-Or, “Fast subspace clustering via sparse representations”, Department of Computer Science, Technion, Tech. Rep, **(2011)**.
[10]. M. Aharon, M. Elad and A. Bruckstein, “K-svd: an algorithm for designing over complete dictionaries for sparse representation”, Signal Processing, IEEE Transactions on, vol. 54, no. 11, **(2006),** pp. 4311–4322.
[11]. A. M. Bruckstein, D. L. Donoho and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images”, SIAM review, vol. 51, no. 1, **(2009),** pp. 34–81.
[12]. B. Cheng, J. Yang, S. Yan, Y. Fu and T. S. Huang, “Learning with l1-graph for image analysis”, Image Processing, IEEE Transactions on, vol. 19, no. 4, **(2010),** pp. 858–866.
[13]. S. F. Cotter, “Sparse representation for accurate classification of corrupted and occluded facial expressions”, In Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on, **(2010),** pp. 838–841.
[14]. M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries”, Image Processing, IEEE Transactions on, vol. 15, no. 12, **(2006),** pp. 3736–3745.
[15]. E. Elhamifar and R. Vidal, “Sparse manifold clustering and embedding”, Advances in Neural Information Processing Systems, vol. 24, **(2011),** pp. 55–63.
[16]. E. Elhamifar and R. Vidal, “Sparse subspace clustering: Algorithm, theory, and applications”, arXiv preprint arXiv, vol. 1203, no. 1005, **(2012)**.
[17]. J.-J. Fuchs, “Recovery of exact sparse representations in the presence of bounded noise”, Information Theory, IEEE Transactions on, vol. 51, no. 10, **(2005),** pp. 3601–3608.
[18]. T. Guha and R. K. Ward, “Learning sparse representations for human action recognition”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 34, no. 8, **(2012),** pp. 1576–1588.
[19]. J. Mairal, F. Bach, J. Ponce and G. Sapiro, “Online dictionary learning for sparse coding”, In Proceedings of the 26th Annual International Conference on Machine Learning ACM, **(2009),** pp. 689–696.
[20]. S. T. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding”, Science, vol. 290, no. 5500, **(2000),** pp. 2323–2326.
[21]. L. K. Saul and S. T. Roweis, “Think globally, fit locally: unsupervised learning of low dimensional manifolds”, The Journal of Machine Learning Research, vol. 4, **(2003),** pp. 119–155.
[22]. M. Stojnic, F. Parvaresh and B. Hassibi, “On the reconstruction of block-sparse signals with an optimal number of measurements”, Signal Processing, IEEE Transactions on, vol. 57, no. 8, **(2009),** pp. 3075–3085.
[23]. J. A. Tropp, “Greed is good: Algorithmic results for sparse approximation”, Information Theory, IEEE Transactions on, vol. 50, no. 10, **(2004),** pp. 2231–2242.
[24]. J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang and S. Yan, “Sparse representation for computer vision and pattern recognition”, Proceedings of the IEEE, vol. 98, no. 6, **(2010),** pp. 1031–1044.
[25]. M. Soltanolkotabi, E. Elhamifar and E. J. Candes, “Robust Subspace Clustering”, arXiv preprint arXiv, vol. 1301, no. 2603, **(2013)**.
[26]. K. Lee, J. Ho and D. Kriegman, “Acquiring linear subspaces for face recognition under variable lighting”, In Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 27, no. 5, **(2005)**, pp. 684–698.
[27]. Cover, Thomas, Hart and Peter, “Nearest neighbor pattern classification”, In Information Theory, IEEE Transactions on, vol. 13, no. 1, **(1967)**, pp. 21–27.
[28]. N. Zhang and J. Yang, “K nearest neighbor based local sparse representation classifier”, In Pattern Recognition (CCPR), 2010 Chinese Conference on, IEEE, **(2010)**, pp. 1-5.
[29]. G. Liu and S. Yan, “Latent low-rank representation for subspace segmentation and feature extraction,” inComputer Vision (ICCV), 2011 IEEE International Conference on, IEEE, **(2011)**, pp. 1615-1622.
[30]. Z. Shi, J. Han, T. Zheng and S. Deng, “Online Learning for Classification of Low-rank Representation Features and Its Applications in Audio Segment Classification,” inarXiv preprint arXiv, vol. 1112, no. 4243, **(2011)**.
[31]. G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu and Y. Ma, “Robust recovery of subspace structures by low-rank representation,” in Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, no. 1, **(2013)**, pp. 171-184.
[32]. G. Ye, D. Liu, I. Jhuo and S. Chang, “Robust late fusion with rank minimization,” in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE, **(2012)**, pp. 3021-3028.
[33]. Z. Lin, M. Chen and Y. Ma, “The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices,” arXiv preprint arXiv, vol. 1009, no. 5055, **(2010)**.

Authors

Jiangshu wei

He received his B.S. degree in Sichuan Normal University in 2003. He received his M.S. degree in University of Electronic Science and Technology of China in 2006. Now, he is a lecturer at the College of information engineering, Sichuan Agricultural University. He is also a Ph.D candidate in the college of computer science at Sichuan University. His main research interests include data mining, pattern recognition and machine learning.

Xiangjun Qi

She received her M.S. degree in Sichuan Agricultural university in 2002. Now, she is an associate professor at the College of information and engineering technology, Sichuan Agricultural University. Her main research interests include mechanical and electrical integration and intelligent controls.

Mantao Wang

He received the B.S. degree in Computer Science and Technology from University of Electronic Science and Technology of China in 1997. He is a lecturer at the School of information engineering, Sichuan Agricultural University. His research interests include image processing and pattern recognition.

