Ideal code constrained supervised sparse coding

Wenjing Liao[a], Robert Williams[b]
[a] School of Informatics, Guizhou University of Finance and Economics, Guiyang 550004, China
Email: liaowenjing1020@163.com
[b] New Mexico Highlands University, Las Vegas, NM 87701, United States
Email: robert_williamsr954@yahoo.com

***Abstract*— In this paper, we proposed a novel sparse coding algorithm by using the class labels to constrain the learning of codebook and sparse code. We not only use the class label to train the classifier, but also use it to construct class conditional codewords to make the sparse code as discriminative as possible. We first construct ideal sparse codes with regarding to the class conditional codewords, and then constrain the learned sparse codes to the ideal sparse codes. We proposed a novel loss function composed of parse reconstruction error, classification error, and the ideal sparse code constrain error. This problem can be optimized by using the transitional KSVD method. In this way, we may learn a discriminative classifier and a discriminative codebook simultaneously. Moreover, using this codebook, the learnt the sparse codes of the same class are similar to each other. Finally, exhaustive experimental results show that the proposed algorithm outperforms other sparse coding methods.**

***Index Terms*— Sparse coding, Codebook, Class labels, K-SVD**

I. Introduction

SPARSE coding (SC) [1]–[7] is a popular representation method, and it has been used to many problems in various fields, especially computer vision fields [8]–[11]. Given a codebook U with few codewords, and a input data sample $\mathbf{x}$, SC approximates $\mathbf{x}$ as a sparse linear combination of the codewords in U. It is a special case of bag-of-features methods [12]–[17], which use a codebook to represent the data samples. The learning of codebook from a training set is very important in this method [18]–[22]. Actually, instead of learning the codebook, other methods can also be used to generate the codebook. For example, by performing the Fourier [23]–[26] or wavelet [27]–[30] transformations, we also can obtain some bases as the codebook. Another alternative method is to use the whole training set as the codebook, but it is very time-consuming to compute the sparse codes when the codebook is large. It could be improved further by applying K-means clustering [31]–[34] to the training set and then using the cluster centers as the codewords. Recently, a novel codebook learning method has been proposed in [35], which is called "K-SVD". It generalizes the K-means clustering and learns the codebook from a training dataset. It has been proven to be very useful in many computer vision problems, including image compression [36]–[40], image recognition [41]–[44], and image retrieval [45]–[48]. Moreover, [49]–[52] proposed to learn the codebook using Lagrange dual and the sparse coding by the feature-sign search algorithm [53], [54]. These codebook generation methods are summarized in Figure 1.

All these methods of codebook generation only consider obtaining a codebook to reconstruct the give sample, but does not consider the class label available. These methods are called "unsupervised codebook learning". However, in most pattern recognition problems, the class labels of the training samples are available. Thus it is better to consider the class label to conduct the sparse coding and codebook learning, so that the generated sparse code and codebook can also benefit the classification problem. For example, in [17], the codebook is learned by using the class labels to maximize the margins of samples [55]–[59].

Inspired by the works of [17], in this paper, we propose a new codebook learning and sparse coding method via class label constrains based on K-SVD. To learn a discriminative and compact codebook for the coding task, we propose to constrain the codebook and sparse codes by class labels with the following two formulas:

- Firstly, we construct different codebooks for differen classes, and then construct an ideal sparse code according to this codebook. To learn the sparse codes, we hope that the sparse codes could be transformed to the ideal codes by a linear transformation. In this way, we constrain that the sparse codes of samples of different classes are different from each other, while the samples of the same class can have similar sparse code.
- Secondly, we also learn a discriminate classifier for each class for the sparse codes. The classifier is only used to classify the sparse code to one of the known classes, but also used to regularize the sparse code so that it could also be discriminative.

The learning of codebook and classifier constrained by the class labels are shown in Figure 2.

The proposed optimization problem is solved by the popular K-SVD algorithm. The outputs of the learning are not only a codebook, but also a classifier. The classifier learned is based on linear regression problem [60]–[64], so it has a very similar formula with sparse coding. Thus the learning of codebook and classifier can be done

Manuscript received August 10, 2011; revised January 2, 2012; accepted April 16, 2012.

doi:10.4304/jcp.9.4.836-844

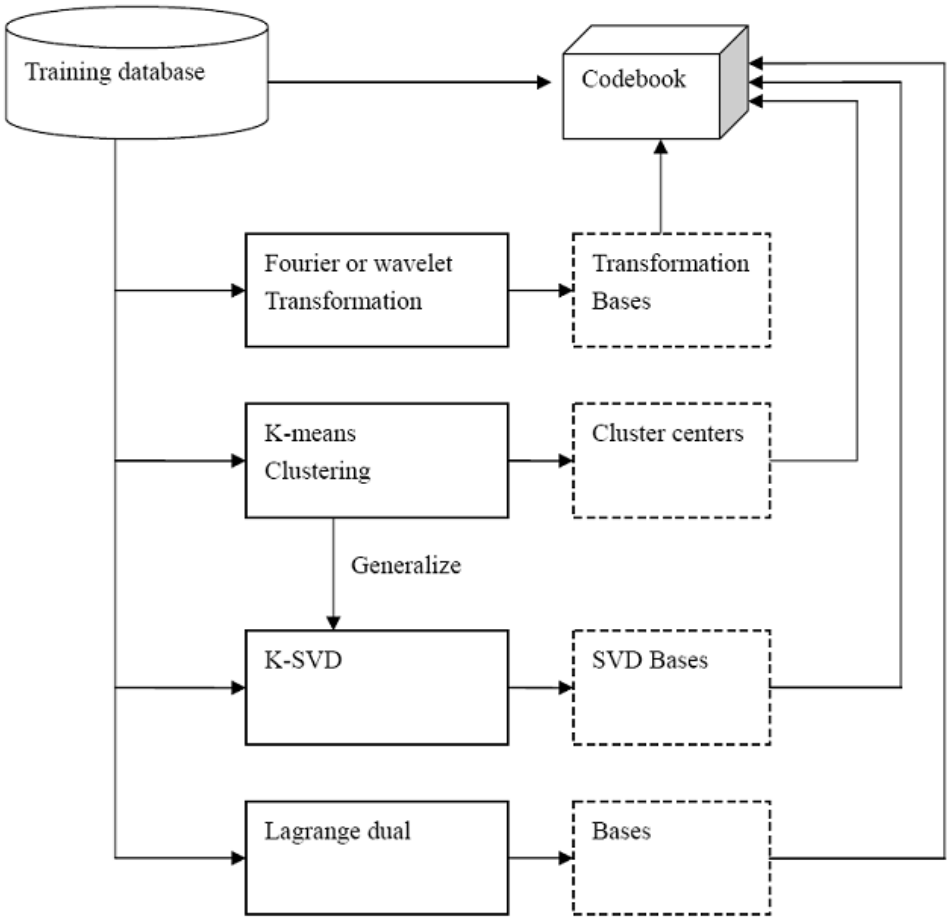

Figure 1. Codebook generation methods

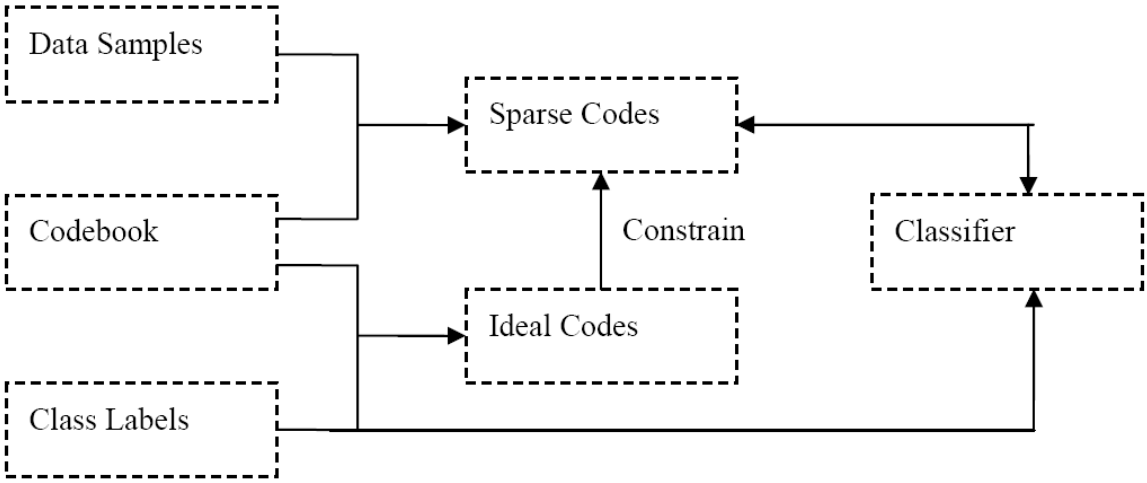

Figure 2. Learning of codebook and classifier constrained by the class labels

jointly and efficiently. This is different from traditional methods which first learn the sparse codes and then learn the classifier for them sequentially. In our method, the reconstruction error and classification error are considered and minimized at the same time. This paper has the following two main contributions:

1) We introduce some novel ideal sparse codes to constrain the learning of sparse codes. This ideal sparse code constrain term is combined with the reconstruction error term an d classification error term to construct a novel unified objective function for the learning problem.
2) To solve this objective function based problem, we propose to use the K-SVD algorithm. Thus the

codebook and classifier can be learned jointly at the same time.

This paper is organized as follows: Section II presents the novel method to learn the codebook and classifier. Section III gives the experimental results. Section IV concludes the paper and discusses future works.

II. Proposed Method

We assume that there are n training samples, denoted as a $d \times n$ data matrix $X = [\mathbf{x}_1, \cdots, \mathbf{x}_n] \in R^{d\times n}$, where d is the dimension of a sample vector. We want to learn a codebook $Y \in R^{d\times m}$ with m codewords to reconstruct the data matrix X, and we also want the reconstruction to be sparse. The reconstruction coefficient matrix is $Z = [\mathbf{z}_1, \cdots, \mathbf{z}_n] \in R^{m\times n}$, so the reconstructed matrix $\widetilde{X} = YZ$. The reconstruction error is denoted as

$$\|X - \widetilde{X}\|_2^2 = \|X - YZ\|_2^2 \tag{1}$$

Moreover, we also want the reconstruction coefficient to be as sparse as possible. So we constrain that the reconstruction coefficient vector $\mathbf{z}_i$ contain less than a non-zero items,

$$\|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \tag{2}$$

Thus the following problem is formulated for sparse coding problem,

$$\begin{aligned} &\min_{Y,Z} \|X - YZ\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \end{aligned} \tag{3}$$

To solve this problem, we use an iterative algorithm to solve Y and Z in turns:

- **solving** Y: When Y is considered, Z is fixed,

$$\min_{Y} \|X - YZ\|_2^2 \tag{4}$$

  The learning of Y is done by minimizing $\|X - YZ\|_2^2$ while constrain $\|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n$ at the same time. To solve this problem with regarding to the codebook matrix Y, we employ the K-SVD algorithm.
- **solving** Z: When Z is considered, Y is fixed,

$$\begin{aligned} &\min_{Z} \|X - YZ\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \end{aligned} \tag{5}$$

  This problem can be solved by Orthogonal Matching Pursuit (OMP) algorithm [65]–[68]. Because the vectors in Z are sparse, we also call them sparse codes.

A. Ideal Code Constraint

Suppose we are dealing the a K class multi-class classification problem [69]–[72]. We first initialize a codebook based on the class labels. That is, learning a sub-codebook for the k-th class $Y_k \in R^{d\times m_k}$, using only the training samples of this class, where m_k is the size of the sub-codebook of the k-th class. Then, the initial overall codebook is the combination of all these class-conditional sub-codebooks:

$$Y = [Y_1, \cdots, Y_K] \in R^{d\times m}, m = \sum_{k=1}^{K} m_k \tag{6}$$

Based on this codebook, we construct a idea sparse code $\mathbf{v}_i$ for the i-th sample belonging to the k-th class as

$$\begin{aligned} \mathbf{v}_i &= [v_{i1}, \cdots, v_{im}]^\top \\ &= [0, \cdots, \underbrace{1, \cdots, 1}_{k-th\ sub-codebook}, \cdots, 0]^\top \in \{0,1\}^m \end{aligned} \tag{7}$$

where v_{ij} is defined as

$$v_{ij} = \begin{cases} 1, & if\ \mathbf{y}_j \in Y_k \\ 0, & otherwise. \end{cases} \tag{8}$$

In this way, given a sample, only its sparse code elements corresponding to the codeword of its class-conational sub-codebook are non-zero, while all other ones are zero. For example, if we have three classes, and for each class we have two codewords, as is shown in Figure 3, then for a sample $\mathbf{x}_i$ belongs to class 2, only two elements of its ideal code corresponding to class 2 are non-zero. So the samples of the same class will have exactly the same ideal sparse code, and samples of different classes will have totally different ideal sparse codes. The ideal code matrix are organized as $V = [\mathbf{v}_1, \cdots, \mathbf{v}_n] \in R^{m\times n}$.

To make the learned sparse codes Z to be more discriminative, we hope it could be transformed to V. So we design a linear transformation to do this,

$$V \approx UZ \tag{9}$$

where $U \in R^{m\times m}$ is the transformation matrix to be solved. To learn this transformation matrix, we hope the transformation error could be as small as possible. The following optimization problem is argued:

$$\begin{aligned} &\min_{U,Z} \|V - UZ\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \end{aligned} \tag{10}$$

In this way, we could learn the sparse codes in Z so that the sparse codes of the same class will be similar, while the sparse codes of different classes will be dissimilar from each other.

B. Classifier Constraint

We also use the class labels to learn a classifier to predict the class label from sparse code. We denote the class label vector for the i-th sample as $\mathbf{f}_i = [f_{i1}, \cdots, f_{iK}]^\top \in \{1,0\}^K \in \{1,0\}^K$, where

$$f_{ik} = \begin{cases} 1, & if\ \mathbf{x}_i\ belongs\ to\ the\ k-th\ class; \\ 0, & otherwise. \end{cases} \tag{11}$$

The class label vectors are organized in a class label matrix $F = [\mathbf{f}_1, \cdots, \mathbf{f}_n] \in R^{K\times n}$. We hope to design

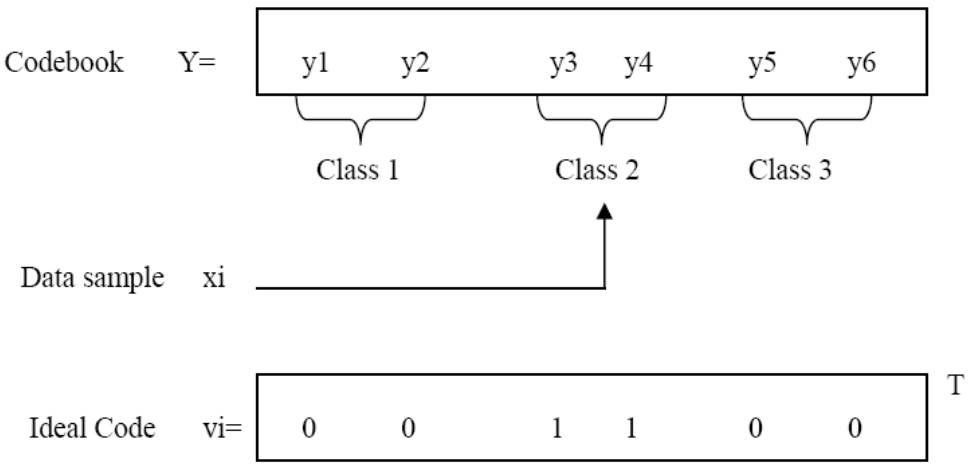

Figure 3. Ideal sparse code

a classifier function to predict the class label vector $\mathbf{f}_i$ of a sample from its sparse code $\mathbf{z}_i$, as

$$\mathbf{f}_i = W\mathbf{z}_i \tag{12}$$

where $W \in R^{K\times m}$ is the classifier parameter matrix. To learn it, we introduce the following optimization problem,

$$\begin{aligned} &\min_{W,Z} \|F - WZ\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \end{aligned} \tag{13}$$

C. Formula and Optimization

The optimization problem in this paper is formulated as follows:

$$\begin{aligned} &\min_{Y,U,W,Z} \|X - YZ\|_2^2 + \|V - UZ\|_2^2 + \|F - WZ\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \end{aligned} \tag{14}$$

In the objective function, the first term $\|X - YZ\|_2^2$ is the reconstruction error, the second term $\|V - VZ\|_2^2$ is the transformation error of the ideal codes, and the last term $\|F - WZ\|_2^2$ is the classification error. The objective function could be rewritten as

$$\begin{aligned} &\min_{Y,U,W,Z} \left\| \begin{bmatrix} X \\ V \\ F \end{bmatrix} - \begin{bmatrix} Y \\ U \\ W \end{bmatrix} Z \right\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \end{aligned} \tag{15}$$

In this formula, it could be seen that the samples, sparse codes, ideal codes and class labels are all involved. Their relations are shown in Figure 4.

We define two new matrix as

$$A = \begin{bmatrix} X \\ V \\ F \end{bmatrix} \in R^{(d+m+K)\times n}, B = \begin{bmatrix} Y \\ U \\ W \end{bmatrix} \in R^{(d+m+K)\times m}, \tag{16}$$

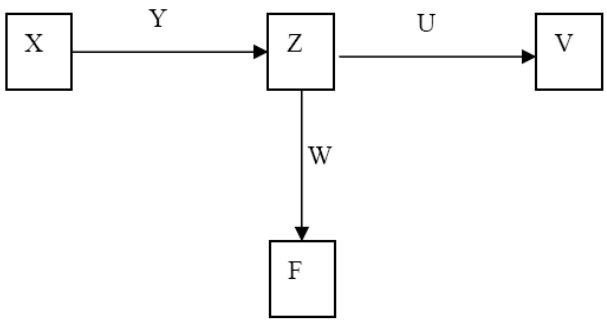

Figure 4. Relations among samples, sparse codes, ideal code, and class labels.

where A is the generalized data matrix, and B is the generalized codebook matrix. then we have

$$\begin{aligned} &\min_{B,Z} \|A - BZ\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \le a, i = 1, \cdots, n \end{aligned} \tag{17}$$

To solve this problem, we use the K-SVD algorithm to solve B, and the OMP algorithm to solve Z:

1) Solving B: To solve B, we have the following problem,

$$\min_{B} \|A - BZ\|_2^2 \tag{18}$$

To solve the generalized codebook $B = [\mathbf{b}_1, \cdots, \mathbf{b}_m]$, where $\mathbf{b}_j$ is the j-th generalized codeword, we solve the generalized codewords one by one. When $\mathbf{b}_j$ is considered, $\mathbf{b}_{j'}, j' \ne j$ are fixed. Similarly, we denote

the j-th row of Z as $\boldsymbol{\alpha}_j$, so that

$$Z = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_m \end{bmatrix} \tag{19}$$

In this way, we have

$$\begin{aligned} \|A - BZ\|_2^2 &= \left\| A - [\mathbf{b}_1, \cdots, \mathbf{b}_m] \times \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_m \end{bmatrix} \right\|_2^2 \\ &= \|A - \sum_{j=1}^{m} \mathbf{b}_j \boldsymbol{\alpha}_j\|_2^2 \\ &= \left\| \left(A - \sum_{j \neq j'}^{m} \mathbf{b}_{j'} \boldsymbol{\alpha}_{j'} \right) - \mathbf{b}_j \boldsymbol{\alpha}_j \right\|_2^2 \\ &= \left\| \widetilde{A}_j - \mathbf{b}_j \boldsymbol{\alpha}_j \right\|_2^2 \end{aligned} \tag{20}$$

where $\widetilde{A}_j = \left(A - \sum_{j \neq j'}^{m} \mathbf{b}_{j'} \boldsymbol{\alpha}_{j'} \right)$ when $\mathbf{b}_{j'}, j' \neq j$ are fixed, $\boldsymbol{\alpha}_{j'}, j' \neq j$ are also fixed. So we only try to solve $\mathbf{b}_j$ with regarding to $\boldsymbol{\alpha}_j$,

$$\min_{\mathbf{b}_j} \left\| \widetilde{A}_j - \mathbf{b}_j \boldsymbol{\alpha}_j \right\|_2^2 \tag{21}$$

because $\boldsymbol{\alpha}_j$ is sparse, and most elements in $\boldsymbol{\alpha}_j$ are zeros, we only consider the non-zeros elements in this vector. We introduce a new non-zero vector $\boldsymbol{\beta}_j$ which contains the non-zero elements. Similarly, we also introduce a new matrix $\widehat{A}_j$ which is a sub-matrix of $\widetilde{A}_j$, which only contains the columns corresponding to the non-zero elements in $\boldsymbol{\alpha}_j$, as shown in Figure 5.

In this way, (21) is turned to

$$\min_{\mathbf{b}_j} \left\| \widehat{A}_j - \mathbf{b}_j \boldsymbol{\beta}_j \right\|_2^2 \tag{22}$$

The optimal solution can be obtained by SVD decomposition [73]–[76]. We perform SVD to matrix $\widehat{A}_j$,

$$\widehat{A}_j = C \quad D^\top \tag{23}$$

Then we set the first column of C as $\mathbf{b}_j$.

2) Solving Z: By fixing B, to solve Z, we have

$$\begin{aligned} &\min_{Z} \|A - BZ\|_2^2 \\ &s.t. \|\mathbf{z}_i\|_0 \leq a, i = 1, \cdots, n \end{aligned} \tag{24}$$

It could be solved using OMP algorithm.

III. Experiments

In the experiments, we use several databases to evaluate the proposed algorithms:

- **YaleB database**: In this database, 2,414 face images are collected from 38 individuals. For each individual, about 64 face images are collected. The size of each image is of 192×168 pixels. In this database, the varying illumination condition and expression make the face recognition very difficult, as shown in Figure 6.
- **Caltech101 database**: This database contains 9144 images of 102 different object classes. These classes include animals, vehicles, flowers, etc. (as shown in Figure 7). The numbers of images for each class are from 31 to 800.

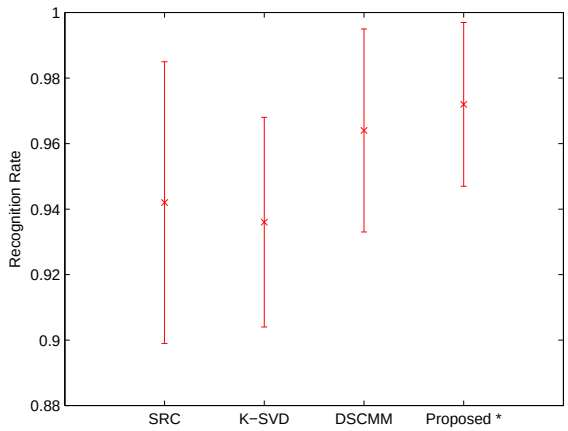

Figure 8. Experiment results on YaleB database.

We used the 10-fold cross validation protocol to conduct the experiment [77]–[79]. We also compare the proposed method against the following methods:

- sparse representation-based classification (SRC) [80], [81];
- K-SVD based sparse coding [35];
- Discriminative Sparse Coding on Multi-Manifolds (DSCMM) [7].

The experiment results are summarized in Figure 8 and 9. It is clear that the proposed approach always achieve better results than other methods. Please note that our approach uses the same codebook size as the others. A possible reason this outperforming is that we use the ideal code to constrain the sparse code, so that the class consistency can be mapped to the sparse code space.

IV. Conclusion

In this paper, we proposed a novel codebook learning method for sparse coding problem. The main contribution is the ideal codes constructed to constrain the sparse code. Moreover, we also use a classifier to regularize the sparse codes. To solve the problem, we formulate it as a K-SVD problem. Different from other methods which first learn the codebook, and then learn the classifier, our approach can learn them together. The encouraging experiment results show that the proposed algorithm outperforms all other methods.

References

[1] J. Wang, P. Liu, M. F.H.She, S. Nahavandi, and A. Kouzani, "Biomedical time series clustering based on

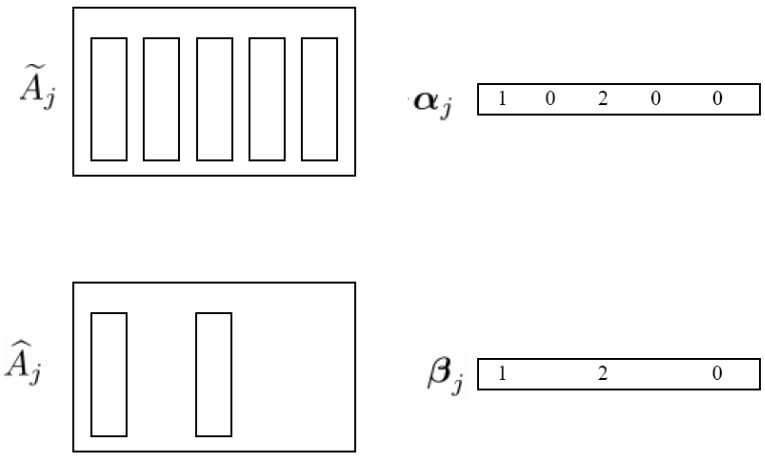

Figure 5. New code vector which only contain the non-zero elements, and it corresponding sub-matrix.

Figure 6. Example images of YaleB database

Figure 7. Example images of Caltech101 database

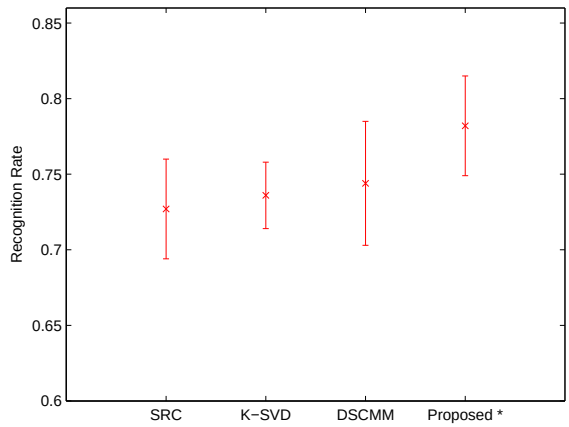

Figure 9. Experiment results on Caltech101 database.

non-negative sparse coding and probabilistic topic model," *Computer Methods and Programs in Biomedicine*, vol. 111, no. 3, pp. 629–641, 2013.

[2] S. Yang, Y. Lv, Y. Ren, and L. Jiao, "Superpixel-wise semi-supervised structural sparse coding classifier for image segmentation," *Engineering Applications of Artificial Intelligence*, 2013.

[3] S. Yang, H. Jin, M. Wang, Y. Ren, and L. Jiao, "Data-driven compressive sampling and learning sparse coding for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, 2013.

[4] Z. Ji, J. Theiler, R. Chartrand, G. Kenyon, and S. Brumby, "Decoupling sparse coding of sift descriptors for large-scale visual recognition," vol. 8750, 2013.

[5] H. Erdoğan, E. Erdem, and A. Erdem, "Group sparsity based sparse coding for region covariances," 2013.

[6] T. Aritake, H. Hino, and N. Murata, "Learning ancestral atom via sparse coding," *IEEE Journal on Selected Topics in Signal Processing*, vol. 7, no. 4, pp. 586–594, 2013.

[7] J. J.-Y. Wang, H. Bensmail, N. Yao, and X. Gao, "Discriminative sparse coding on multi-manifolds," *Knowledge-Based Systems*, 2013.

[8] C. Zhang, S. Wang, Q. Huang, C. Liang, J. Liu, and

Q. Tian, "Laplacian affine sparse coding with tilt and orientation consistency for image classification," *Journal of Visual Communication and Image Representation*, vol. 24, no. 7, pp. 786–793, 2013.

[9] J. Cui, J. Qi, and D. Li, "A fast approximate sparse coding networks and application to image denoising," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7951 LNCS, no. PART 1, pp. 620–626, 2013.

[10] T. Tong, R. Wolz, P. Coupé, J. Hajnal, and D. Rueckert, "Segmentation of mr images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling," *NeuroImage*, vol. 76, pp. 11–23, 2013.

[11] K. Huang, R. Hu, Z. Han, T. Lu, J. Jiang, and F. Wang, "Face image superresolution via locality preserving projection and sparse coding," *Journal of Software*, vol. 8, no. 8, pp. 2039–2046, 2013.

[12] G. Amato, F. Falchi, and C. Gennaro, "On reducing the number of visual words in the bag-of-features representation," vol. 1, 2013, pp. 657–662.

[13] K. Kashihara, "Classification of individually pleasant images based on neural networks with the bag of features," 2013, pp. 291–293.

[14] C. Wang, B. Zhang, Z. Qin, and J. Xiong, "Spatial weighting for bag-of-features based image retrieval," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8032 LNAI, pp. 91–100, 2013.

[15] Y. Xu, J. Liu, Z. Zhang, N. Tan, D. Wong, S. Saw, and T. Wong, "Learn to recognize pathological myopia in fundus images using bag-of-feature and sparse learning approach," 2013, pp. 888–891.

[16] Z. Wu, J. Cao, H. Tao, and Y. Zhuang, "A novel noise filter based on interesting pattern mining for bag-of-features images," *Expert Systems with Applications*, vol. 40, no. 18, pp. 7555–7561, 2013.

[17] J. J.-Y. Wang, H. Bensmail, and X. Gao, "Joint learning and weighting of visual vocabulary for bag-of-feature based tissue classification," *Pattern Recognition*, vol. 46, no. 12, pp. 3249–3255, 2013.

[18] J. Lin, L.-Y. Duan, J. Chen, R. Ji, S. Luo, and W. Gao, "Learning multiple codebooks for low bit rate mobile visual search," 2012, pp. 933–936.

[19] M. Jiu, C. Wolf, C. Garcia, and A. Baskurt, "Supervised learning and codebook optimization for bag-of-words models," *Cognitive Computation*, vol. 4, no. 4, pp. 409–419, 2012.

[20] Q. Li, C. Yao, L. Wang, and Z. Tu, "Randomness and sparsity induced codebook learning with application to cancer image classification," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7766 LNCS, pp. 181–193, 2013.

[21] X. Tian and Y. Lu, "Discriminative codebook learning for web image search," *Signal Processing*, vol. 93, no. 8, pp. 2284–2292, 2013.

[22] L. Cao, R. Ji, W. Liu, H. Yao, and Q. Tian, "Weakly supervised codebook learning by iterative label propagation with graph quantization," *Signal Processing*, vol. 93, no. 8, pp. 2274–2283, 2013.

[23] S. Beyhan, J.-M. Léger, and F. Kadirgan, "Understanding the influence of ni, co, rh and pd addition to ptsn/c catalyst for the oxidation of ethanol by in situ fourier transform infrared spectroscopy," *Applied Catalysis B: Environmental*, vol. 144, no. 1, pp. 66–74, 2014.

[24] J. Lima and L. Novaes, "Image encryption based on the fractional fourier transform over finite fields," *Signal Processing*, vol. 94, no. 1, pp. 521–530, 2014.

[25] E. Kurniawati, A. Rohman, and K. Triyana, "Analysis of lard in meatball broth using fourier transform infrared spectroscopy and chemometrics," *Meat Science*, vol. 96, no. 1, pp. 94–98, 2014.

[26] N. Le Bihan, S. Sangwine, and T. Ell, "Instantaneous frequency and amplitude of orthocomplex modulated signals based on quaternion fourier transform," *Signal Processing*, vol. 94, no. 1, pp. 308–318, 2014.

[27] S. Banerjee and M. Mitra, "Ecg beat classification based on discrete wavelet transformation and nearest neighbour classifier," *Journal of Medical Engineering and Technology*, vol. 37, no. 4, pp. 264–272, 2013.

[28] Y. Kim, "Rapid response calculation of lng cargo containment system under sloshing load using wavelet transformation," *International Journal of Naval Architecture and Ocean Engineering*, vol. 5, no. 2, pp. 227–245, 2013.

[29] M. Tarighat, M. Mohammadizadeh, and G. Abdi, "Simultaneous spectrophotometric determination of cd2+, cu 2+, and zn2+ in rice and vegetal samples with dimethylspiro[isobenzofurane-1,6-pyrorolo[2,3-d]pyrimidine]-2,3, 4,5(1 h,3 h,7 h)tetraone using wavelet transformation-feed forward neural networks," *Journal of Agricultural and Food Chemistry*, vol. 61, no. 28, pp. 6832–6840, 2013.

[30] M. Cyriac and C. Chellamuthu, "A wavelet based approach for near-lossless image compression using logarithmic transformation," *International Review on Computers and Software*, vol. 8, no. 6, pp. 1334–1340, 2013.

[31] F. Yang, T. Sun, and C. Zhang, "Erratum: An efficient hybrid data clustering method based on k-harmonic means and particle swarm optimization (expert systems with applications (2009) 36:6 (9847-9852))," *Expert Systems with Applications*, vol. 40, no. 11, p. 4735, 2013.

[32] K. Liao, G. Liu, L. Xiao, and C. Liu, "A sample-based hierarchical adaptive k-means clustering method for large-scale video retrieval," *Knowledge-Based Systems*, vol. 49, pp. 123–133, 2013.

[33] L. Bai, J. Liang, C. Sui, and C. Dang, "Fast global k-means clustering based on local geometrical information," *Information Sciences*, vol. 245, pp. 168–180, 2013.

[34] A. George, "Efficient high dimension data clustering using constraint-partitioning k-means algorithm," *International Arab Journal of Information Technology*, vol. 10, no. 6, 2013.

[35] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: Design of dictionaries for sparse representation," *Proceedings of SPARS*, vol. 5, pp. 9–12, 2005.

[36] Z. Haddad, A. Beghdadi, A. Serir, and A. Mokraoui, "Wave atoms based compression method for fingerprint images," *Pattern Recognition*, vol. 46, no. 9, pp. 2450–2464, 2013.

[37] K. Thangavelu and T. Krishnan, "Lossless color medical image compression using adaptive block-based encoding for human computed tomographic images," *International Journal of Imaging Systems and Technology*, vol. 23, no. 3, pp. 227–234, 2013.

[38] C. Rawat and S. Meher, "A hybrid image compression scheme using dct and fractal image compression," *International Arab Journal of Information Technology*, vol. 10, no. 6, 2013.

[39] E. Puthooran, R. Anand, and S. Mukherjee, "Lossless compression of medical images using a dual level dpcm with context adaptive switching neural network predictor," *International Journal of Computational Intelligence Systems*, vol. 6, no. 6, pp. 1082–1093, 2013.

[40] K. Vidhya and S. Shenbagadevi, "A novel threshold approach for medical image compression," *Advanced Science Letters*, vol. 19, no. 12, pp. 3713–3717, 2013.

[41] J.-X. Mi, "Face image recognition via collaborative representation on selected training samples," *Optik*, vol. 124, no. 18, pp. 3310–3313, 2013.

[42] F. Xiang, Z. Wang, and X. Yuan, "Image reconstruction based on sparse and redundant representation model: Local vs nonlocal," *Optik*, vol. 124, no. 18, pp. 3636–3641, 2013.

[43] N. Onkarappa and A. Sappa, "A novel space variant image representation," *Journal of Mathematical Imaging and Vision*, vol. 47, no. 1-2, pp. 48–59, 2013.

[44] W. Liu and S. Li, "Multi-morphology image super-resolution via sparse representation," *Neurocomputing*, vol. 120, pp. 645–654, 2013.

[45] P. Li, J. Cheng, and H. Lu, "Hashing with dual complementary projection learning for fast image retrieval," *Neurocomputing*, vol. 120, pp. 83–89, 2013.

[46] S. Jayasekara, H. Dassanayake, and A. Fernando, "A novel image retrieval system based on histogram factorization and contextual similarity learning," *Applied Mechanics and Materials*, vol. 380, pp. 4148–4151, 2013.

[47] J. Yu, Z. Qin, T. Wan, and X. Zhang, "Feature integration analysis of bag-of-features model for image retrieval," *Neurocomputing*, vol. 120, pp. 355–364, 2013.

[48] R. Das Gupta, J. Dash, and M. Sudipta, "Rotation invariant textural feature extraction for image retrieval using eigen value analysis of intensity gradients and multi-resolution analysis," *Pattern Recognition*, vol. 46, no. 12, pp. 3256–3267, 2013.

[49] H. Lee, A. Battle, R. Raina, and A. Ng, "Efficient sparse coding algorithms," 2007, pp. 801–808.

[50] S. Zhang, H. Yao, X. Sun, and X. Lu, "Sparse coding based visual tracking: Review and experimental comparison," *Pattern Recognition*, vol. 46, no. 7, pp. 1772–1788, 2013.

[51] B. He, D. Xu, R. Nian, M. van Heeswijk, Q. Yu, Y. Miche, and A. Lendasse, "Fast face recognition via sparse coding and extreme learning machine," *Cognitive Computation*, pp. 1–14, 2013.

[52] J. Xu, G. Yang, H. Man, and H. He, "L1 graph based on sparse coding for feature selection," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7951 LNCS, no. PART 1, pp. 594–601, 2013.

[53] K. Ramamurthy, J. Thiagarajan, P. Sattigeri, and A. Spanias, "Learning dictionaries with graph embedding constraints," 2012, pp. 1974–1978.

[54] H. Guo, Z. Jiang, and L. Davis, "Discriminative dictionary learning with pairwise constraints," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7724 LNCS, no. PART 1, pp. 328–342, 2013.

[55] X. Chen, J. Yang, D. Zhang, and J. Liang, "Complete large margin linear discriminant analysis using mathematical programming approach," *Pattern Recognition*, vol. 46, no. 6, pp. 1579–1594, 2013.

[56] H. Cevikalp and B. Triggs, "Hyperdisk based large margin classifier," *Pattern Recognition*, vol. 46, no. 6, pp. 1523–1531, 2013.

[57] H. Watanabe, T. Ohashi, S. Katagiri, M. Ohsaki, S. Matsuda, and H. Kashioka, "Robust and efficient pattern classification using large geometric margin minimum classification error training," *Journal of Signal Processing Systems*, pp. 1–14, 2013.

[58] M. Kirmse and U. Petersohn, "Large margin principle in hyperrectangle learning," *Neurocomputing*, 2013.

[59] B. Liu, B. Fang, X. Liu, J. Chen, Z. Huang, and X. He, "Large margin subspace learning for feature selection," *Pattern Recognition*, vol. 46, no. 10, pp. 2798–2806, 2013.

[60] G. Liang, J. Xu, and L. Liu, "Qspr analysis for melting point of fatty acids using genetic algorithm based multiple linear regression (ga-mlr)," *Fluid Phase Equilibria*, vol. 353, pp. 15–21, 2013.

[61] M. Černý, J. Antoch, and M. Hladík, "On the possibilistic approach to linear regression models involving uncertain, indeterminate or interval data," *Information Sciences*, vol. 244, pp. 26–47, 2013.

[62] A. Blanco-Fernández, A. Colubi, and M. García-Bárzana, "A set arithmetic-based linear regression model for modelling interval-valued responses through real-valued variables," *Information Sciences*, vol. 247, pp. 109–122, 2013.

[63] M. Aslam, T. Riaz, and S. Altaf, "Efficient estimation and robust inference of linear regression models in the presence of heteroscedastic errors and high leverage points," *Communications in Statistics: Simulation and Computation*, vol. 42, no. 10, pp. 2223–2238, 2013.

[64] Y.-D. Wu and X.-Q. Liu, "Non-diagonal-type estimator in linear regression," *Communications in Statistics: Simulation and Computation*, vol. 43, no. 1, pp. 142–150, 2014.

[65] Q. Zhao, J. Wang, Y. Han, and P. Han, "Compressive sensing of block-sparse signals recovery based on sparsity adaptive regularized orthogonal matching pursuit algorithm," 2012, pp. 1141–1144.

[66] G. Gui, A. Kuang, L. Wang, and A. Zhang, "Mip-mitigated sparse channel estimation using orthogonal matching pursuit algorithm," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 5, no. 1, pp. 180–186, 2013.

[67] M. Suchithra, P. Sukanya, P. Prabha, O. Sikha, V. Sowmya, and K. Soman, "An experimental study on application of orthogonal matching pursuit algorithm for image denoising," 2013, pp. 729–736.

[68] A. Fsheikh, M. Wheeler, and I. Hoteit, "An ensemble based nonlinear orthogonal matching pursuit algorithm for sparse history matching of reservoir models," vol. 1, 2013, pp. 84–95.

[69] K. Polat, "Data weighting method on the basis of binary encoded output to solve multi-class pattern classification problems," *Expert Systems with Applications*, vol. 40, no. 11, pp. 4637–4647, 2013.

[70] J. Abellán, "An application of non-parametric predictive inference on multi-class classification high-level-noise problems," *Expert Systems with Applications*, vol. 40, no. 11, pp. 4585–4592, 2013.

[71] H. Jabeen and A. Baig, "Two-stage learning for multi-class classification using genetic programming," *Neurocomputing*, vol. 116, pp. 311–316, 2013.

[72] Z. Wang and J. Feng, "Multi-class learning from class proportions," *Neurocomputing*, vol. 119, pp. 273–280, 2013.

[73] A. Sbrizzi, A. Raaijmakers, H. Hoogduin, J. Lagendijk, P. Luijten, and C. van den Berg, "Transmit and receive rf fields determination from a single low-tip-angle gradient-echo scan by scaling of svd data," *Magnetic Resonance in Medicine*, 2013.

[74] J. Wang, "Wavelet digital watermarking algorithm on the basis of svd decomposition," *Applied Mechanics and Materials*, vol. 333-335, pp. 1056–1059, 2013.

[75] G. Melo, A. Malcolm, D. Mikesell, and K. van Wijk, "Using svd for improved interferometric green's function retrieval," *Geophysical Journal International*, vol. 194, no. 3, pp. 1596–1612, 2013.

[76] M. Ali and C. Ahn, "An optimized watermarking technique based on self-adaptive de in dwt-svd transform domain," *Signal Processing*, vol. 94, pp. 545–556, 2014.

[77] I. Ashoor, N. Najafian, Y. Korin, E. Reed, T. Mohanakumar, D. Ikle, P. Heeger, and M. Lin, "Standardization and cross validation of alloreactive ifn elispot assays within the clinical trials in organ transplantation consortium," *American Journal of Transplantation*, vol. 13, no. 7, pp. 1871–1879, 2013.

[78] X. Zhang, X. Zhao, G. Liu, Q. Kang, and D. Wu, "Radioactive quality evaluation and cross validation of data from the hj-1a/b satellites' ccd sensors," *Sensors (Switzerland)*, vol. 13, no. 7, pp. 8564–8576, 2013.

[79] D. Rojatkar, K. Chinchkhede, and G. Sarate, "Handwritten devnagari consonants recognition using mlpnn with five fold cross validation," 2013, pp. 1222–1226.
[80] J. Wright, A. Ganesh, Z. Zhou, A. Wagner, and Y. Ma, "Demo: robust face recognition via sparse representation," 2008.
[81] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

Wenjing Liao received her B.E. degree in Computer Science from Guizhou University of Finance and Economics, China in June 2005 and her M.E. degree in Software Engineering from Sichuan University, China in December 2010. Her current research interest includes data mining technology and algorithm research and applied cryptography.

Robert Williams received his M.S degree from New Mexico Highlands University, US in June 2009. His current research interest includes multimedia technology and pattern recognition.