

# Low-Rank Optimization Dictionary Training for Image Classification

Ly Xuan<sup>1,2</sup>, Ma Zezhong<sup>1</sup>, and Liu Qing<sup>3,\*</sup>

<sup>1</sup>Chongqing Land Resources Housing Surveying and Planning Institute, 401121, Chongqing, China

<sup>2</sup>School of Remote Sensing and Information Engineering, Wuhan University, 430079, Wuhan, China

<sup>3</sup>Chongqing industrial& commercial school, 472289, Chongqing, China

**Abstract.** Bag-of-words model has been extremely popular in image categorization. The method of constructing the dictionary is important. In this paper a category constrained low-rank optimization dictionary training approach is proposed for the dictionary construction. Through the low-rank optimization, the rank of the coefficient matrix constructed by same category images is minimized. Experimental results show that the proposed method can obtain better performance on two standard image databases (Caltech-101 and Caltech-256) than not employing the category constrained low-rank optimization.

## 1 INTRODUCTION

If we do not have enough images of each category for training, it is not probable to use deep learning approaches. Bag of Words (BOW) [1] is a description method for statistic feature images which describes the histogram of an image via the frequencies of any feature words appearing in a statistic dictionary. Currently, there are numerous image feature description methods, such as those introduced in literature [2] that corresponds to different feature detection methods, feature description methods as well as feature organization and representation methods. Due to its powerful image feature description capacity, BOW is widely used in areas of image retrieval [3][4], classification [5] and coding [6][7]. BOW has different creation and usage methods in different applications. When it is used for image denoising and super-resolution image restructuring, an image is always divided into little blocks of the same size. In literature [8], feature analysis is performed to these local little blocks to form a feature BOW allowing local blocks description. For image classification, the model will perform feature analysis to the entire image, which normally consists of feature detection, feature description and feature clustering, and then a feature BOW will be formed by each cluster center. See literature [9-12].

In the issues of image classification based on BOW, it is important to know how to create a stable and high-resolution feature dictionary with high descriptive power. A good feature dictionary can maximize resolvability with minimum image restructuring errors. Normally there are two ways to construct a BOW: one is using images in all categories to construct a feature dictionary and the other is constructing a separate feature dictionary for each category of images so that multiple feature

dictionaries are constructed. For issues like a single dictionary having no sufficient utilization category information, and the very high calculation cost and easy loss of common feature information of each category of images useful for the classification of constructing a multiple-class dictionaries, particularly when there are a particularly large number of categories, this paper introduces a low rank optimization feature dictionary training approach with category constraint, which enables feature dictionaries to generate representation coefficients as similar as possible when images of the same category are represented by introduction of the idea of low rank, so as to introduce class information into dictionary learning and to increase the resolvability of dictionary description of images.

## 2 RELATED WORKS

Common-used dictionary methods for image classification are SPM [11] (Spatial Pyramid Matching) and BoF (Bag-of-Features) [13] and a number of expansions to BoF, such as local generation model described in literature [14] [15], potential generation model in literature [16] and discriminative dictionary learning in literature [17] [18]. Furthermore, BOW is also widely used in target monitoring and [19] and feature descriptor [20]. In these methods, feature dictionaries play a very important role. "Key feature dictionary" was first introduced in literature [9], which is also one of the first literatures where such method is used in image classification. In this literature, methods of affine covariant region and SIFT (Scale-Invariant Features Transform) feature descriptors were used for feature description, K-means was used to get a feature dictionary, and then Bayes classifier and support vector machine (SVM) were used to perform category

\* Corresponding author: shixuanlv305@126.com

experiments. Later, literature [10] made improvement to it, where standardized non-affine transformed SIFT locality description was used and SVM was replaced with Boost with additional minor geometric consistency constraints. In literature [12], feature BOW methods were studied in a more systematic way and experiments were performed to feature detection methods and feature descriptors, such as SIFT, RIFT (Rotation Invariant Features Transform) and SVM kernel function. In literature [21][22], sparse optimization was applied to BOW method; such a method could construct a dictionary with high resolvability, so it has become one of the mainstream BOW construction methods.

The procedure of describing images with feature BOW is: first, use the feature vectors in a dictionary to represent each feature vector detected in an image; then describe an image with the corresponding dictionary representation coefficients in the BOW of all the image features; and at last construct a classifier with the dictionary description feature vectors.

Key issues in a BOW requiring consideration:

- (1) Feature selection, including feature detection methods and feature description methods;
- (2) Methods of constructing a feature dictionary when features are extracted, such as conventional clustering method, sparse constraint method or KSVD [23][24] methods, etc.
- (3) How should the feature dictionary represent new features, that is to say how to quantify vectors, such as by means of nearest neighbor coding, flexible weighting and sparse coding;
- (4) How to pool after vectors have been quantified, that is how to construct the feature description of the entire image when the representation coefficients of all feature points in an image in the BOW;
- (5) Identify using feature vector construction classifier formed after feature dictionary description.

In the conventional dictionary training approach, literature [11] get the feature dictionary by clustering the features detected in all images (such as  $K$ -means), where the number of cluster centers are the number of words in the feature dictionary. If feature set  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M] \in R^{D \times M}$  is the feature points under  $M$   $D$ -dimension feature space and  $\mathbf{V}$  is the feature dictionary of  $K$  cluster centers, that is the size of the dictionary is  $K$ , then vectoring problem is to find the best optimization formula (1)

$$\min_{U, V} \sum_{m=1}^M \|\mathbf{x}_m - \mathbf{u}_m \mathbf{V}\|^2 \quad (1)$$

$$s.t. \text{Card}(\mathbf{u}_m) = 1, |\mathbf{u}_m| = 1, \mathbf{u}_m \geq 0, \forall m$$

where,  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_M]^T$ ,  $\text{Card}(\mathbf{u}_m) = 1$  indicate that only one atom of  $\mathbf{u}_m$  is nonzero, atoms of  $\mathbf{u}_m$  are nonnegative,  $|\mathbf{u}_m|$  is  $L1$  norm.

In the conventional approach, nearest neighbor feature words are used to represent new characteristics and then statistical histogram of words in dictionaries are calculated to represent images. This approach can reflect

the characteristic distribution in an image to an extent, but cannot accurately describe new characteristics as it only uses the nearest neighbor words to represent new characteristics, which have relatively large reconstruction error. Later, the approach based on flexible weighting value made improvement to the conventional approach, using the  $k$  atoms in the nearest dictionary to the feature vector of input signal, the weighting value of the  $k$  atoms is determined by the similarity between the atoms and the signal, and the more similar they are, the greater the weighting value. Although this approach is simple, easy to implement and has smaller signal reconstruction error comparing to  $K$ -means approach, such description method features relatively poor discrimination.

### 3 CATEGORY CONSTRAINED LOW-RANK OPTIMIZATION DICTIONARY TRAINING

#### 3.1 Basic Idea

Feature dictionary training is crucial, as a good feature dictionary will reduce the error in reconstruction of new features as much as possible while ensuring dictionaries for images of the same category have more similar dictionary representation, and dictionary for images of different categories have more distinctive differences. In formula (1), the category information of images which is helpful to train dictionaries with better resolvability is not employed when the feature dictionary is trained. The description of a dictionary to an image is represented on the description coefficient, that is the coefficient of each linear atom combination in the dictionary. Images of same category have similar coefficient matrix under the description of BOW, while images of different categories have relatively distinctive coefficient matrix under the description of BOW.

The low-rank idea is very widely used, such as for image correction in literature [25], for descriptor space structure in literature [26], for feature extraction in literature [27], for popular optimization in literature [28], for matrix restoration in literature [29], for subspace partitioning in literature [30], etc. During image correction, the rank for images corrected by transformation parameters will become lower because the corrected images will have more similar or same data in each column (or row), resulting in lower ranks of image matrix. Solution to such an issue can be transformed into solution to the following optimization problem:

$$\min_{I^0, E, \tau} \text{rank}(I^0) + \gamma \|E\|_0 \quad (2)$$

$$s.t. \quad I \circ \tau = I^0 + E$$

The part of rank solution in formula (2) is unfavorable to solve this optimization problem, and will normally be replaced with kernel norm solution:

$$\min_{I^0, E, \tau} \|I^0\|_* + \gamma \|E\|_0 \quad (3)$$

s.t.  $I \circ \tau = I^0 + E$

In view of the idea of low rank, the present paper suggest to apply low-rank optimization to dictionary creation, that is introducing category constraint information through minimizing the rank of coefficient matrix of images of the same category in the dictionary description so that images of the same category have more similar dictionary representation. For this purpose, the following optimization issues are constructed to solve low-rank constrained feature dictionary  $D$ :

$$\min \sum_{i=1}^C (\|Z_i\|_* + \lambda \|E_i\|_{2,1})$$

s.t.  $X_i = DZ_i + E_i \quad i = 1, 2, \dots, C$  (4)

$$\sum_{m=1}^p d_{mm}^2 \leq 1 \quad n = 1, 2, \dots, k$$

where,  $\|Z_i\|_*$  is the nuclear norm, the sum of the singular values,  $\|E\|_{2,1} = \sum_j \sqrt{\sum_i [E]_{ij}^2}$  is  $l_2/l_1$  norm, which is some kind of regularization strategy. There are  $C$  classes and  $Z_i$  is the coefficient matrix of the class  $i$  images using  $D$  as the dictionary.

### 3.2 Solving the Low-Rank Optimization Problem

The optimization problem in (4) can be solved by ALM (Augmented Lagrange Method). Let  $Z_i = J_i$ , The above formulation (4) naturally leads to the following optimization problem:

$$\min \sum_{i=1}^C \left( \|J_i\|_* + \lambda \|E_i\|_{2,1} + tr\langle W_i, Z_i - J_i \rangle + tr\langle Y_i, X_i - DZ_i - E_i \rangle \right. \\ \left. + \frac{\mu}{2} \|X_i - DZ_i - E_i\|_F^2 + \frac{\mu}{2} \|Z_i - J_i\|_F^2 \right) \\ + tr(B(D^T D - I)). \quad (5)$$

Where  $B$  is a diagonal matrix. Update  $J_i$  :

$$\|J_i\|_* + \frac{\mu}{2} \left( \frac{2}{\mu} tr\langle W_i, Z_i - J_i \rangle + \|Z_i - J_i\|_F^2 + \frac{\|W_i\|_F^2}{\mu^2} - \frac{\|W_i\|_F^2}{\mu^2} \right) \\ \Rightarrow \|J_i\|_* + \frac{\mu}{2} \left\| J_i - \left( Z_i + \frac{W_i}{\mu} \right) \right\|_F^2 \quad (6)$$

Thus, we end up with the following optimization problem:

$$J_i^* = \arg \min \frac{1}{\mu} \|J_i\|_* + \frac{1}{2} \left\| J_i - \left( Z_i + \frac{W_i}{\mu} \right) \right\|_F^2 \quad (7)$$

A singular value thresholding algorithm can be used to solve (7). It is defined that

$$S_\varepsilon[x] \doteq \begin{cases} x - \varepsilon, & \text{if } x > \varepsilon \\ x + \varepsilon, & \text{if } x < -\varepsilon \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

with  $x \in R$  and  $\varepsilon > 0$ , for the above problem in(7), we end up with

$$US_\varepsilon[S]V^T = \arg \min_X \varepsilon \|X\|_* + \frac{1}{2} \|X - W\|_F^2, \quad (9)$$

where  $USV^T$  is the SVD of  $W$ . Update  $Z_i$  :

$$tr\langle W_i, Z_i - J_i \rangle + tr\langle Y_i, X_i - DZ_i - E_i \rangle + \frac{\mu}{2} \|X_i - DZ_i - E_i\|_F^2 \\ + \frac{\mu}{2} \|Z_i - J_i\|_F^2 \\ \Rightarrow \frac{\mu}{2} \left( \frac{2}{\mu} tr\langle W_i, Z_i - J_i \rangle + \|Z_i - J_i\|_F^2 + \frac{\|W_i\|_F^2}{\mu^2} - \frac{\|W_i\|_F^2}{\mu^2} \right) \\ + \frac{\mu}{2} \left( \frac{2}{\mu} tr\langle Y_i, X_i - DZ_i - E_i \rangle + \|X_i - DZ_i - E_i\|_F^2 \right) \\ + \frac{\mu}{2} \left( \frac{\|Y_i\|_F^2}{\mu^2} - \frac{\|Y_i\|_F^2}{\mu^2} \right) \\ \Rightarrow \frac{\mu}{2} \left( \left\| Z_i - J_i + \frac{W}{\mu} \right\|_F^2 + \left\| X_i - DZ_i - E_i + \frac{Y_i}{\mu} \right\|_F^2 \right) \quad (10)$$

The formulation (10) derivatives at  $Z_i$

$$\frac{\delta \frac{\mu}{2} \left( \left\| Z_i - J_i + \frac{W}{\mu} \right\|_F^2 + \left\| X_i - DZ_i - E_i + \frac{Y_i}{\mu} \right\|_F^2 \right)}{\delta Z_i} = 0. \quad (11)$$

We get

$$Z_i = (I + D^T D)^{-1} \left( J_i + D^T (X_i - E_i) + \frac{D^T Y_i - W_i}{\mu} \right) \quad (12)$$

Update  $E_i$  :

$$\lambda \|E_i\|_{2,1} + \frac{\mu}{2} \left\| X_i - DZ_i - E_i + \frac{Y_i}{\mu} \right\|_F^2 + tr\langle Y_i, X_i - DZ_i - E_i \rangle \\ \Rightarrow \lambda \|E_i\|_{2,1} + \frac{\mu}{2} \left\| E_i - \left( X_i - DZ_i + \frac{Y_i}{\mu} \right) \right\|_F^2 \quad (13)$$

$$E_i^* = \arg \min \frac{\lambda}{\mu} \|E_i\|_{2,1} + \frac{1}{2} \left\| E_i - \left( X_i - DZ_i + \frac{Y_i}{\mu} \right) \right\|_F^2 \quad (14)$$

We can solve the optimization problem(14)through the approach in [31], with  $Q = [q_1, q_2, \dots, q_i, \dots]$ , if  $W^*$  is the optimal solution:

$$\min_W \lambda \|W\|_{2,1} + \frac{1}{2} \|W - Q\|_F^2 \quad (15)$$

Then the column vector of  $W^*$  can be calculated as:

$$[W^*]_{:,i} = \begin{cases} \frac{\|q_i\|_2 - \lambda}{\|q_i\|_2} q_i, & \text{if } \lambda < \|q_i\|_2, \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

Update  $D$ :

$$\begin{aligned} & \sum_{i=1}^C \left( \text{tr} \langle \mathbf{Y}_i, \mathbf{X}_i - \mathbf{DZ}_i - \mathbf{E}_i \rangle + \frac{\mu}{2} \|\mathbf{X}_i - \mathbf{DZ}_i - \mathbf{E}_i\|_F^2 \right) + \text{tr}(\mathbf{B}(\mathbf{D}^T \mathbf{D} - \mathbf{I})) \\ \Rightarrow & \sum_{i=1}^C \frac{\mu}{2} \left( \left\| \mathbf{DZ}_i - \mathbf{X}_i + \mathbf{E}_i - \frac{\mathbf{Y}_i}{\mu} \right\|_F^2 \right) + \text{tr}(\mathbf{B}(\mathbf{D}^T \mathbf{D} - \mathbf{I})) \end{aligned} \quad (17)$$

The formulation (17) derivatives at  $\mathbf{D}$ :

$$\frac{\delta \sum_{i=1}^C \frac{\mu}{2} \left( \left\| \mathbf{DZ}_i - \mathbf{X}_i + \mathbf{E}_i - \frac{\mathbf{Y}_i}{\mu} \right\|_F^2 \right) + \text{tr}(\mathbf{B}(\mathbf{D}^T \mathbf{D} - \mathbf{I}))}{\delta \mathbf{D}} = 0 \quad (18)$$

$$\mathbf{D} = \sum_{i=1}^C \left( \mathbf{X}_i - \mathbf{E}_i + \frac{\mathbf{Y}_i}{\mu} \right) \mathbf{Z}_i^T \left( \sum_{i=1}^C \mu \mathbf{Z}_i \mathbf{Z}_i^T + 2\mathbf{B} \right)^{-1} \quad (19)$$

Update the multipliers iteratively until it is converged:

$$\begin{aligned} \mathbf{W}_i &= \mathbf{W}_i + \mu(\mathbf{Z}_i - \mathbf{J}_i) \\ \mathbf{Y}_i &= \mathbf{Y}_i + \mu(\mathbf{X}_i - \mathbf{DZ}_i - \mathbf{E}_i) \\ \mathbf{B} &= \mathbf{B} + \mu(\mathbf{D}^T \mathbf{D} - \mathbf{I}) \end{aligned} \quad (20)$$

## 4 EXPERIMENT AND ANALYSIS

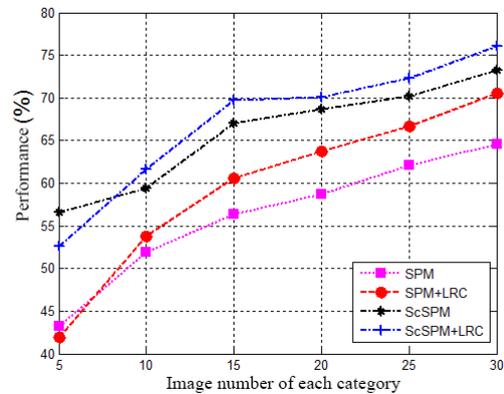
To verify the effectiveness of the proposed image description approach, this paper selects Caltech-101[32] and Caltech-256[33] image database for experiment. First, transform images into grey-scale images, and then detect the key feature points by means of DOF, describe the key points through SIFT feature descriptors where the  $4 \times 4$  total seed points obtained via local pixel locality block probability density sampling are used for description. Each key point generates 128 data with moving step length of 8 pixels. Dictionary initialization is obtained by means of K-means, and the size of a feature dictionary is  $128 \times 1024$ . During the experiment, all the classifiers are LIBSVM support vector classifiers.

Caltech-101 image database contains 101 categories of images for animals, flowers and cars and a total 9144 pieces of images. Each category includes 31 to 800 images and most images are pixel images. For the experiment, some images from each category are selected respectively to form an exercise set while the remaining images form a testing set. Table 1 shows the comparison between results of the experiment with Caltech-101 by this and other approaches. Wherein LRC-DT (Low Rank Constrained-Dictionary Training) is the approach proposed in this paper with the feature dictionary obtained from low rank constrained optimization. From the results, it can be seen that the proposed LRC-DT dictionary training approach effectively increases the accuracy of image classification, and with the size increase of exercise sample of each category, addition of low-rank constraint (LRC) can better adjust the base vector of the feature dictionary, realizing better classification effect; however, small exercise sample of each category will result in relatively smaller increase or smaller decrease. To evaluate the experiment effectiveness before and after the low-rank

optimization feature dictionaries are used, the SPM approach proposed by Lazebnik et al. [11] and ScSPM coding approach in literature [21] are replaced with the feature dictionary with LRC low-rank constraints and compare the classification effect. Figure 1 shows the experiment results, from which we can see that when the calculation methods in other steps are kept unchanged, replacing feature dictionary training approach with low-rank constrained approach can improve the classification accuracy by various extents.

**Table 1.** Image classification accuracy on Caltech-101 (%)

Image number of each category	5	10	15	20	25	30
Zhang[34]	46.6	55.8	59.1	62.0	-	66.2
Lazebnik [11]	43.2	51.9	56.4	58.7	62.1	64.6
Yang[21]	<b>56.5</b>	59.4	67.0	68.7	70.1	73.2
Wang[22]	51.1	59.7	65.4	67.7	70.1	73.4
LRC-DT	52.6	<b>61.6</b>	<b>68.7</b>	<b>70.0</b>	<b>72.3</b>	<b>76.1</b>

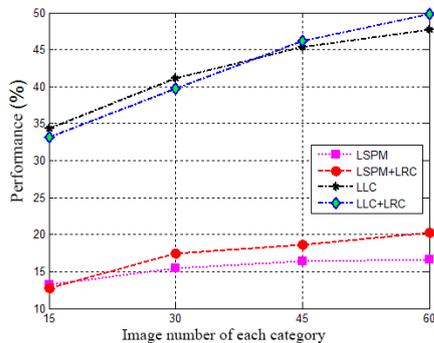


**Figure 1.** Comparison Results on Caltech-101

Caltech-256 image database has 256 image categories with 30607 images. Comparing to Caltech-101 image database, it has more category and image variance, and each category includes at least 80 images. Table 2 shows the experiment results that when there are 45 and 60 exercise samples, which means there are relatively a large number of exercise samples in each category, the approach proposed in this paper has the highest classification accuracy. Figure 2 shows the comparison results of SPM approach and LLC coding approach with the feature dictionary replaced with the low-rank optimization constrained feature dictionary. It can be seen that with LRC, the accuracy has increased to a certain extent.

**Table 2.** Image classification accuracy on Caltech-256 (%)

Image number of each category	15	30	45	60
Griffin[35]	28.30	34.10	-	-
Yang[21]	27.73	34.02	37.46	40.14
Wang[22]	<b>34.36</b>	<b>41.19</b>	45.31	47.68
LRC-DT	33.15	39.76	<b>46.17</b>	<b>49.78</b>



**Figure 2.** Comparison Results on Caltech-101

## 5 CONCLUSION

Constructing a multiple BOW with rich representation is crucial and a good feature BOW can not only represent abundant image features, realize better resolution to descriptions of images of different categories, but also give descriptions with similarity to images of the same category. The approach to apply low-rank constraint to the coefficient matrix of images of the same category integrates category information into the creation process of feature dictionaries, improving performance of such feature dictionaries and increase the accuracy of image classification based on BOW to a certain extent. It is probable for the situation that we do not have enough images of each category for training.

The approach proposed in this paper integrates the similarity of feature description of images of the same category into the construction of feature dictionaries, but how to include the relatively great difference between feature descriptions of images of different categories will be further researched in subsequent studies. Furthermore, due to time consumption of low-rank optimization calculation, the efficient calculation of low-rank optimization will be another direction of study in the future.

## ACKNOWLEDGMENTS

This work was supported by Image Recognition Based High Resolution Remote Sensing Data Analysis Research, CQGT-KJ-2014032.

## References

- Jegou H, Douze M, Schmid C. Packing bag-of-features[C]. // In IEEE 12th International Conference on Computer Vision (ICCV), Piscataway: IEEE, 2009, 2357-2364.

- Peyre G. A Review of Adaptive Image Representations[J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 896-911.
- Philbin J, Chum O, Isard M, et al. Object retrieval with large vocabularies and fast spatial matching[C]. // In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway: IEEE, 2007, 1-8.
- Chum O, Philbin J, Sivic J, et al. Total recall: Automatic query expansion with a generative feature model for object retrieval[C]. //In IEEE 11th International Conference on Computer Vision (ICCV), Piscataway: IEEE, 2007: 1-8.
- Li F F, Perona P. A Bayesian hierarchical model for learning natural scene categories[C]. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway: IEEE, 2005, 2: 524-531.
- Deng C Z, Cao H Q. Construction of Multiscale Ridgelet Dictionary and Its Application for Image Coding[J]. Journal of Image and Graphics, 2009,14(7):1273-1278.
- Sun Y B, Wei Z H, Xiao L, et al. Multimorphology Sparsity Regularized Image Super-Resolution. ACTA ELECTRONICA SINICA, 2010, 38(12):2898-2903.
- Aharon M, Elad M, Bruckstein A. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311-4322.
- Csurka G, Dance C R., Fan L, et al. Visual categorization with bags of keypoints[C]. // In ECCV International Workshop on Statistical Learning in Computer Vision, Berlin: Springer, 2004: 1-22.
- Csurka G, Dance C R, Perronnin F, et al. Generic visual categorization using weak geometry[J]. Toward Category-Level Object Recognition Lecture Notes in Computer Science, 2006: 207-224.
- Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories[C]. // In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway: IEEE, 2006: 2169-2176.
- Zhang J, Marszalek M, Lazebnik S, et al. Local features and kernels for classification of texture and object categories: a comprehensive study[J]. International Journal of Computer Vision, 2007, 73(2): 213-238.
- Sivic J, Zisserman A. Video google: A text retrieval approach to object matching in videos[C]. // In IEEE International Conference on Computer Vision (ICCV), Piscataway: IEEE, 2003: 1470-1477.
- Boiman O, Shechtman E, Irani M. In defense of nearest-neighbor based image classification[C]. // In IEEE Computer Society Conference on Computer

- Vision and Pattern Recognition (CVPR), Piscataway: IEEE , 2008: 1-8.
15. Li F F, Fergus R, PERONA P. Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories[J]. *Computer Vision and Image Understanding*, 2007, 106(1): 59-70.
  16. Bosch A, Zisserman A, Munoz X. Scene classification using a hybrid generative/discriminative approach[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2008, 30(4): 712-727.
  17. Liu Y, Jin R., Sukthankar R., et al. Unifying discriminative visual codebook generation with classifier training for object category recognition[C]. // In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Piscataway: IEEE , 2008: 1-8.
  18. Jurie F, Triggs B. Creating efficient codebooks for visual recognition[C]. // In *IEEE International Conference on Computer Vision (ICCV)*, Piscataway: IEEE , 2005, 1: 604-610.
  19. Burghouts, G J, Schutte K. Spatio-temporal layout of human actions for improved bag-of-words action detection[J]. *Pattern Recognition Letters*, 2013, 34(15): 1861-1869.
  20. Banerji, S, Sinha A, et al.. A New Bag of Words LBP (BoWL) Descriptor for Scene Image Classification[J]. *Computer Analysis of Images and Patterns*, 2013, 8047: 490-497.
  21. Yang J C, Yu K, Gong Y H, et al. Linear spatial pyramid matching using sparse coding for image classification[C]. // In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Piscataway: IEEE , 2009: 1794-1801.
  22. Wang J J, Yang J C, Yu K, et al. Locality-constrained linear coding for image classification[C]. // In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Piscataway: IEEE , 2010: 3360-3367.
  23. Rubinstein, R, Peleg T, et al.. Analysis K-SVD: A Dictionary-Learning Algorithm for the Analysis Sparse Model[J]. *IEEE Transactions on Signal Processing*, 2013, 61(3): 661-677.
  24. Jiang Z, Lin Z, et al. Label Consistent K-SVD: Learning A Discriminative Dictionary for Recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, (99): 1-13.
  25. Zhang Z, Ganesh A, et al. TILT: Transform Invariant Low-Rank Textures[J]. *International Journal of Computer Vision*, 2012, 99(1): 1-24.
  26. Liu G C, lin Z C, et al. Robust Recovery of Subspace Structures by Low-Rank Representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 171-184.
  27. Zhang N, Yang J. Low-rank representation based discriminative projection for robust feature extraction[J]. *Neurocomputing*, 2013, 111(0): 13-20.
  28. Shalit U, Weinshall D, et al.. Online learning in the embedded manifold of low-rank matrices[J]. *Journal of Machine Learning Research*, 2012, 13(1): 429-458.
  29. Liu Y Y, Jiao L C, et al. An efficient matrix bi-factorization alternative optimization method for low-rank matrix recovery and completion[J]. *Neural Networks*, 2013, 48(0): 8-18.
  30. Zhang X, Sun F, et al.. Fast Low-Rank Subspace Segmentation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2013, (99): 1-6.
  31. Yang J F, Yin W T, Zhang Y, et al. A fast algorithm for edge preserving variational multichannel image restoration[J]. *SIAM Journal on Imaging Sciences*, 2009, 2(2): 569-592.
  32. Boiman O, Shechtman E, Irani M. In defense of nearest-neighbor based image classification[C]. // In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Piscataway: IEEE , 2008: 1-8.
  33. Shotton J, Winn J, Rother C, et al. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling appearance, shape and context[J]. *International Journal of Computer Vision*, 2009, 81(1): 2-23.
  34. Zhang H, Berg A C, MAIRE M, MALIK J. Svm-knn: Discriminative nearest neighbor classification for visual category recognition[C]. //In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Piscataway: IEEE, 2006: 2126-2136.
  35. Griffin G, Holub A, Perona P. Caltech-256 Object Category Dataset. Technical Report 7694, California Institute of Technology, 2007.