Online Codebook Reweighting Using Pairwise Constraints for Image Classification

Xin Zhao^{1,2}, Weiqiang Ren², Kaiqi Huang², and Tieniu Tan²

1. Department of Automation, University of Science and Technology of China, Hefei, China

2. National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China

Email: {xzhaoh, wqren, kqhuang, tnt}@nlpr.ia.ac.cn

Abstract—Bag-of-words (BoW) model is widely used for image classification. Recently, the framework of sparse coding and max pooling proved an effective approach for image classification. Max pooling adopts a winner-take-all strategy. Thus, it can be regarded as a codebook weighting process. The results of this process are the weights of the associated codebook. However, there are high intra-class variations and strong background clutters in many image classification tasks. The weights obtained by max pooling only have limited information. This paper presents a codebook reweighting algorithm using pairwise constraints to improve the performance of sparse coding and max pooling framework. Pairwise constraints are the natural way of encoding the relationships between pairs of images. Therefore, the reweighted codebook is more effective to describe the relevance between pairs of images. An efficient online learning algorithm is presented based on passive-aggressive training strategy. We compare our method with other state-of-the-art methods on Graz-01 & 02 datasets. Experimental results illustrate the effectiveness and efficiency of our method for image classification.

Index Terms—Image classification, Codebook reweighting, Online learning, Pairwise constraints

I. INTRODUCTION

Image classification is a major problem in computer vision and pattern recognition. Image representation is important because it is the fundamental step in image classification task. In order to get a good representation, bag-of-words model [1]–[4] has recently gained a lot of interest in the community. This model is simple and effective. First, the classic BoW model extracts local feature descriptors of an image. Then, the descriptors are voted to the codebook which is obtained by clustering. At last, the voting results are represented as a histogram. The used voting strategy is hard voting [1]. This model still has two shortcomings: (1) hard voting strategy only considers the occurrence frequency of codes and it is prone to noise of background feature descriptors; (2) it does not take spatial information into account. This representation has limited distinction.

To overcome the problems of the classic BoW model, the spatial pyramid matching (SPM) [5] has been proposed to combine the spatial information with the original representation. This method can be considered as a simple image region partition strategy. It achieved an improvement over an orderless image representation. In [6], Yang et al. have proposed a method (ScSPM) which combined sparse coding with max pooling to overcome the limitations of hard voting strategy and improve the SPM. In the left part of Fig. 1, firstly,

Fig. 1. (Left) The framework of the image representation process with the proposed codebook reweighting method. (Right) A demonstration of the our motivation. The rectangular points and the triangle points are data points in two different classes. The red lines and green lines stand for cannot-link and must-link respectively. After codebook reweighting, the yellow data points will be closer to the data points with the same category.

this method generates sparse coding by a reconstruction error minimum process. Unlike normal vector quantization (VQ) method, in this step, a codebook is generated by the alternating training process as well. After sparse coding, a max pooling strategy is used to improve the original SPM by extracting the max response of each codeword in a SPM region. Although the learned representation by [6] is more salient and robust to local translation, the codewords of the induced representation do not have informative weights. This might be a problem when there are high intra-class variations and strong background clutters. On one hand, images in the same class may be quite different, since there are background clutters. On the other hand, images in different classes may have similar feature representation because the objects of interest are not the major things in the images.

The shortcomings of existing methods motivate us to get more informative representation for images by evaluating codewords with appropriate weights. In this paper, we propose a codebook reweighting method to give more informative weights to the codebook in ScSPM by incorporating constraints between pairs of images. The pairwise constraints have two common forms, namely must-link and cannot-link. They indicate whether two images belong to the same category or not. Therefore, they are considered as a natural way to simulate the human senses of similarity between pairs of images. Our reweighting strategy has an intuitive goal: the

must-link images have larger similarity than the cannot-link images after codebook reweighting, as shown in the right part of Fig. 1. This is also a common idea that has been captured by other works [7]–[10] especially the distance metric learning methods [11].

The most obvious drawback of using pairwise constraints is the prohibitive computational complexity and storage complexity. Due to the consideration of pairwise similarity, there are a large number of constraints in the codebook reweighting task even for a small dataset. In order to reduce the complexity, we propose an online learning algorithm to do codebook reweighting based on passive-aggressive [13], [14] learning strategy. First, we construct the pairwise constraints using a similarity measure which is similar to intersection kernel in [15]. Then, we combine the pairwise constraints with the hinge loss function and minimize it in a large margin manner. The whole learning process is conducted by adding each constraint sequentially. Compared to the batch learning algorithm such as [9], the storage complexity of our online learning algorithm is much lower. The computational complexity is much lower as well because a closed-form solution is captured without using any optimization tools. The proposed algorithm is an efficient learning method while maintaining a good performance for image classification task.

In this paper, we make the following contributions:

- analyze the BoW model based methods especially the ScSPM framework. Due to their shortcomings, we propose to reweight the codebook generated by ScSPM.
- utilize pairwise constraints to encode the relationships between pairs of images in order to give codewords more informative weights.
- formulate the pairwise constraints into an online learning framework. The proposed online learning algorithm is much faster than the related batch learning algorithm.

The remainder of the paper is organized follows. In Section II we introduce the methods that are related to our work. We formulate the codebook reweighting with pairwise constraints in Section III. Then in Section IV, we give the details of the proposed online learning algorithm. Experimental results and analysis are presented in Section V. Finally, we draw conclusions in Section VI.

II. RELATED WORK

In this section, we will briefly review the existing methods that are related to our method. Due to BoW model's simplicity and good performance on image classification problem, many researchers focused their attention on this model and did some work to improve the classic BoW model.

The primary step of BoW model is codebook generation. There is much work on this problem. The widely used codebook generation method is the k-means clustering method. In [16], [17], radius based clustering method is introduced to get more informative codebook than normal k-means and its variants. Another direction of the BoW model improvement is the descriptor coding strategy such as [6], [18]–[20]. In [6], a sparse coding strategy and a max pooling strategy

are combined in order to get state-of-the-art performance on several benchmark datasets. In the training process of ScSPM, a codebook is generated by minimizing a reconstruction error. Thus, this induced codebook is more informative than the codebook in other methods [16]–[20]. Our method is based on ScSPM. We try to reweight the codebook that is generated by ScSPM using the results of max pooling step.

Compared with hard-voting strategy, a soft-voting strategy [17] has been proposed to represent each descriptor with multiple codewords. Instead of the occurrence frequency of each codeword, the responses of codewords are replaced by the distance between descriptors and codewords. As [6] and [17] generating more informative codebook than other methods, we can consider them as weighted codebook methods. In order to further improve the BoW model, codebook reweighting is proposed to deal with the shortcomings of weak weighted codebook methods. Similar reweighting motivation can also be found in [22]. Inspired by [9], [21], we combine pairwise constraints with a large margin training process to do codebook reweighting. Pairwise constraints are considered as the natural way to simulate the human senses of similarity between pairs of images. Thus, pairwise constraints are utilized to get more informative codebook weights for image classification.

We also propose an online learning algorithm based on passive-aggressive strategy for efficient codebook reweighting using pairwise constraints. There are some related methods that have been proposed for batch learning. The intuitive strategy is to choose some of the constraints such as in [21]. However, the strategy need to be designed carefully, since some reasonable constraints might be discarded. In [9], Cai et al. proposed a method that combined a constraint selection strategy with an alternating optimization process. Although the alternating optimization was much faster than the global optimization in [9], it still needs to solve a quadratic problem with the help of some optimization tools to get the weighting results. As shown in Section IV, the proposed online learning algorithm has low complexity for codebook reweighting, since a closed-form solution is obtained.

III. CODEBOOK REWEIGHTING

In order to make codebook weights more informative, we propose to use pairwise constraints to measure the importance of codewords. We denote $I = \{I_i | i = 1, 2, \dots, n\}$ as n images and $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$ are the associated labels. For each image $I_i \in I$, we use sparse coding with max pooling to represent it. In the following, we introduce the image representation briefly. Then, we formulate our codebook reweighting problem with pairwise constraints.

A. Image Representation

For an image, let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{D \times N}$ to be N descriptors (such as SIFT descriptor) where each descriptor is a D -dims vector. The initial codebook obtained by clustering is $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_M\} \in \mathbb{R}^{D \times M}$ which has M cluster centers. $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N\} \in \mathbb{R}^{M \times N}$ are the responses of descriptors to each codeword respectively. For the normal VQ

image representation, the following constrained least square problem is solved:

$$\begin{aligned} \arg \min \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{C}\mathbf{b}_i\| \\ \text{s.t. } \|\mathbf{b}_i\|_0 = 1, \sum_{j=1}^M b_{ij} = 1 \end{aligned} \quad (1)$$

where $\|\bullet\|_0$ stands for the l^0 -norm. This constraint controls the non-zero elements of the responses.

To relax the strong l^0 constraint, in sparse coding, the l^1 constraint is used in the following reconstruction error minimization problem:

$$\begin{aligned} \arg \min \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{C}\mathbf{b}_i\| + \lambda \|\mathbf{b}_i\|_1 \\ \text{s.t. } \sum_{j=1}^M b_{ij} = 1 \end{aligned} \quad (2)$$

where $\|\bullet\|_1$ stands for l^1 -norm. In the training process, codebook and responses are updated iteratively using alternating optimization. Then, the learned codebook is used in the coding process for computing the responses of descriptor. For the whole image representation, max pooling is utilized for statistics of descriptors' responses. We denote the pooling results as n image feature vectors $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\} \in \mathbb{R}^{M \times n}$. Although the learned codebook is more informative than the VQ method, the codebook does not have informative weights, since max pooling can only be considered as a weak weighting process. This might be a problem when there are high intra-class variations and strong background clutters.

B. Pairwise Constrained Codebook Reweighting

To overcome the shortcomings of ScSPM, we propose a codebook reweighting method using pairwise constraints to give the codebook more informative weights than those in ScSPM. As demonstrated in the literature [7]–[11], the pairwise constraints are the natural way of encoding the relationships between images.

Fig. 2. Examples of pairwise constraints from Graz datasets [12]: the left two columns are the cannot-link constraints (red lines); the right two columns are the must-link constraints (green lines).

As shown in Fig. 2, let I_i and I_j to be a must-link pair if $y_i = y_j$ while I_i and I_k to be a cannot-link pair if $y_i \neq y_k$. The image feature vector $\mathbf{p}_i$ generated by max pooling is the associated representation of image I_i . Each element of $\mathbf{p}_i$ can be regarded as the weight of the corresponding codeword. Inspired by [9], [15], we measure the similarity of two images by the intersection of the corresponding image feature vectors. For example, the intersection of I_i and I_j is a vector $\mathbf{d}_{i,j} = \min(\mathbf{p}_i, \mathbf{p}_j)$ with M elements as defined in [15]. Then, the similarity of I_i and I_j is $d_{i,j} = \sum_{m=1}^M \mathbf{d}_{i,j}(m)$ where $\mathbf{d}_{i,j}(m)$ is the m -th element of $\mathbf{d}_{i,j}$. In order to reweight the codebook, we define the reweighted image similarity as $d_{i,j}^{\mathbf{w}} = \mathbf{w}^T \mathbf{d}_{i,j}$ in which $\mathbf{w}$ is the weight vector. Moreover, we construct the pairwise constraints as follows:

$$d_{i,j}^{\mathbf{w}} > d_{i,k}^{\mathbf{w}}, y_i = y_j, y_i \neq y_k, \forall i, j, k. \quad (3)$$

In Equation (3), the reweighted image similarity between images in the same category tends to be larger than those in the different categories. This will reduce the intra-class difference and the influence of the background clutter intuitively. However, along with the desired properties of pairwise constraints, both the computational complexity and the storage complexity will be very high, since every relationship of each image pair is considered. Our solution is to use an online learning algorithm for efficiency.

IV. ONLINE LEARNING ALGORITHM

In this section, we suppose to learn the weights of the codebook which is generated by ScSPM with pairwise constraints in an online manner.

A. Online Formulation

We denote the images $\{(I_i, I_j, I_k) | \forall i, j, k\} \in \mathcal{I}$ as a triplet, $y_i = y_j, y_i \neq y_k$. $d_{i,j,k}^{\mathbf{w}} = d_{i,j}^{\mathbf{w}} - d_{i,k}^{\mathbf{w}} = \mathbf{w}^T (\mathbf{d}_{i,j} - \mathbf{d}_{i,k}) = \mathbf{w}^T \mathbf{d}_{i,j,k}$ is denoted as the margin of the triplet. Similar to SVM, we extend Equation (3) to be:

$$d_{i,j,k}^{\mathbf{w}} > 1, y_i = y_j, y_i \neq y_k, \forall i, j, k. \quad (4)$$

This means that our aim is to achieve a margin of at least 1. Thus, we construct the hinge loss function to be:

$$\ell(d_{i,j,k}^{\mathbf{w}}) = \begin{cases} 0 & , d_{i,j,k}^{\mathbf{w}} > 1 \\ 1 - d_{i,j,k}^{\mathbf{w}} & , \text{otherwise} \end{cases} \quad (5)$$

For a batch learning process, the sum of the hinge loss (5) of the whole triplets should be minimized. However, as we demonstrated in the previous sections, this leads to a large scale task even for small datasets. So we perform an online learning algorithm based on passive-aggressive strategy to update weights iteratively over all triplets.

At each iteration, for example at the t -th iteration, a convex optimization problem with soft margin is solved. The triplets are added to the following update process sequentially:

$$\begin{aligned} \mathbf{w}^t = \arg \min_{\mathbf{w} \in \mathbb{R}^M} \frac{1}{2} \|\mathbf{w} - \mathbf{w}^{t-1}\|_2^2 + C\xi \\ \text{s.t. } \ell(d_{i,j,k}^{\mathbf{w}}) \leq \xi, \xi \geq 0, \forall i, j, k \end{aligned} \quad (6)$$

In this iterative update process, the initial weight vector $\mathbf{w}^0$ is set to be zero vector. The hinge loss $\ell(d_{i,j,k}^{\mathbf{w}})$ at t -th iteration is minimized while the updated weights $\mathbf{w}^t$ should keep as close as possible to the previous weights $\mathbf{w}^{t-1}$. It is regarded as a passive-aggressive process, since it updates the model aggressively when the hinge loss is large and passively in order to maintain the information of last iterations. The passive and aggressive trade-off is controlled by the parameter C .

B. Optimization Process

Equation (6) can be rewritten as the following Lagrangian:

$$L(\mathbf{w}, \tau_1, \tau_2, \xi) = \frac{1}{2} \|\mathbf{w} - \mathbf{w}^{t-1}\|_2^2 + C\xi + \tau_1(1 - \mathbf{w}^T \mathbf{d}_{i,j,k} - \xi) - \tau_2\xi, \quad (7)$$

where $\tau_1 \geq 0$ and $\tau_2 \geq 0$ are the Lagrangian multipliers.

We set the partial derivative of $L(\mathbf{w}, \tau_1, \tau_2, \xi)$ with respect to ξ to zero and get

$$C - \tau_1 - \tau_2 = 0. \quad (8)$$

Since problem (6) is convex, the optimal solution should fulfill that the partial derivatives of $L(\mathbf{w}, \tau_1, \tau_2, \xi)$ with respect to the elements of $\mathbf{w}$ are zero,

$$\frac{\partial L(\mathbf{w}, \tau_1, \tau_2, \xi)}{\partial \mathbf{w}} = \mathbf{w} - \mathbf{w}^{t-1} - \tau_1 \mathbf{d}_{i,j,k} = 0. \quad (9)$$

Thus, the optimal $\mathbf{w}$ should be updated as

$$\mathbf{w} = \mathbf{w}^{t-1} + \tau_1 \mathbf{d}_{i,j,k}. \quad (10)$$

This is a function of τ_1 .

Substituting Equations (8) and (10) to the Lagrangian (7), we can get the following result after grouping the terms:

$$L(\tau_1) = -\frac{1}{2} \tau_1^2 \|\mathbf{d}_{i,j,k}\|^2 + \tau_1(1 - (\mathbf{w}^{t-1})^T \mathbf{d}_{i,j,k}), \quad (11)$$

then we take the derivative of $L(\tau_1)$ with respect to τ_1 to zero and have

$$\tau_1 = \frac{1 - (\mathbf{w}^{t-1})^T \mathbf{d}_{i,j,k}}{\|\mathbf{d}_{i,j,k}\|^2}. \quad (12)$$

Since $\tau_2 \geq 0$, we have $\tau_1 \leq C$ according to Equation (10). Then we get

$$\tau_1 = \min(C, \frac{1 - (\mathbf{w}^{t-1})^T \mathbf{d}_{i,j,k}}{\|\mathbf{d}_{i,j,k}\|^2}). \quad (13)$$

Thus, the final update process has a closed-form solution as Equation (10) and (13) for each triplet iteratively.

V. EXPERIMENTAL EVALUATION

In this section, we perform experiments on Graz-01 and Graz-02 datasets [12] which are standard image classification datasets. We compare the proposed method with other state-of-the-art methods on Graz datasets. This comparison demonstrates the effectiveness of the proposed method. Then, we compare our method with the similar batch learning algorithm [9] to evaluate the efficiency of our online learning algorithm.

A. Experimental Setting

We use Graz datasets for experimental evaluation. In Graz datasets, there are three object classes which are “person”, “bike” and “car” as well as the “background” class. The classification task is “object” vs. “background” (“no object”). The datasets are challenging because of the high intra-class variations and strong background clutters. Graz-01 dataset contains “person”, “bike” object classes and “background” class. Graz-02 dataset has all three object classes and “background” class. The examples are referred to [12].

For Graz-01 dataset, we use 100 images per “object” and 100 images per “background” for training and 100 images for test with each class has a half of the images. For Graz-02 dataset, these three numbers equal 150.

As introduced in Section III-A, we use ScSPM for image representation. First, we extract SIFT descriptors densely. Then, the codebook size that we use equals 2000. We choose SPM level as $L = 0$ for Graz-01 and $L = 2$ for Graz-02. In our experiments, we choose 1×1 , 2×2 and 3×1 cells for the associated three SPM levels when $L = 2$. At last, we use kNN classifier for the final classification in all experiments. The experiments of codebook reweighting are carried out on a standard 2.83GHz Intel Core2 Quad Q9550 CPU with Matlab.

B. Image Classification on Graz-01 & 02

We compare the performance of our method on Graz datasets with the state-of-the-art methods including some codebook weighting methods. This comparison is used to demonstrate the effective of our method. We adopt the Equal Error Rate (EER) of the Receiver-Operating Characteristic (ROC) for evaluation with the same settings in [9]. All the results reported in this section are averaged over 10 runs with randomly selected training and test images.

Method	Bike	Person
SPM [5]	86.3	82.3
Cai et al. [9]	86.7±4.6	84.0±2.0
Opelt et al. [12]	83.5	76.5
Zhang et al. [23]	92.0	88.0
Equal_w	85.2±2.4	89.4±3.1
Our method	87.2±2.4	95.4±2.2

TABLE I
THE ROC EER ON GRAZ-01 DATASET.

Method	Bike	Car	Person
Cai et al. [9]	84.9±2.1	82.2±2.0	84.9±2.0
Opelt et al. [12]	76.5	70.7	81.0
Tuytelaars et al. [24]	89.5	80.2	85.2
Yang et al. [25]	85.1	87.1	90.2
Equal_w	84.5±2.7	74.1±1.9	88.1±2.8
Our method	89.3±1.3	81.7±1.1	95.3±2.0

TABLE II
THE ROC EER ON GRAZ-02 DATASET.

Table I summaries the results on Graz-01 dataset. The results of our method are better than the results without reweighting (“Equal_w” in Table I). Our method outperforms the classic SPM [5] and the first work [12] on Graz datasets. Compared with the codebook weighting method [9], we get significant improvement on “person” and better performance

on “bike”. The best performance so far is obtained by [23]. We get better performance on “person”. Classifying “bike” is more challenging because [23] used two types of feature detectors, two descriptors and SVM classifier [9].

Table II summarizes the results on Graz-02 dataset. Compared with the method without reweighting (“Equal_w” in Table II), the reweighted results get a significant improvement, especially for “person” and “car”. Our method also outperforms [12] significantly. For codebook weighting method [9], our method makes an improvement on “bike” and “person”. Compared with the best performance so far [24], [25], we obtain lower performance on “car”, but better on “person” and comparable on “bike”. In this paper, we only use simple dense sampling strategy. However, an biased sampling strategy was used in [25]. This strategy combined a prior map learned from the training data and a bottom-up saliency map.

In order to illustrate the efficiency of our method, we compare the proposed online learning algorithm with [9]. As pairwise constraints were also used in [9], an Alternating Optimization (AO) algorithm has been proposed to improve the Global Optimization (GO). Our online learning algorithm is an online extension of the above algorithms. Since the computational complexities of AO and GO were only reported on Graz-02 dataset, we conduct the experiments on this dataset as well. The experiments are carried out 10 times. The GO algorithm needs **36 hours** in average with 30,000 triplets. The AO algorithm needs **1 hours and 15 minutes** for the same settings. The proposed online learning algorithm for codebook reweighting only needs **17 minutes** in average with 3,352,500 ($149 \times 150 \times 150$) triplets. Our method is much faster than GO and AO because it has a closed-form solution for training without any optimization tools.

VI. CONCLUSIONS

In this paper, we have proposed an online learning algorithm using pairwise constraints for image classification. We have analyzed the advantages of the classic BoW model and ScSPM. However, the relationships between pairs of images are not taken into account in previous work. This might be a problem when there are high intra-class variations and background clutters. Therefore, we have proposed a codebook reweighting method. We have utilized pairwise constraints in order to encode the relationships between pairs of images. After codebook reweighting, the must-link images tend to have higher similarity than the cannot-link images. Since using pairwise constraints induced a large scale problem, we have proposed an online learning algorithm to deal with this problem. Our algorithm is computationally efficient because it learns incrementally and has closed-form solution. The quality of the codebook reweighting is evaluated on Graz datasets and is compared to the state-of-the-art methods. The results demonstrated the effectiveness and efficiency of our method.

In future, we will study the proposed codebook reweighting method combined with other BoW model based methods, and experimentally analyze its performance on other existing image classification datasets.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China (Grant No.60875021,60723005). The authors would like to thank Zhang Zhang for his suggestions and help.

REFERENCES

- [1] G. Csurka, C. Bray, C. Dance, and L. Fan, “Visual categorization with bags of keypoints,” in Proc. ECCV, 2004.
- [2] J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman, and W. T. Freeman, “Discovering object categories in image collections,” in Proc. CVPR, 2005.
- [3] J. Mutch and D. Lowe, “Multiclass object recognition with sparse, localized features,” in Proc. CVPR, 2006.
- [4] C. Wallraven, B. Caputo, and A. Graf, “Recognition with local features: the kernel recipe,” in Proc. ICCV, 2003.
- [5] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in Proc. CVPR, 2006.
- [6] J. Yang, K. Yu, Y. Gong, and T. Huang, “Linear spatial pyramid matching using sparse coding for image classification,” in Proc. CVPR, 2009.
- [7] M. Schultz and T. Joachims, “Learning a distance metric from relative comparisons,” in Proc. NIPS, 2004.
- [8] A. Frome, Y. Singer, F. Sha, and J. Malik, “Learning globally-consistent local distance functions for shape-based image retrieval and classification,” in Proc. ICCV, 2007.
- [9] H. Cai, F. Yan, and K. Mikołajczyk, “Learning weights for codebook in image classification and retrieval,” in Proc. CVPR, 2010.
- [10] Pavan K. Mallapragada, Rong Jin, and Anil K. Jain, “Online Visual Vocabulary Pruning Using Pairwise Constraints,” in Proc. CVPR, 2010.
- [11] Kilian Q. Weinberger and Lawrence K. Saul, “Distance metric learning for large margin nearest neighbor classification,” Journal of Machine Learning Research, 2009.
- [12] A. Opelt, A. Pinz, M. Fussenegger, and P. Auer, “Generic object recognition with boosting,” TPAMI, 28(3):416-431, 2006.
- [13] Koby Crammer, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, and Yoram Singer, “Online passive-aggressive algorithms,” Journal of Machine Learning Research, 2006.
- [14] Gal Chechik, Varun Sharma, and Uri Shalit, “An Online Algorithm for Large Scale Image Similarity Learning,” in Proc. NIPS, 2009.
- [15] S. Maji, A.C. Berg, and J. Malik, “Classification using intersection kernel support vector machines is efficient,” in Proc. CVPR, 2008.
- [16] F. Jurie and B. Triggs, “Creating efficient codebooks for visual recognition,” in Proc. ICCV, 2005.
- [17] J. C. van Gemert, J. M. Geusebroek, C. J. Veenman, and A. W. M. Smeulders, “Kernel codebooks for scene categorization,” in Proc. ECCV, 2008.
- [18] B. Fulkerson, A. Vedaldi, and S. Soatto, “Localizing objects with smart dictionaries,” in Proc. ECCV, 2008.
- [19] L. Yang, R. Jin, R. Sukthankar, and F. Jurie, “Unifying discriminative visual codebook generation with classifier training for object category recognition,” in Proc. CVPR, 2008.
- [20] J. Mairal, M. Leordeanu, F. Bach, M. Hebert, and J. Ponce, “Discriminative sparse image models for class-specific edge detection and image interpretation,” in Proc. ECCV, 2008.
- [21] R. Yan, J. Zhang, J. Yang, and Alexander G. Hauptmann, “A discriminative learning framework with pairwise constraints for video object classification,” TPAMI, vol. 28, no. 4, 2006.
- [22] C. Zhang, J. Liu, J. Wang, Q. Tian, C. Xu, H. Lu, and S. Ma, “Image classification using spatial pyramid coding and visual word reweighting,” in Proc. ACCV, 2010.
- [23] J. Zhang, M. Marszałek, S. Lazebnik, and C. Schmid, “Local features and kernels for classification of texture and object categories: A comprehensive study,” IJCV, 73(2):213-238, 2007.
- [24] T. Tuytelaars and C. Schmid, “Vector quantizing feature space with a regular lattice,” in Proc. ICCV, 2007.
- [25] L. Yang, N. Zheng, J. Yang, M. Chen, and H. Chen, “A biased sampling strategy for object categorization,” in Proc. ICCV, 2009.