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 summaries 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. Mikolajczyk, "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. Marszalek, 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.