

Acknowledgments

This work was supported by National Hi-Tech Research and Development Program (863 Program) of China under Grants 2014AA015102 and 2012AA012503, National Natural Science Foundation of China under Grant 61371128, and Ph.D. Programs Foundation of Ministry of Education of China under Grant 20120001110097.

References

- Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; and Süsstrunk, S. 2012. Slic superpixels compared to state-of-the-art superpixel methods. *TPAMI* 34(11):2274–2281.
- Chen, Y.; Zhu, L.; Yuille, A.; and Zhang, H. 2009. Un-supervised learning of probabilistic object models (poms) for object classification, segmentation, and recognition using knowledge propagation. *TPAMI* 31(10):1747–1761.
- Everingham, M.; Van Gool, L.; Williams, C.; Winn, J.; and Zisserman, A. 2010. The PASCAL visual object classes (VOC) challenge. *IJCV* 88(2):303–338.
- Farabet, C.; Couprie, C.; Najman, L.; and LeCun, Y. 2013. Learning hierarchical features for scene labeling. *TPAMI* 35(8):1915–1929.
- Han, Y.; Wu, F.; Shao, J.; Tian, Q.; and Zhuang, Y. 2012. Graph-guided sparse reconstruction for region tagging. In *CVPR*, 2981–2988.
- He, R.; Zheng, W.-S.; Hu, B.-G.; and Kong, X.-W. 2013. Two-stage nonnegative sparse representation for large-scale face recognition. *TNNLS* 24(1):35–46.
- Liu, X.; Cheng, B.; Yan, S.; Tang, J.; Chua, T. S.; and Jin, H. 2009a. Label to region by bi-layer sparsity priors. In *ACM MM*, 115–124.
- Liu, X.; Yan, S.; Yan, J.; and Jin, H. 2009b. Unified solution to nonnegative data factorization problems. In *ICDM*, 307–316.
- Liu, D.; Yan, S.; Rui, Y.; and Zhang, H.-J. 2010. Unified tag analysis with multi-edge graph. In *ACM MM*, 25–34.
- Liu, S.; Yan, S.; Zhang, T.; Xu, C.; Liu, J.; and Lu, H. 2012a. Weakly supervised graph propagation towards collective image parsing. *TMM* 14(2):361–373.
- Liu, X.; Yan, S.; Cheng, B.; Tang, J.; Chua, T.-S.; and Jin, H. 2012b. Label-to-region with continuity-biased bi-layer sparsity priors. *ACM TOMCCAP* 8(4):50.
- Liu, Y.; Liu, J.; Li, Z.; Tang, J.; and Lu, H. 2013. Weakly-supervised dual clustering for image semantic segmentation. In *CVPR*, 2075–2082.
- Liu, C.; Yuen, J.; and Torralba, A. 2009. Nonparametric scene parsing: Label transfer via dense scene alignment. In *CVPR*, 1972–1979.
- Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. *IJCV* 60(2):91–110.
- Lu, Z., and Peng, Y. 2013. Latent semantic learning with structured sparse representation for human action recognition. *PR* 46(7):1799–1809.
- Ng, A. Y.; Jordan, M. I.; and Weiss, Y. 2001. On spectral clustering: Analysis and an algorithm. *NIPS* 14:849–856.
- Shotton, J.; Winn, J.; Rother, C.; and Criminisi, A. 2009. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *IJCV* 81(1):2–23.
- Shotton, J.; Johnson, M.; and Cipolla, R. 2008. Semantic texton forests for image categorization and segmentation. In *CVPR*, 1–8.
- Tang, J.; Hong, R.; Yan, S.; Chua, T.-S.; Qi, G.-J.; and Jain, R. 2011. Image annotation by knn-sparse graph-based label propagation over noisily tagged web images. *ACM TIST* 2(2):14.
- Tighe, J., and Lazebnik, S. 2013. Superparsing: scalable nonparametric image parsing with superpixels. *IJCV* 101(2):329–349.
- Verbeek, J., and Triggs, B. 2007. Region classification with markov field aspect models. In *CVPR*, 1–8.
- Vezhnevets, A., and Buhmann, J. M. 2010. Towards weakly supervised semantic segmentation by means of multiple instance and multitask learning. In *CVPR*, 3249–3256.
- Vezhnevets, A.; Ferrari, V.; and Buhmann, J. M. 2011. Weakly supervised semantic segmentation with a multi-image model. In *ICCV*, 643–650.
- Vezhnevets, A.; Ferrari, V.; and Buhmann, J. M. 2012. Weakly supervised structured output learning for semantic segmentation. In *CVPR*, 845–852.
- Winn, J., and Jojic, N. 2005. Locus: Learning object classes with unsupervised segmentation. In *ICCV*, 756–763.
- Yan, S., and Wang, H. 2009. Semi-supervised learning by sparse representation. In *SDM*, 792–801.
- Yang, Y.; Yang, Y.; Huang, Z.; Shen, H. T.; and Nie, F. 2011. Tag localization with spatial correlations and joint group sparsity. In *CVPR*, 881–888.
- Yang, Y.; Huang, Z.; Yang, Y.; Liu, J.; Shen, H. T.; and Luo, J. 2013. Local image tagging via graph regularized joint group sparsity. *PR* 46(5):1358–1368.
- Yuan, J.; Li, J.; and Zhang, B. 2008. Scene understanding with discriminative structured prediction. In *CVPR*, 1–8.
- Yuille, A. L., and Rangarajan, A. 2003. The concave-convex procedure. *Neural Computation* 15(4):915–936.
- Zhang, D.; Islam, M. M.; Lu, G.; and Sumana, I. J. 2012. Rotation invariant curvelet features for region based image retrieval. *IJCV* 98(2):187–201.
- Zhang, K.; Zhang, W.; Zheng, Y.; and Xue, X. 2013a. Sparse reconstruction for weakly supervised semantic segmentation. In *IJCAI*, 1889–1895.
- Zhang, L.; Song, M.; Liu, Z.; Liu, X.; Bu, J.; and Chen, C. 2013b. Probabilistic graphlet cut: Exploiting spatial structure cue for weakly supervised image segmentation. In *CVPR*, 1908–1915.
- Zhu, G.; Yan, S.; and Ma, Y. 2010. Image tag refinement towards low-rank, content-tag prior and error sparsity. In