TPA 14: Comparative Study of the Performances of very recent Feature Extractors, used for Detection, Matching and Recognition

January, 2015

Problem Statement: Detailed analysis of recent feature extractors for various high level tasks like detection, matching (stereo, motion, mosaic) and recognition. (Select any two tasks)

Some Examples of Recent Features: CNN or Decaf, VLAD, VM-1SIFT, BF-DSIFT, SV-DSIFT, LL-MO1SIFT, Fisher Vector Pyramid, CDH, MSD, 3D Harris, ST-SIFT, CHOG-3D, 4-D LST etc.

Input:

- Recognition: Training Images with Labels
  Testing Images with Labels (Labels to measure the Performance)
- Detection: Training Images with Bounding Boxes and Labels
  Testing Images with Labels (Bounding Boxes to measure the performance)
- Matching: Images with Labels (Labels to measure the Performance), for estimating Homography, creating MOSAIC etc.

Expected Output:

- Comparison of different feature extractors or techniques (minimum 6 recent ones) in terms of accuracy, average precision and precision-recall. (depending on task)

Dataset: PASCAL VOC 2007 [9] , Caltech-101 [10] and other video or image datasets. (Minimum 4 datasets)

Note: First use of feature extractors published in rich literature, should not be more than a two years old publications in top conferences and journals.

References

1. Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." CVPR, 2014.
2. Kantorov, Vadim, and Ivan Laptev. "Efficient feature extraction, encoding and classification for action recognition.", CVPR, 2014.
3. Zhang, Hao, et al. "Simplex-Based 3D Spatio-temporal Feature Description for Action Recognition.", CVPR, 2014
4. Chatfield, Ken, et al. "Return of the Devil in the Details: Delving Deep into Convolutional Nets." *BMVC, 2014*.

5. He, Kaiming, et al. “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition.”, ECCV 2014.
6. Xie, Lingxi “Hierarchical Part Matching for Fine-Grained Visual Categorization”, ICCV 2013.
7. Mittelman, Roni et al. “Weakly Supervised Learning of Mid-Level Features for Object Recognition.”, CVPR 2013.
8. Wang, Hua, “Heterogeneous Visual Features Fusion via Sparse Multimodal Machine”, CVPR 2013.
9. http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2007/
10. http://www.vision.caltech.edu/Image_Datasets/Caltech101