DIRAC Training - Report

LUO Jie (Roger)
IDIAP Research Institute
Centre du Parc, Av. des Pres-Beudin 20, Martigny, Switzerland
jluo@idiap.ch

December 18, 2007

Since the visiting period was six weeks, which was relative short, this period was mainly dedicated to learn different techniques and make a plan for the future research works after my return to IDIAP. Thus, experiments were only conducted within small scales in order to test the applicability of the developed ideas. And these techniques are of potential interest to DIRAC applications and closely related to the new DIRAC "white paper". I will briefly summarize what have been done in the following part of this report.

1. At the beginning of this visit, we submitted a joint conference paper [1] on the topic of object category detection using audio-visual cues. We plan to extend this multimodal system using hierarchical representation [2]. Under the same framework, I conducted a few experiments on detecting new object category in leaf level using only visual inputs. The preliminary results turned out to be not very good. The false negative rate is high. We identified that possible problems might come from the optimization process which is not designed for detecting novel objects. By extracting more information such as separate patches and scale from the constellation models [3] would help reduce false alarm rate. In general, this problem was difficult and the system framework would require careful design for the task while keeping the same performance on known object categories. It would be a promising direction for future work.

2. Unexpected combinations in certain known context could be considered as an interesting type of rare events. Using the context information for doing classification is a much research topic in vision groups at MIT (e.g. [4] & [5] and reference therein). The context information could be used to predict the likely location of an object, and improve the accuracy of object detection. For detecting unexpected out-of-context events, it is important to model the context information first. I learned these related works and their limitations from people in the lab.

3. One of main technical difficulties in doing novelty and outlier detection, perhaps other machine learning problems as well, is that data is usually represented by feature in high dimension and these observations are usually noisy. Thus, a powerful while task-specific feature prune/selection techniques would significantly improve the system's performance. Learning a distance function [6] from data with side constrain informations would be a good solutions for the task at hand. During this period, I studied different distances functions[7] develop by Prof. Weinshall's students.

Acknowledge

I am grateful to Prof. Daphna Weinshall for supervising my work. I would also like to acknowledge Jake Bouvrie, Sharat Chikkerur, Tony Ezzat, Hueihan Jhuang and Cheston Tan for their fruitful discussions, and Roie Kliper for giving me an introduction on distance functions. Also many thanks to all staffs of CBCL, MIT for their friendliness and support.

References

[1] L. Jie, B. Caputo, A. Zweig, J. Bach and J. Anemüller. Object Category Detection Using Audio-visual Cues. IDIAP-RR-07-58, Submitted to *International Conference on Computer Vision System (ICVS'08)*.

[2] A. Zweig and D. Weinshall. Exploiting Object Hierarchy: Combining Models from Different Category Levels. In *Proc. of 11th IEEE International Conference of Computer Vision (ICCV'07)*, Rio de Janeiro Brazil, October 2007.

[3] A. Bar-Hillel and D. Weinshall. Efficient Learning of Relational Object Class Models. *International Journal of Computer Vision (IJCV)*, 2007 (in press)

[4] K. Murphy, A. Torralba, D. Eaton, W. T. Freeman. Object detection and localization using local and global features. *Lecture Notes in Computer Science*. Sicily workshop on object recognition, 2005.

[5] L. Wolf and S. Bileschi. A critical view of context. *International Journal of Computer Vision (IJCV)*, 2006.

[6] A. Bar-Hillel, T. Hertz, N. Shental and D. Weinshall. Learning a Mahalanobis Metric from equivalence constraints. *Journal of Machine Learning Research (JMLR)*, 2005.

[7] Tomer Hertz. Learning Distance Functions with Applications in Image Retrieval, Clustering, Classification and Immunoinformatics. Ph.D Thesis. Hebrew University of Jerusalem, 2007.