Local-LDA: Open-Ended Learning of Latent Topics for 3D Object Recognition in Robotics

S. Hamidreza Kasaei[1,3], Luís Seabra Lopes[1,2], Ana Maria Tomé[1,2]

1 — Department of Electronics, Telecommunications and Informatics & IEETA, University of Aveiro
2 — Department of Electronics, Telecommunications and Informatics, University of Aveiro
3 — Currently at the University of Groningen, The Netherlands

Object recognition is still a challenging problem because of ill-definition of objects, large variations in object appearance and concept drift in dynamic environments. Moreover, since in open-ended domains the set of categories to be learned is not known in advance, autonomous robots must have the ability to continuously execute learning and recognition in a concurrent and interleaved fashion. The training instances are extracted from on-line experiences of the robot and become gradually available over time, rather than being completely available at the beginning of the learning process.

We developed an open-ended 3D object recognition system capable of concurrently learning the object categories as well as the features used to encode them. A multi-layered object representation approach is used to enhance open-ended learning (Fig. 1). Each object view is hierarchically described as a random mixture over a set of Latent Dirichlet Allocation (LDA) topics, where each topic is discrete distribution over visual words. Topic modelling is suitable for open-ended learning because it provides short object descriptions and enables efficient processing of large collections. Given the open-ended setting, we propose an extension of LDA to learn topics for each category independently (Local-LDA). These local topics are discovered in an unsupervised fashion and updated incrementally using new object views. In the lowest layer, the spin-image descriptor is used to represent the local shapes of the objects in different key points; in a bag-of-words layer, the given object view is described by a histogram of local shape features; in the topic layer, we use the visual words to compute a set of LDA topics for the given object. Since the views of an object can vary significantly depending on the perspective, and the categories themselves are often heterogeneous, the category layer stores distinctive views of different objects in each object category.

An extensive set of experiments was carried out to assess the performance of Local-LDA and compare it with other state-of-art approaches. Local-LDA outperformed the other approaches, achieving appropriate descriptiveness and scalability with respect to increasing number of categories (Fig. 2).

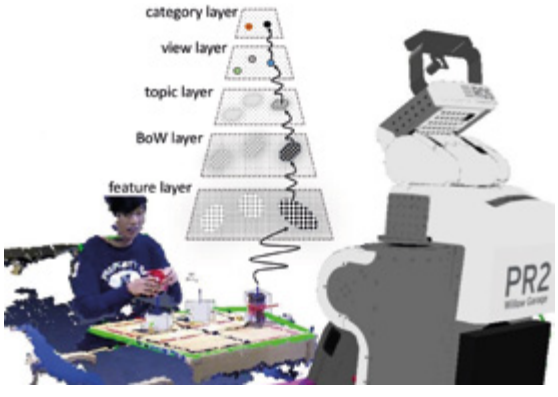
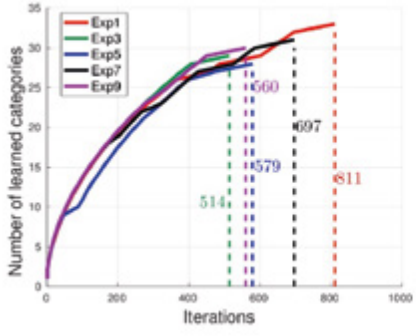


FIGURE 1
Layered approach.



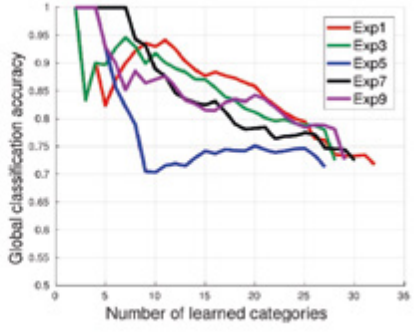


FIGURE 2
Experimental results for Local-LDA