A bottom-up and top-down optimization framework for learning a compositional hierarchy of object classes

Sanja Fidler Marko Boben Aleš Leonardis
Faculty of Computer and Information Science
University of Ljubljana, Slovenia
{sanja.fidler, marko.boben, ales.leonardis}@fri.uni-lj.si

Learning hierarchical representations of object structure in a bottom-up manner faces several difficult issues. First, we are dealing with a very large number of potential feature aggregations. Furthermore, the set of features the algorithm learns at each layer directly influences the expressiveness of the compositional layers that work on top of them. However, we cannot ensure the usefulness of a particular local feature for object class representation based solely on the local statistics. This can only be done when more global, object-wise information is taken into account.

We build on the hierarchical compositional approach [1] that learns a hierarchy of contour compositions of increasing complexity and specificity. Each composition models spatial relations between its constituent parts.

The main drawback of the original approach is that learning is performed strictly bottom-up — once a layer is learned it accepts no further revisions. Performed in this way, important information gets either lost or a very large, possibly redundant set of features must be chosen at each layer in order to compensate for the potential loss. Since the ultimate goal is to learn a joint hierarchical representation of a higher number of object classes, this problem becomes even more pronounced.

Here we seek for a compact (non-redundant) multi-class hierarchical vocabulary while also ensuring it contains all object relevant information. The idea of this work is to cast layer learning into a stochastic optimization framework that iteratively improves the hierarchy as a whole. Optimization is two-fold: one that learns and selects the compositions in a bottom-up phase and the other that extends/improves them by top-down feedback from the higher layers.

In the bottom-up phase, we first learn a large set of compositions at each layer and employ a stochastic optimization process to select the least redundant subset that explains the data sufficiently well. Since this procedure may select a set of compositions that, when further combined, produce lower expressiveness in the higher layers, we additionally perform a top-down optimization step. Here we revise the set of compositions as to maximally improve the expressiveness of the layer above. The algorithm then loops between the two learning stages.

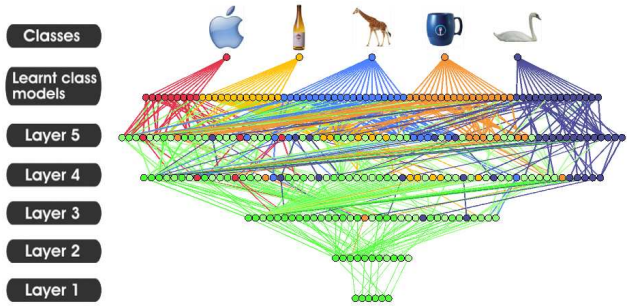

Figure 1. A compact multi-class hierarchical compositional representation. The nodes denote distinct compositions and the links represent the compositional relations. The bottom layer contains a small number of different types of contour fragments. The green nodes depict the compositions shareable between several classes.

We evaluated the approach on several object classes. Even for a single class, the learned hierarchy is about 3 times more compact and inference is 2.5 times faster than that of the original approach [1]. Moreover, we improve the detection performance by almost 10%. As such, the proposed optimization framework enables a significantly more efficient and reliable coding of multiple object classes.

In the multi-class experiments, we used ten diverse object classes. The joint representation is highly compact and takes only 1Mb on disk. We additionally demonstrate several important issues. The proposed approach achieves a sub-linear growth in size of the hierarchy and, consequently, a sub-linear inference times as the number of modeled classes increases. Furthermore, we demonstrate a competitive detection performance with respect to the current state-of-the-art results for several object classes. Fig. 1 shows an example of a learned hierarchy for five classes.

References

[1] S. Fidler and A. Leonardis. Towards scalable representations of visual categories: Learning a hierarchy of parts. In *CVPR*, 2007.

978-1-4244-3993-5/09/$25.00 ©2009 IEEE