Recursive Compositional Models: Representation, Learning, and Inference

Long (Leo) Zhu
CSAIL
MIT
leozhu@csail.mit.edu

Alan Yuille
Dept. Statistics, Psychology, and Computer Science
UCLA
yuille@stat.ucla.edu

Abstract

Recursive compositional models *(RCMs) [1, 2, 3, 4] are hierarchical models which enable us to represent the shape/geometry and visual appearance of objects and images at different scales. The key design principle is* recursive compositionality. *Objects are represented by RCMs in a hierarchical form where complex structures are composed of more elementary structures. Formally, they are represented by probability distributions defined over graphs with variable topology. Learning techniques are used to learn these models from a limited number of examples of the object by exploiting the recursive structure (some of our papers use supervised learning while others are unsupervised and induce the object structure). In addition, we can exploit this structure to develop algorithms that can perform inference on these RCMs to rapidly detect and recognize objects. This differs from more standard "flat models" of objects which have much less representational power if they wish to maintain efficient learning and inference. The basic properties of an RCM are illustrated in figures (1,2). Because RCMs give a rich hierarchical description of objects and images they can be applied to a range of tasks including object detection, segmentation, parsing and image parsing. In all cases, we achieved state of the art results when evaluated on datasets with groundtruth.*

References

[1] L. Zhu, Y. Chen, Y. Lu, C. Lin, and A.L. Yuille. Max Margin AND/OR Graph Learning for Parsing the Human Body. CVPR 2008.

[2] L. Zhu, Y. Chen, X. Ye, and A.L. Yuille. Structure-Perceptron Learning of a Hierarchical Log-Linear Model. CVPR 2008.

[3] L. Zhu, Y. Chen, Y. Lin, and A.L. Yuille. Recursive Segmentation and Recognition Templates for Parsing. NIPS 2008.

[4] L. Zhu, C. Lin, H. Huang, Y. Chen, and A.L. Yuille. Unsupervised Structure Learning: Hierarchical Composition, Suspicious Coincidence and Competitive Exclusion. ECCV 2008.
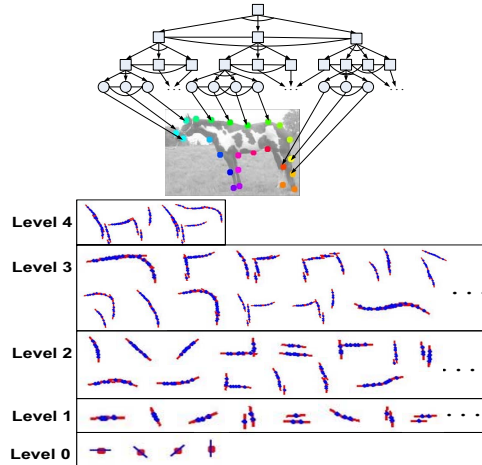
Figure 1. Top panel: The compositional representation of the object [4]. The boxes represent non-leaf nodes. The circles denote leaf nodes that directly relate to properties of the input image. Bottom panel: This figure shows the mean shapes of elements of the hierarchical dictionary.
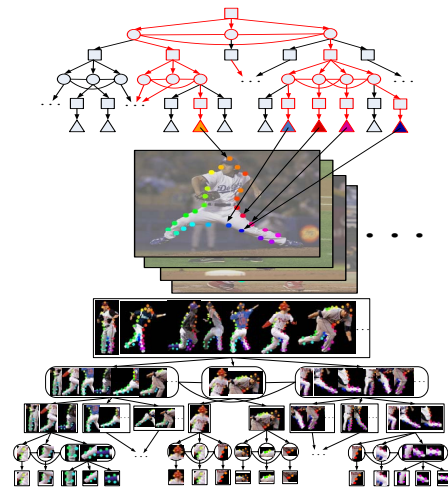

Figure 2. Top Panel: AND-OR RCMs use OR nodes (circles) to alter the graph topology making them flexible enough to deal with the different poses of baseball players [1]. Bottom Panel: 100 different topologies, player poses, can be represented by a graph with only 40 nodes.