Automatic Discovery of Groups of Objects for Scene Understanding

Congcong Li (Cornell) Devi Parikh (TTIC) Tsuhan Chen (Cornell)

Motivation

How would we understand an image?

Scene?

Outdoor Restaurant

Objects?

3 Picnic-umbrellas, 3 Tables,6 Chairs

Group of Objects!

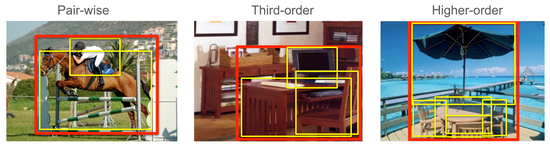
3 *Sets* of picnic-umbrella, table & chairs

Groups of objects: composites of two or more objects which have mutually *consistent* spatial, scale, and view point relationships.

Problem: It is **NOT** feasible to manually compile a list of all possible groups with arbitrary number of participating objects!

Contributions

❖ Modeling a full spectrum of **arbitrary high-order** object interactions for deeper scene understanding

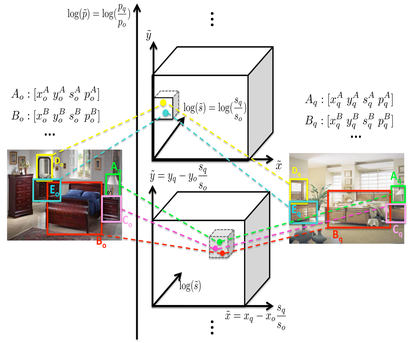


❖ **Automatically discovering** groups from images annotated only with object labels

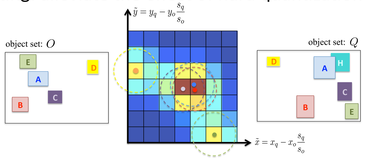
❖ **Improving object detection and scene recognition** performance on a variety of datasets: UIUC phrasal, PASCAL VOC07, SUN09, MIT indoor

Approach: Group Discovery

Step1: Find common object patterns between every image-pair through a **4-dimensional transform space**.



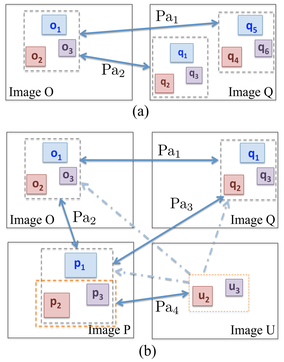
• Soft voting: alleviate the effect of hard quantization



Step2: Clustering patterns into groups.

• Assume transitivity between patterns

• Allow missing participating objects: lower-order groups instantiations are merged with corresponding higher-order group instantiations



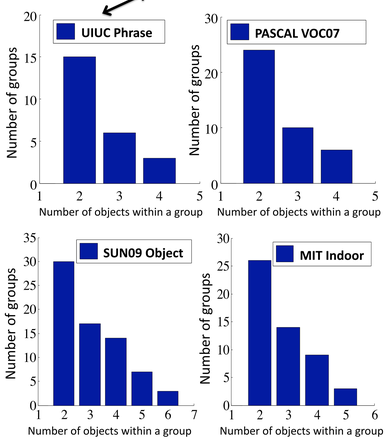
Step3: Training group detectors.

• Generate a bounding box for each instantiation of the group: the smallest box that encompasses all participating objects including the hallucinated missing object.

• Utilize any off-the-shelf object detection method to train group detectors. We used the deformable part-based model.

Results: Discovered Groups

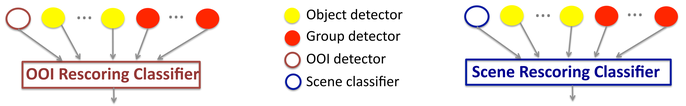
Manual labeling [Sadeghi & Farhadi CVPR 2011]: 12 pair-wise phrases

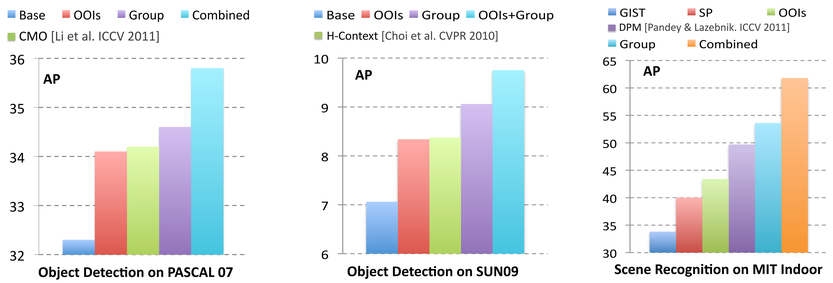


High-order groups are discovered on multiple datasets!

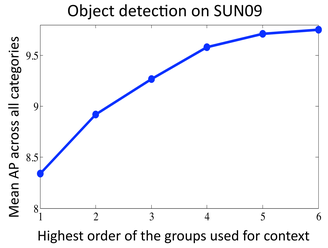
Results: Improved Scene Understanding

Contextual reasoning





Higher-order groups provide useful contextual information!



For more details, please visit: http://chenlab.ece.cornell.edu/projects/objectgroup