Hierarchical Semantics of Objects (hSOs)
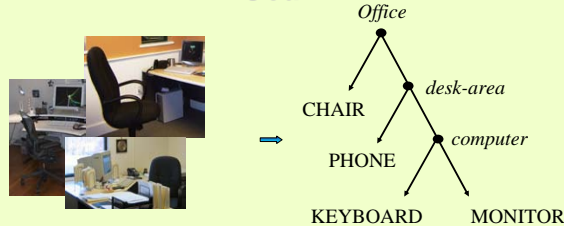
Devi Parikh and Tsuhan Chen, Department of Electrical and Computer Engineering, Carnegie Mellon University

What is an hSO?

What do you see?

Offices

Monitors, keyboards, phones and chairs

Computers, phones and chairs

Desk-stuff and chairs

Goal

Office
desk-area
CHAIR
computer
PHONE
KEYBOARD   MONITOR

Images from a scene ⟹ Corresponding hSO

UNSUPERVISED !

Motivations

❖ **Higher level understanding of the scene**
❖ **Compact scene representation**
❖ **Context**
❖ **Avoid extensive human labeling – let the data speak!**

Contributions

❖ **LEARNING: Unsupervised learning of hSO**
❖ **APPLICATION: Using hSO for robust object detection**

LEARNING: Approach

Input

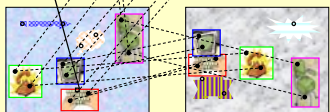### Feature extraction

Geometrically consistent correspondences
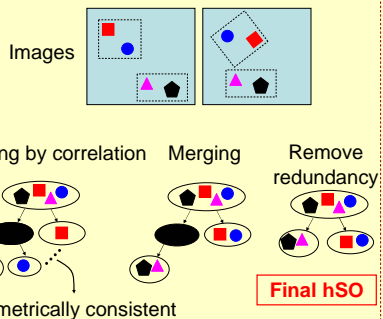
[Leordeanu, ICCV 2005]

Feature selection

Did not find correspondences in enough images

Did not find geometrically consistent correspondences

Split-n-merge clustering

Images

Splitting by correlation   Merging   Remove redundancy

Geometrically consistent

Final hSO

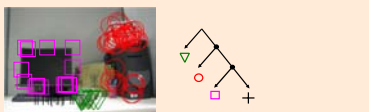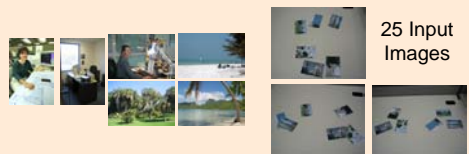### Interaction among feature pairs
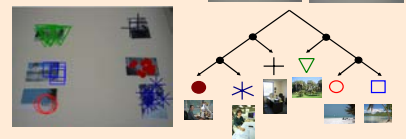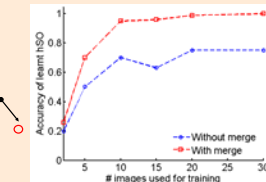
Recursive clustering ?
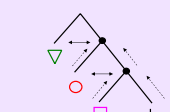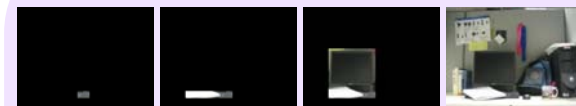
LEARNING: Results

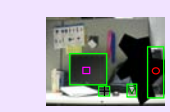30 Input Images

Results

25 Input Images

Results

30 Input Images

Results

Accuracy of learnt hSO

images used for training

-- Without merge
-- With merge

APPLICATION

run individual detectors

context by ▢   refined ✛   detector
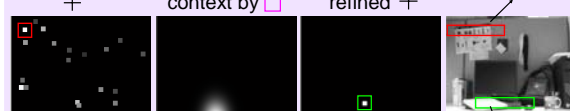
detector + context

detector   detector + context

Conclusion

❖ **Identify foreground**
❖ **Separate foreground into multiple objects**
❖ **Learn interactions among objects**
❖ **Infer hierarchical semantics among objects**
❖ **Learn models for objects**
❖ **Use models & semantics for robust object detection**

UNSUPERVISED!