OBJECT BANK: A high level image representation for scene classification and semantic feature sparsification

Li-Jia Li, Hao Su, Eric P. Xing, Li Fei-Fei

Presentation by Vivek Dubey, Guided by Prof. Yong Jae Lee
Motivation
Motivation
Motivation
Object Bank

- Off the shelf classifiers such as LR and SVM
- Provides superior performance on high level visual recognition tasks
- Used sparsity algorithms to make this work more efficient and hence scalable for large scale datasets
A typical Classification Task
Comparisons!
Object Bank

An image is represented as a scale invariant response map of a large number of pre-trained generic object detectors
The Object Bank Representation of Images

- Pre trained detectors are outsourced
Latent SVM detectors
(Felzenszwalb et. al. TPAMI 2010)
Texture Classifier
(Hoeim et. al., SIGGRAPH, 2005)
The Object Bank Representation

Original Image

Object Detector Responses
- Sailboat
- Water
- Bear

Spatial Pyramid

Object Bank Representation

Max Response (OB)
- Sailboat
- Water
- Sky
- Bear
Dataset- 15-Scene
(Lazebnik et.al. CVPR, 2006)
Dataset- LabelMe
(Torralba et.al, MIT AI Lab Demo, 2005)
Dataset- MIT Indoor
(A.Torralba et.al., CVPR 2009)
Dataset-UIUC Sports
(Li Fei-Fei et.al., ICCV 2007)
Results across different test datasets
Discussions about OB

- Why 200 objects? Why not all?
- Robustness wrt Training Sample Size
- Near losslessness of Content Based Compression
- Profitability Over Growing OB
Robustness
Near Losslessness of Content Based Compression

![Graphs showing compression of image representation with accuracy and dimension percentage.]
Profitability of Growing OB
To sum it up!

- Semantic meaning to scenes implies better categorization
- We need to take care of the humongous volume of visual data that we have and this work helps us move towards organization of that data
- It is the stuff for the future!