

CS331 Autumn 2012-13

Week 2, October 4th

Bangpeng Yao

Maji et al, CVPR 2011 (9:08am – 9:58am)

Presenter: Qifeng Chen

- Poselets
 - What is poselet? (Bourdev & Malik, ICCV 2009).
 - Clusters of human pose configurations.
 - Need annotation of key point locations of humans for each training image.
 - Can we determine the human's action from a poselet?
 - Local feature; not necessary to tell the human's global pose or action.
 - There are also poselets covering the whole body.
- Poselet activation vector
 - Just confidence score of independent detectors; No spatial location encoded.
 - Computation is very expensive: search image windows in many locations and scales for poselet; But it is parallelizable.
 - Refer to the object bank paper (Li et al, NIPS 2010)
- 3D pose estimation from still images
 - View classification
- Static action classification
 - Select action specific poselets, and then compute poselet activation vectors (PAV)
 - Add spatial interactions of humans and objects (horse, bike, etc); Mixture of model
- General comments:
 - The paper seems to be a mixture of some different models.
 - The key technical contribution is PAV; but only using PAV does not achieve good performance.
 - Interesting comparison between deformable part model (DPM, Felzenszwalb et al, CVPR 2008) and PAV.
 - DPM: much less parts (6 in DPM vs 150 in PAV)
 - DPM: high dimensions of low level representation vs. PAV: confidence scores of object detection

Yang et al, CVPR 2010 (10:03am – 10:34am)

Presenter: Kevin Tang

- Joint pose estimation and action classification; Main goal: action classification; Poselet as hidden variables.
- Poselets
 - Not exactly the same as the original Bourdev paper or the Maji paper.
 - Clustering of 2D joint locations; Similar semantic meaning.
- Model
 - Conditional random field (CRF): appearance, pairwise, root, and global components.
 - When representing an image, we always should consider both appearance (unary) and shape configurations (pairwise).
 - Learning:
 - Different from typical latent SVM, here the pose assignments are observable during training.
 - Avoid over-fitting: Pose assignments are abandoned during training.
- Results
 - The method improves the performance of all the classes except running.
 - Pose of running is similar with walking?
- Summary
 - Baseline is too weak (HOG+SVM).