Learning a Discriminative Hidden Part Model for Human Action Recognition (Wang et al)
Presenter: Evan
9:05am

Focuses on Video rather that still images.
Three main approaches: Global, Bag of Words, Part based models

Global and Bag of Words are not mutually exclusive. One could use local HoG features as words in along with BoW/features.

HoG vs SIFT: Both gradient based. Both are somewhat rotation invariant
SIFT: Detector + Key Point descriptor.
HoG: More global representation.

In paper we want to detect Salient feature. For a given image, first detect action classes in the image and then use hidden features to find relation between action classes (hCRF). The hidden variables are learnt from training (they are initialized randomly to begin with)

Motion tracking is done using Lucas-Kanade optical flow. The key assumption is that the people/action is located at the center of the image. Motion descriptors and patch location compromise a part based model

For testing, run all the action feature through an image to find the stingiest salient feature.

Voting across frames is used to indicate the action in the video.
The data set being used is very weak: one person in the center, fixed background, fixed aspect ratio.

Action Bank: A High-Level Representation of Activity in Video (Sreemananth CVPR 2012) (10:00 am)
Presenter: Kyunghee

template based method.

Use action bank inspired from Object bank

Use 7 different energies: Left, right up, down, flicker, static and lack of orientation
Use steerable filter to calculate the above enrages

Use UCF sport and KTH data set

Use reference [6] as their starting point.

Max pooling is is used for every template.

A total length of action back feature vector is : Na * Ns * 73 (Na: action class, Ns: different scale)

They seem to perform the best in both KTH and UCF dataset.

Extremely large processing time compare to others.