Social Interactions : A First person Perspective [Fathi et al CVPR 2012]

Problem: Detecting recognizing social interactions [dialogue, discussion, monologue, walk dialogue, walk discussion]

Setup:

- Head mounted cameras (first person views / egocentric view)
- 8 persons
- 3 days in amusement park (42 hours)

Advantage: First person camera always captures where person is attending

Building blocks of the algorithm

1. Location of faces
   - view angle from camera
   - distance from camera
2. Model for Attention

   Markov Random field with graph of face locations
   - Unary potential terms
     - Gaussian function: possibility of f_i looking at location l
     - Sigmoid term : threshold the distance between faces- avoid face looking at itself
     - Bias faces to look at faces rather than objects
   - Pairwise potentials
     - Bias faces to look at same location in the scene

Method

Features to capture social interation:

1. Location of faces around first person
2. Roles based on patterns of attention
   a. Number of faces looking at person x
   b. Whether first person looks at person x
   c. Mutual attention between first person and x
   d. Number of faces looking at where x is attending
3. First person head movement (based on optical flow)

Social interactions are inferred from the temporal patterns of the above features using Hidden Conditional Random Field

Results

- ROC curves and confusion matrix presented, which seems to be doing well.
- This is first work and introduces new dataset, so comparison with other methods is not presented
- Also show how social network can be built using this method.