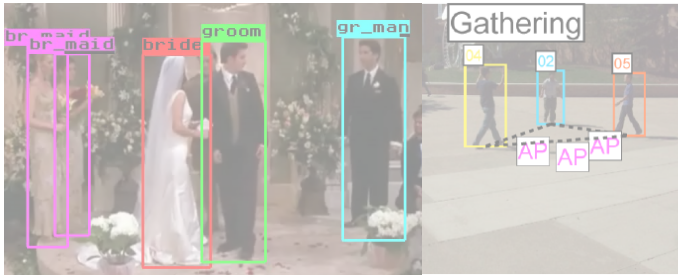
Scene dynamism

Dynamic scene

Static scene

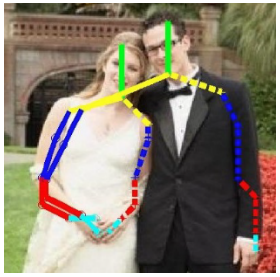
Rehg, CVPR13
Prabhaker, ECCV12
Prabhakar, CVPR12
Patron-Perez, BMVC10



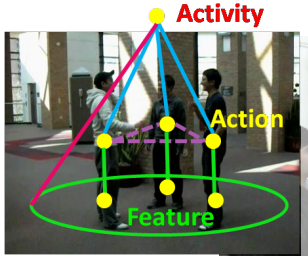
Lan, CVPR12
Ramanathan, CVPR13
Antic, ECCV14

Ding, ECCV10
Choi, ECCV12, CVPR14
Direkoglu, ECCV12

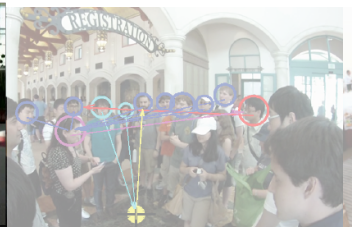
Rodriguez, ICCV11a, ICCVb
Mehran, CVPR09
Alahi, CVPR14



Yang, CVPR12
Hoai, CVPR14



Fathi, CVPR12
Choi, ECCV14
Park, NIPS12, ICCV13



Cristani, BMVC11
Park, CVPR15
Arev, SIGGRAPH14

Wang, ECCV10
Gallagher, CVPR09

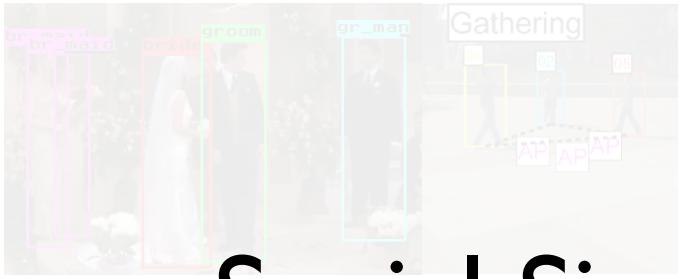
Dyadic interaction

Crowd interaction

Number of group members

Body Pose as Social Signals

Rehg, CVPR13
Prabhaker, ECCV12
Prabhakar, CVPR12
Patron-Perez, BMVC10



Antic, ECCV14
Direkoglu, ECCV12

Rodriguez, ICCV11a, ICCVb
Mehran, CVPR09
Alahi, CVPR14

Yang, CVPR12
Hoai, CVPR14

Fathi, CVPR12
Choi, ECCV14
Park, NIPS12, ICCV13

Cristani, BMVC11
Park, CVPR15
Arev, SIGGRAPH14

Wang, ECCV10
Gallagher, CVPR09

Scene dynamism

Dynamic scene

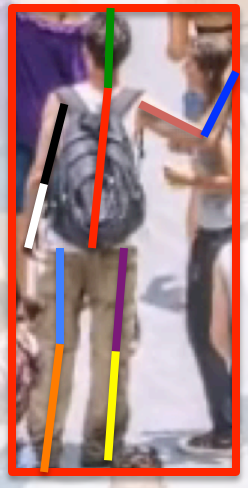
Static scene

Dyadic interaction

Crowd interaction

Number of group members

Pose as a skeleton



+ Almost complete representation.
+ Provide fine details of the person.

- Hard to estimate in complex scene.
- Large amount of variation.

Pose as a finite set of semantic classes

Standing
Sitting
Running
Facing-Front
Facing-Right
...

Standing and Facing-forward

+ Compact representation.
+ Higher accuracy.

- Require definition of classes.
- Less details.

Standing and Facing-forward

The left touches the right's head.

Proxemics

Group Activity Recognition

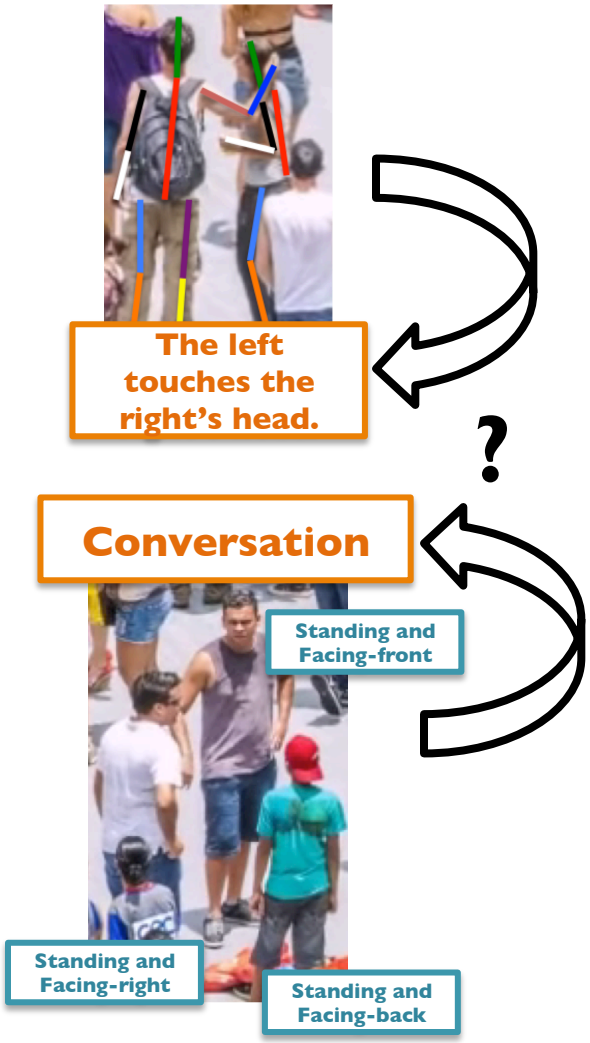
Conversation

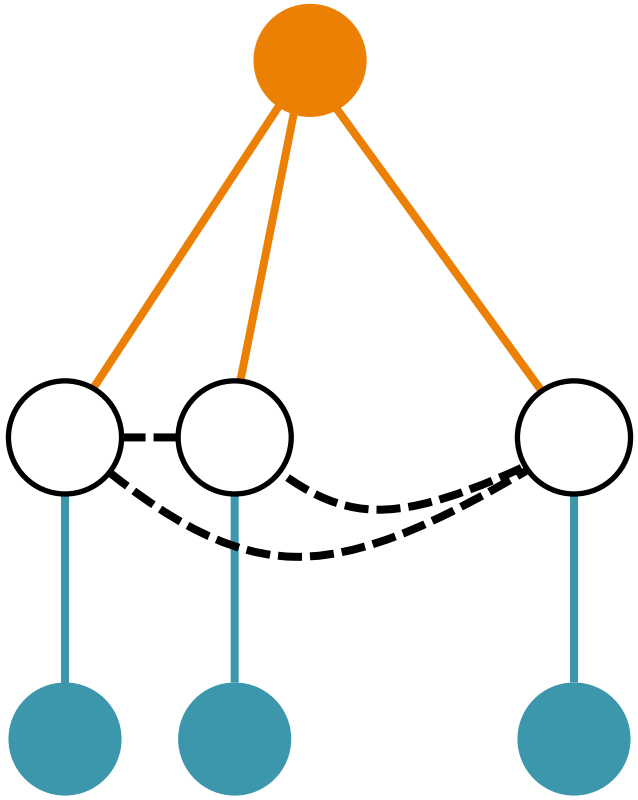
Standing and Facing-front

Standing and Facing-right

Standing and Facing-back

Exploiting Structures in Group Behavior

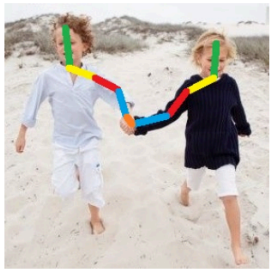


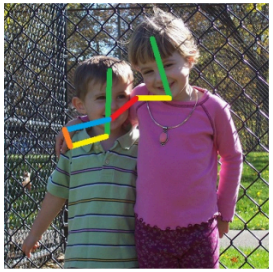


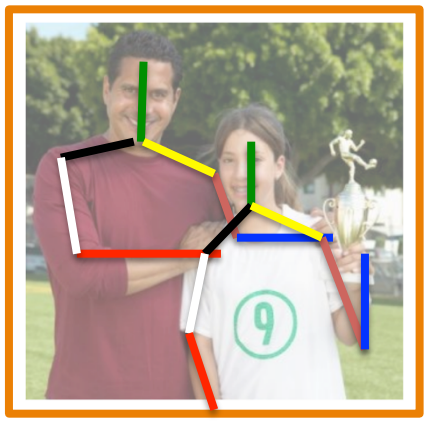
Proxemics in Personal Photos (Yang CVPR12)

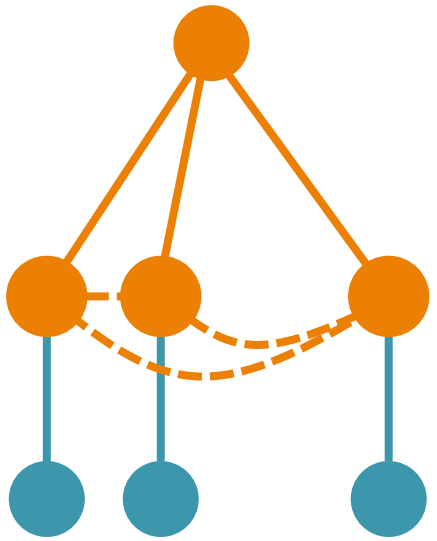
Input: image

Output: proxemics label
and skeletons









[Yang CVPR12]

Proxemics

Proxemics: the study of spatial arrangement of people as they interact

- anthropologist Edward T. Hall in 1963

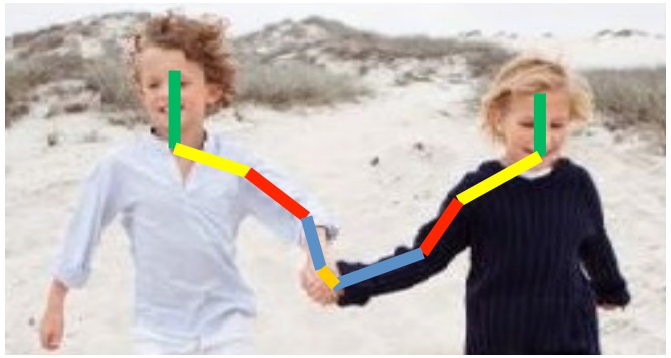
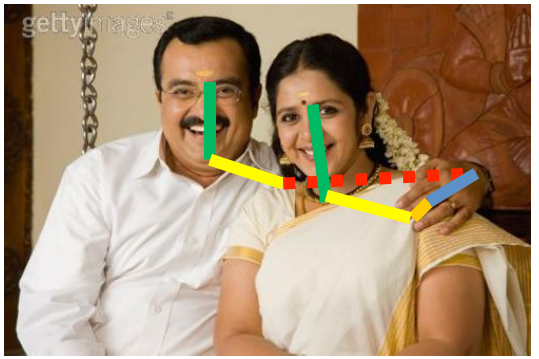
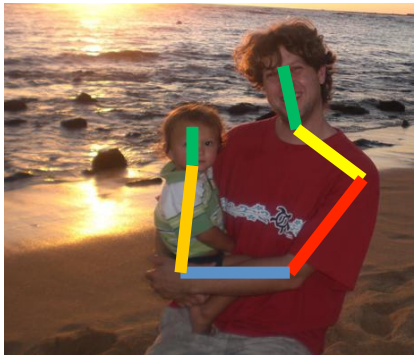
[Yang CVPR12]

Proxemics in Personal Photos

Hand touch Hand	Hand touch Shoulder
	
Shoulder touch Shoulder	**Arm touch Torso**
	

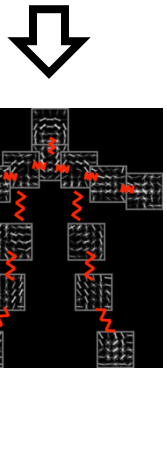
[Yang CVPR12]

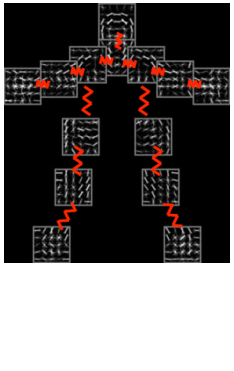
Proxemics in Personal Photos

Hand touch Hand	Hand touch Shoulder
	
Shoulder touch Shoulder	**Arm touch Torso**
	

[Yang CVPR12]

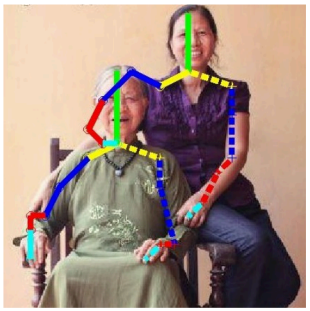
Naïve Solution



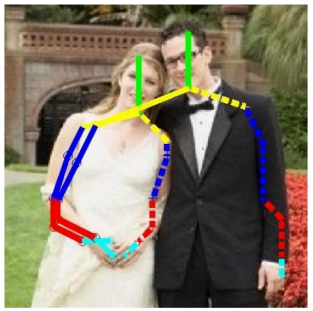


Hand touch Hand

(a) Extreme Pose



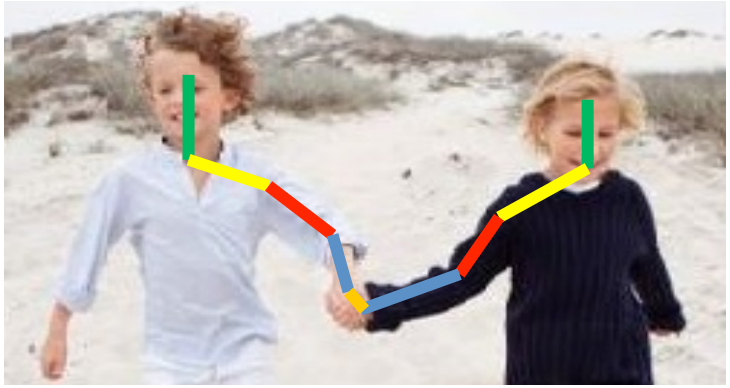
(b) Occlusion



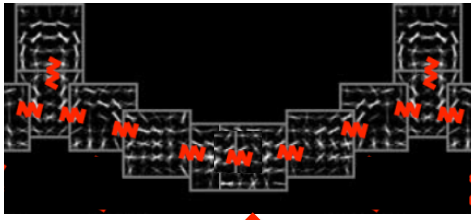
(c) Part Ambiguity

[Yang CVPR12]

Joint Pose Estimation



Hand touch Hand

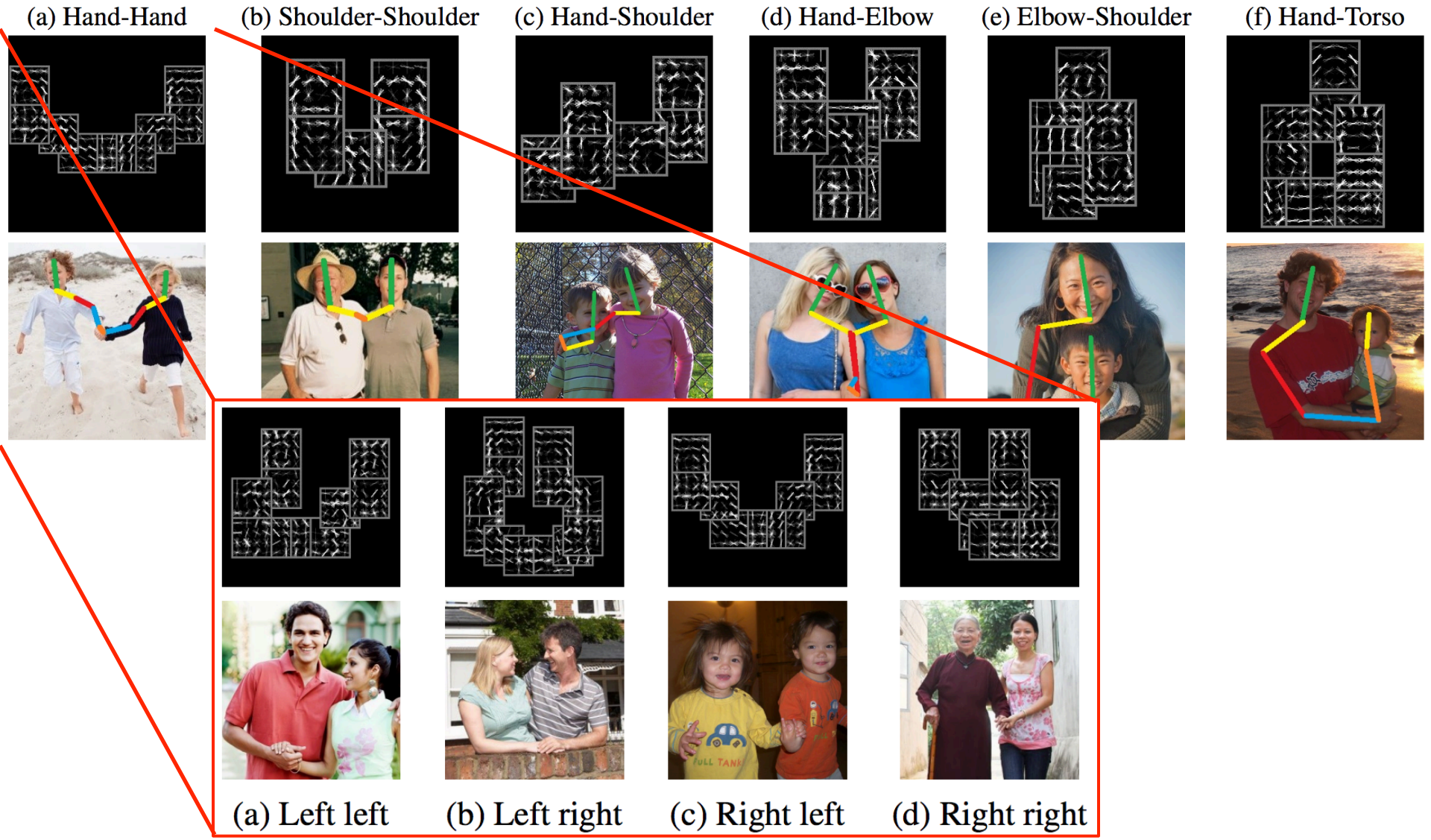


Touch code

Tree structure!
Efficient Inference.

[Yang CVPR12]

Proxemic Dependent Pose Models



[Yang CVPR12]

Experimental Setup

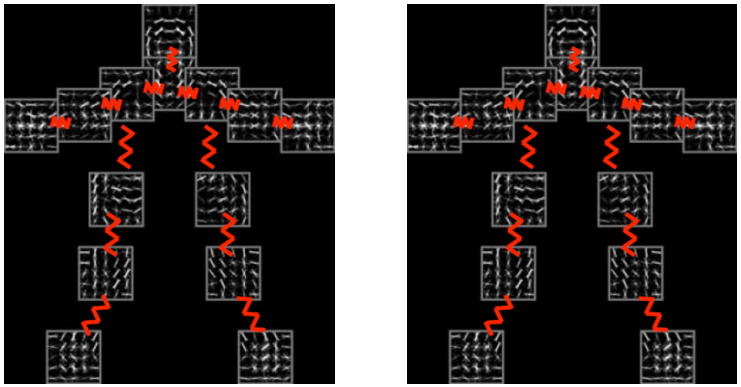
- Proxemic Dataset
  - Available at: http://www.ics.uci.edu/~yyang8/research/proxemics/index.html
  - 589 images, 1207 people, 1332 pairs.





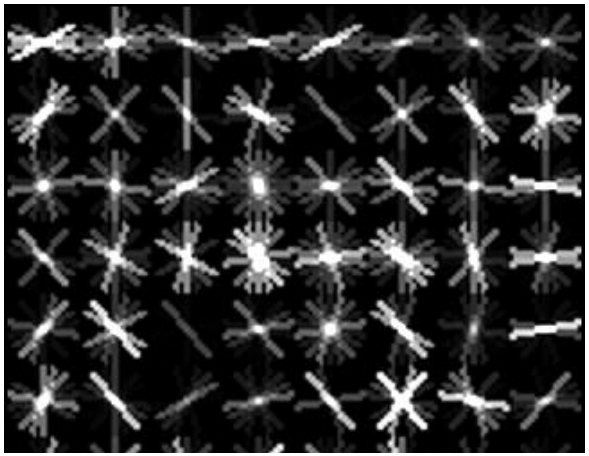

[Yang CVPR12]

Baselines



Distance?

Baseline 1: Sequential

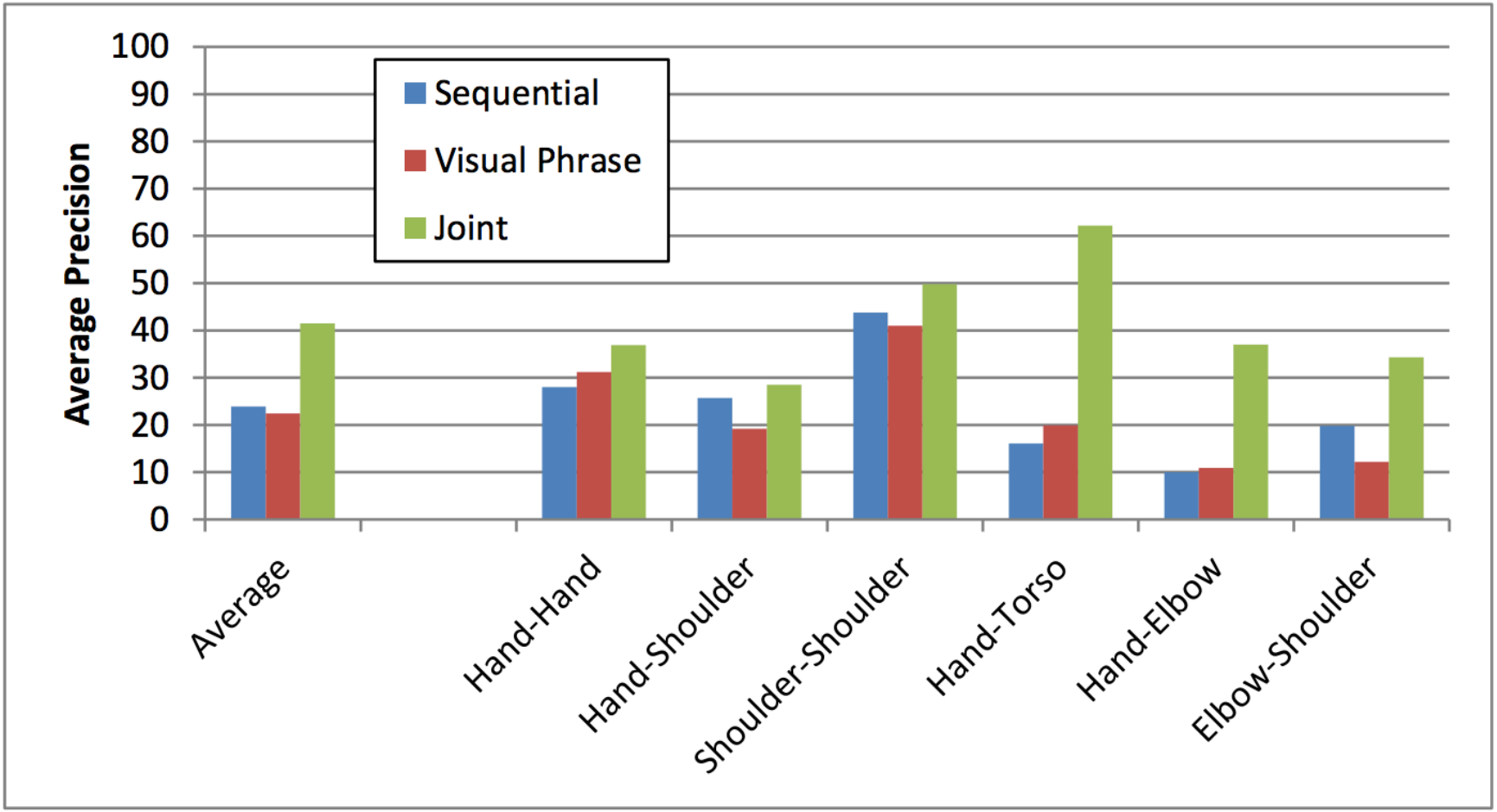


HoG Template

Baseline 2: Phrases

[Yang CVPR12]

Quantitative Evaluation



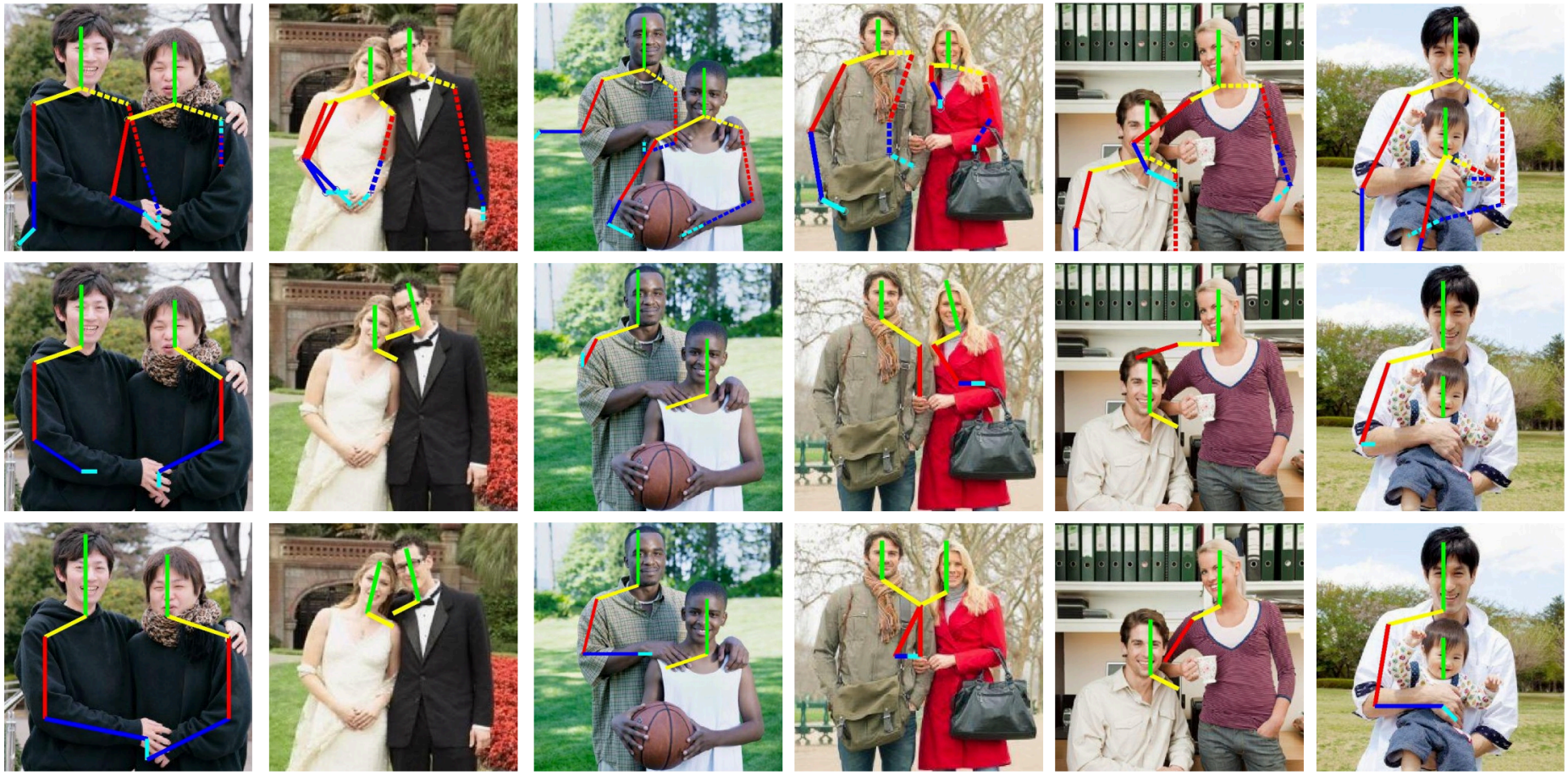
[Yang CVPR12]

Quantitative Evaluation

- Pose Accuracy:
  - Sequential Estimation: 47.5%
  - Joint Estimation: 73.6%

[Yang CVPR12]

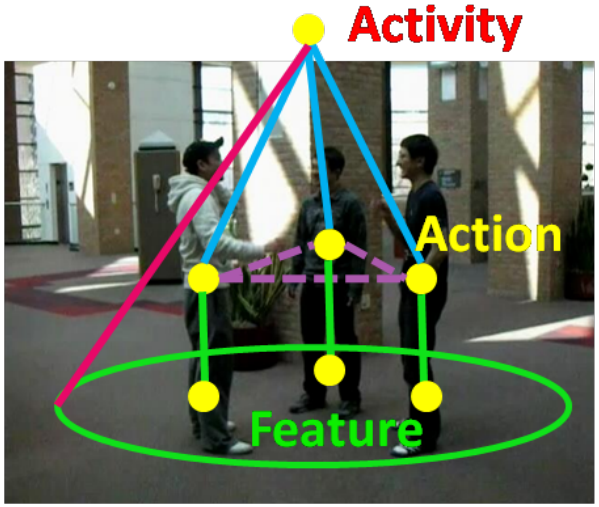
Qualitative Examples

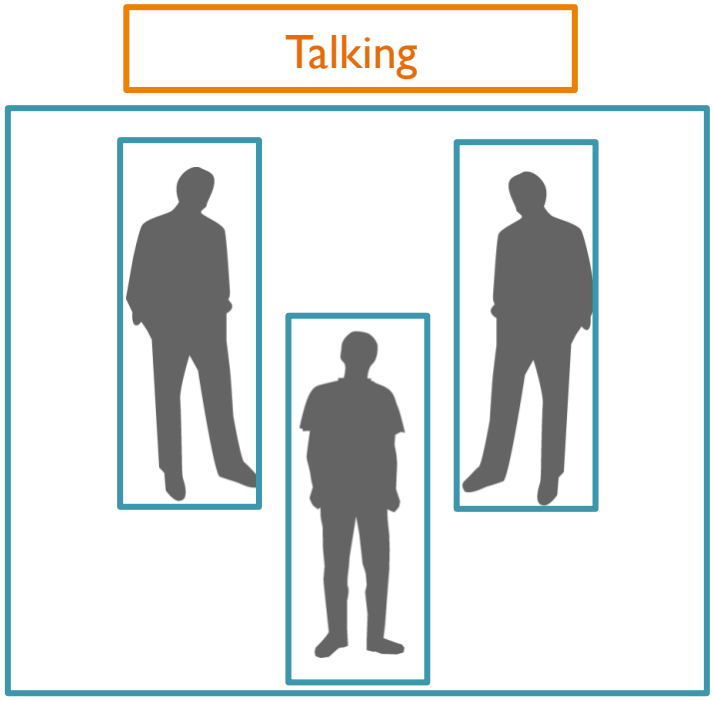


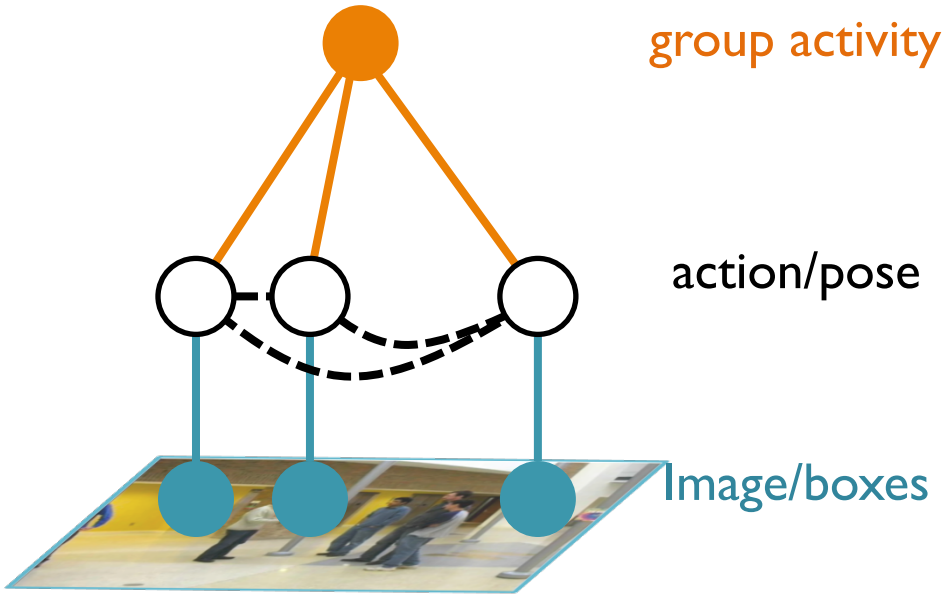
Group Activity Recognition (Lan NIPS10)

Input: image and bounding boxes

Output: group activity label







Collective Activities

Definition:

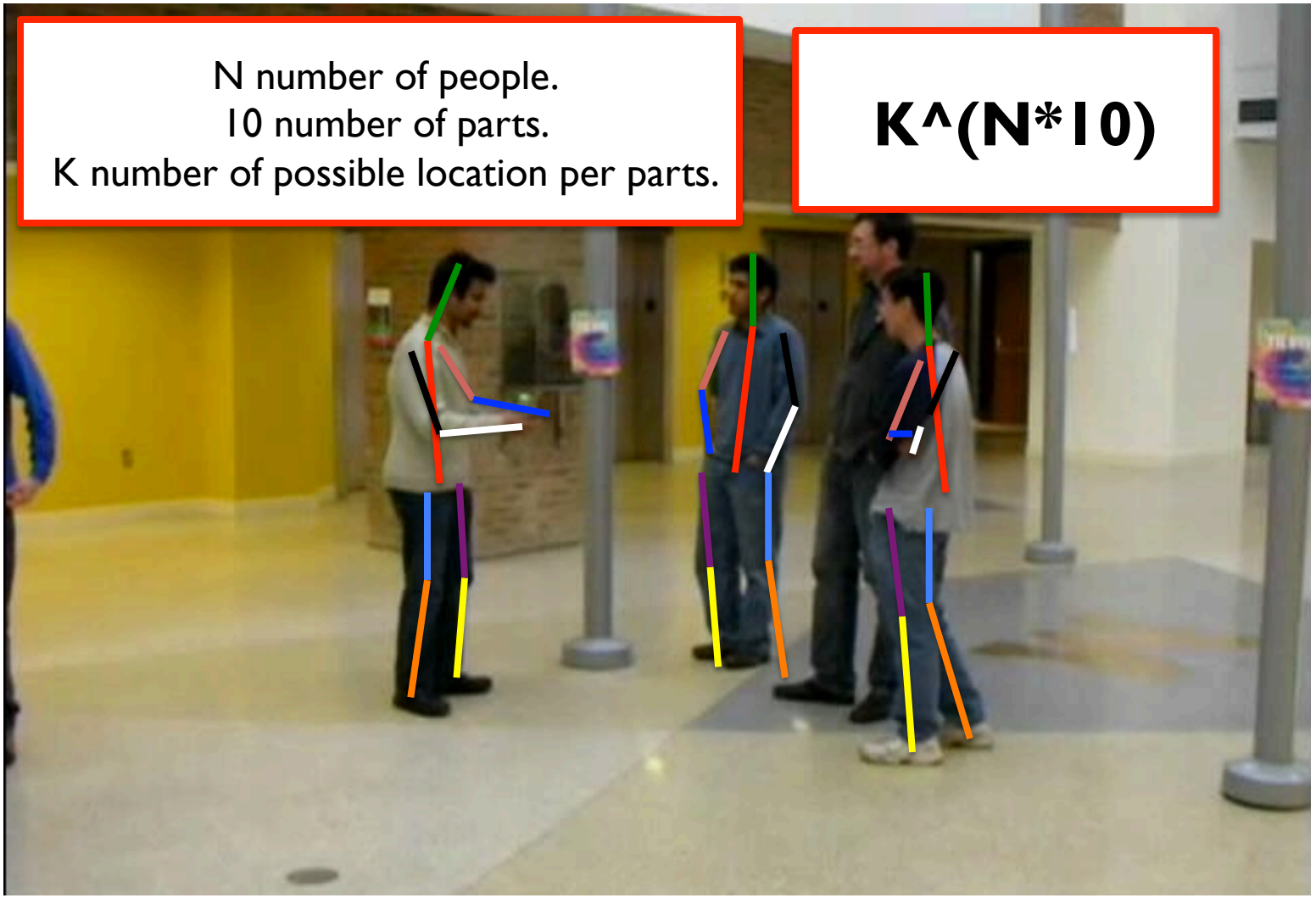
Activities that are defined or reinforced by the existence of a coherent behavior of a group of individuals in time and space.

Waiting

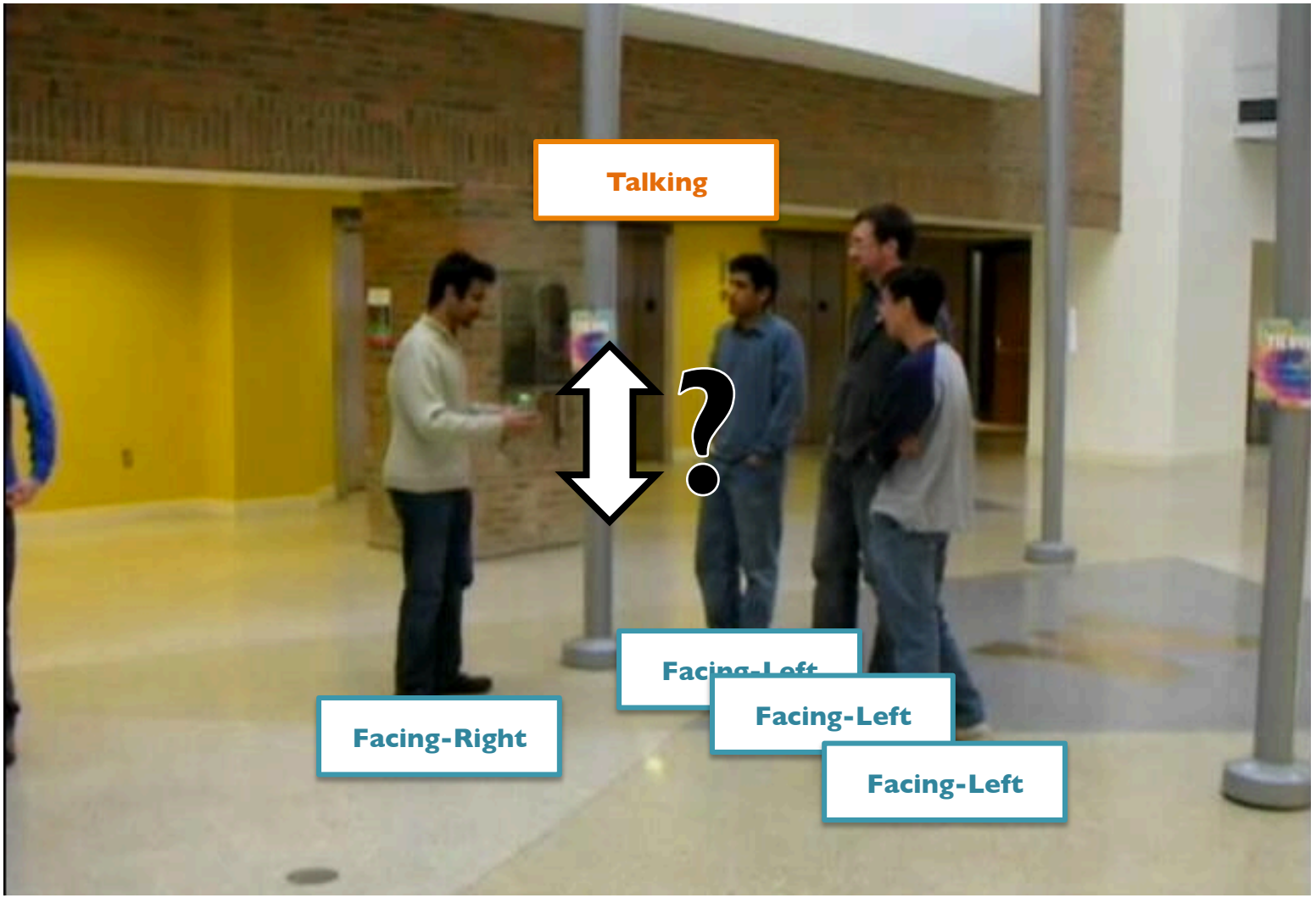
Queuing

Talking

Pose Representation for Groups

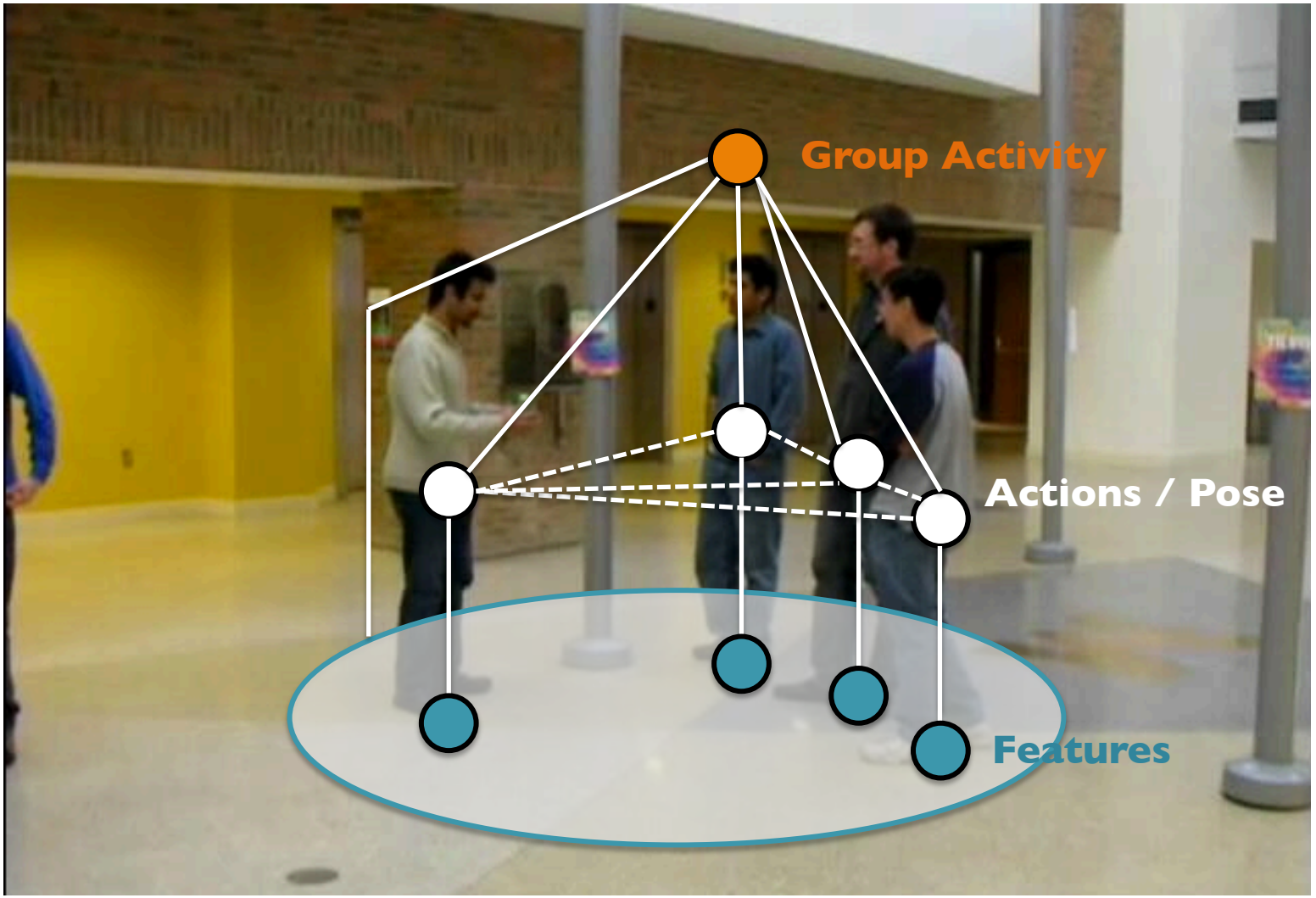


Pose to Group Behavior?



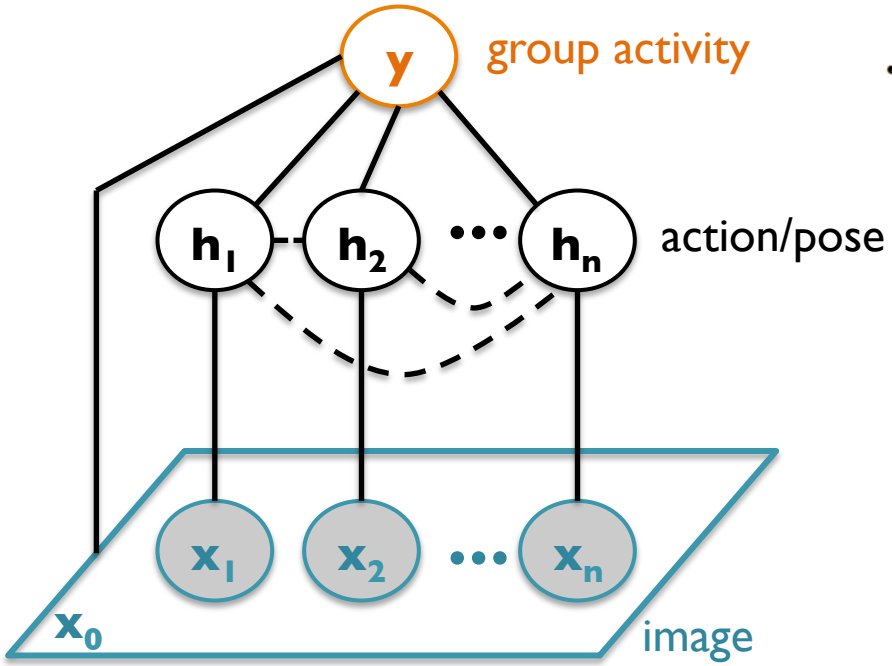
[Lan NIPS10]

Hierarchical Model For Group Activity Recognition



[Lan NIPS10]

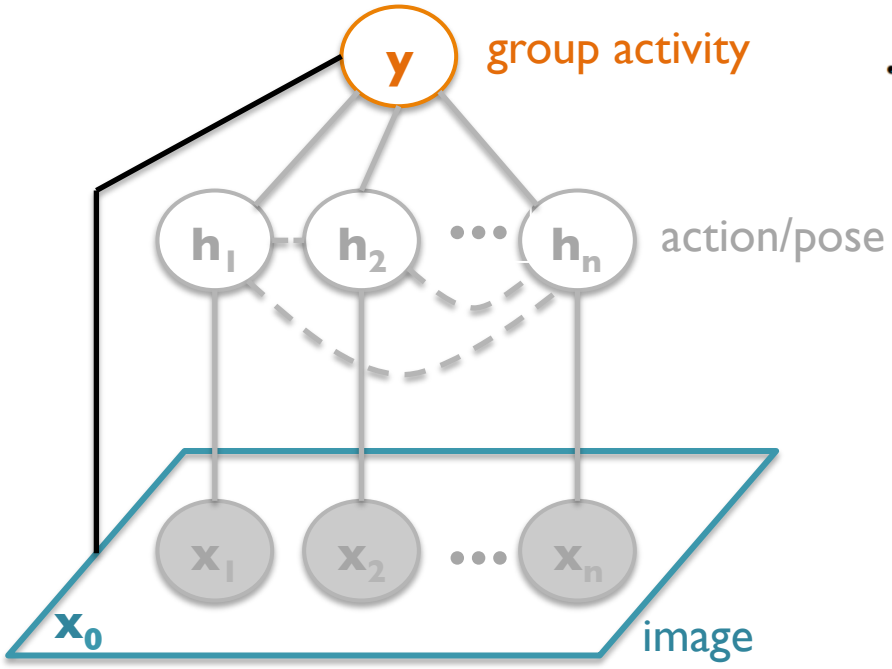
Hierarchical Model For Group Activity Recognition



$$
\begin{aligned}
f_w(\mathbf{x}, \mathbf{h}, y; \mathcal{G}) &= w^\top \Psi(y, \mathbf{h}, \mathbf{x}; \mathcal{G}) \\
&= w_0^\top \phi_0(y, x_0) \\
&+ \sum_{j \in \mathcal{V}} w_1^\top \phi_1(x_j, h_j) \\
&+ \sum_{j \in \mathcal{V}} w_2^\top \phi_2(y, h_j) \\
&+ \sum_{j,k \in \mathcal{E}} w_3^\top \phi_3(y, h_j, h_k)
\end{aligned}
$$

[Lan NIPS10]

Hierarchical Model For Group Activity Recognition



$$f_w(\mathbf{x}, \mathbf{h}, y; \mathcal{G}) = w^\top \Psi(y, \mathbf{h}, \mathbf{x}; \mathcal{G})$$

$$= w_0^\top \phi_0(y, x_0)$$

$$+ \sum_{j \in \mathcal{V}} w_1^\top \phi_1(x_j, h_j)$$

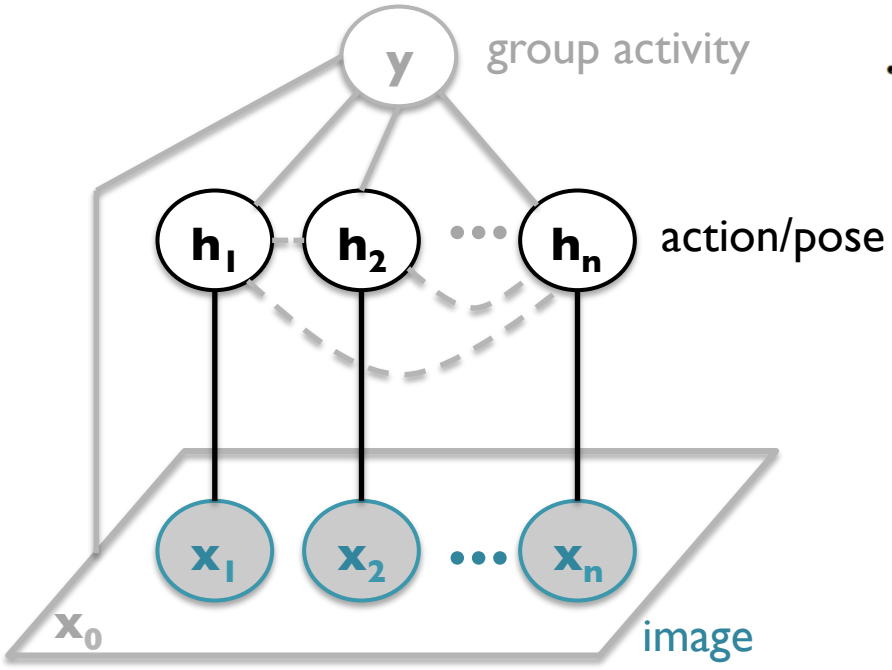
$$+ \sum_{j \in \mathcal{V}} w_2^\top \phi_2(y, h_j)$$

$$+ \sum_{j,k \in \mathcal{E}} w_3^\top \phi_3(y, h_j, h_k)$$

Direct Measurement for Groups.

[Lan NIPS10]

Hierarchical Model For Group Activity Recognition



$$f_w(\mathbf{x}, \mathbf{h}, y; \mathcal{G}) = w^\top \Psi(y, \mathbf{h}, \mathbf{x}; \mathcal{G})$$

$$= w_0^\top \phi_0(y, x_0)$$

$$+ \sum_{j \in \mathcal{V}} w_1^\top \phi_1(x_j, h_j)$$

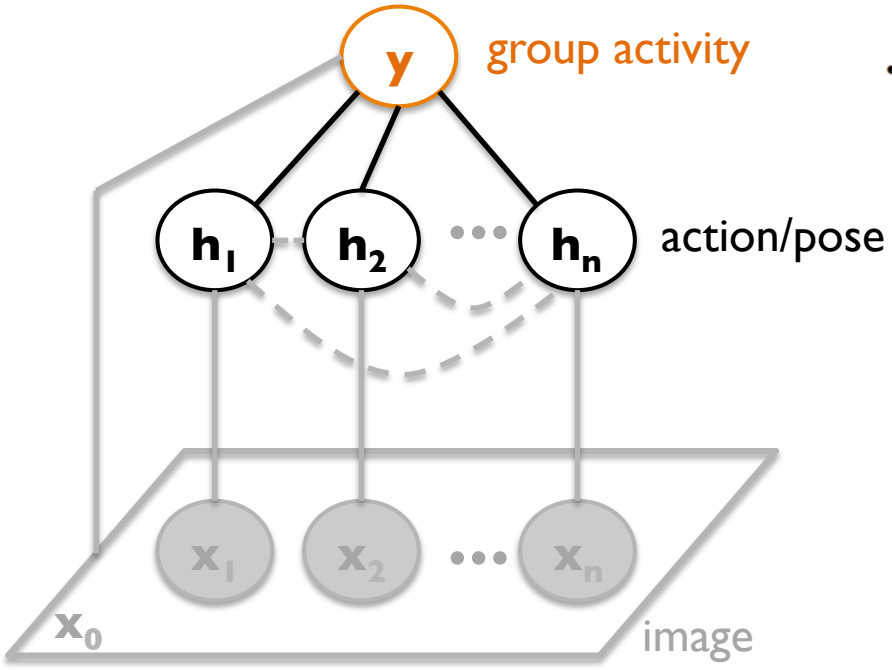
$$+ \sum_{j \in \mathcal{V}} w_2^\top \phi_2(y, h_j)$$

$$+ \sum_{j,k \in \mathcal{E}} w_3^\top \phi_3(y, h_j, h_k)$$

Measurements for Action/Pose.

[Lan NIPS10]

Hierarchical Model For Group Activity Recognition



$$f_w(\mathbf{x}, \mathbf{h}, y; \mathcal{G}) = w^\top \Psi(y, \mathbf{h}, \mathbf{x}; \mathcal{G})$$

$$= w_0^\top \phi_0(y, x_0)$$

$$+ \sum_{j \in \mathcal{V}} w_1^\top \phi_1(x_j, h_j)$$

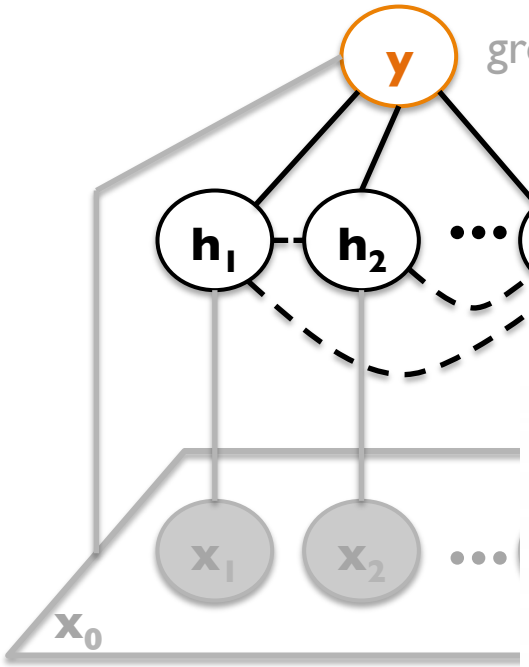
$$+ \sum_{j \in \mathcal{V}} w_2^\top \phi_2(y, h_j)$$

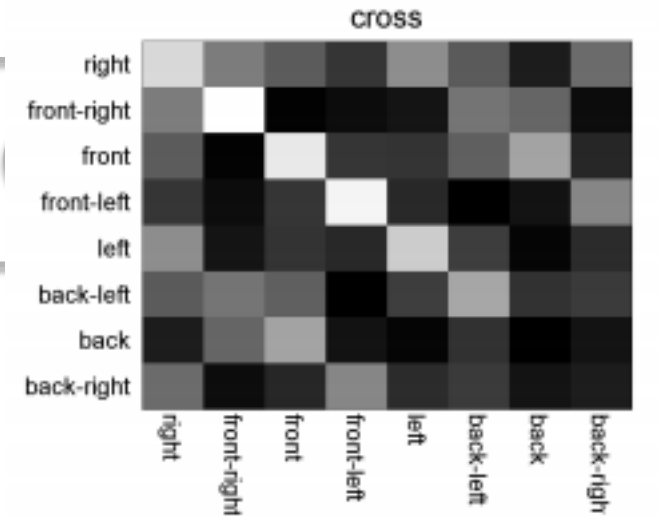
$$+ \sum_{j,k \in \mathcal{E}} w_3^\top \phi_3(y, h_j, h_k)$$

Group-Individual Relationship via co-occurrence.

[Lan NIPS10]

Hierarchical Model For Group Activity Recognition





$$f_w(\mathbf{x}, \mathbf{h}, y; \mathcal{G}) = w^\top \Psi(y, \mathbf{h}, \mathbf{x}; \mathcal{G})$$

$$= w_0^\top \phi_0(y, x_0)$$

$$+ \sum_{j \in \mathcal{V}} w_1^\top \phi_1(x_j, h_j)$$

$$+ \sum_{j \in \mathcal{V}} w_2^\top \phi_2(y, h_j)$$

$$+ \sum_{j,k \in \mathcal{E}} w_3^\top \phi_3(y, h_j, h_k)$$

Group Dependent
Pairwise Individual Relationship
via co-occurrence.

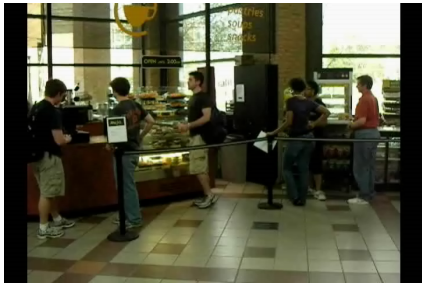
[Lan NIPS10]

Dataset

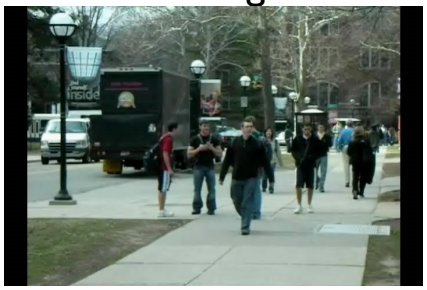
- Collective Activity Dataset
  - Available at http://wwweb.eecs.umich.edu/vision/activity-dataset.html
  - 44 videos with multiple people
  - Crossing, Waiting, Queuing, Walking, Talking

Crossing

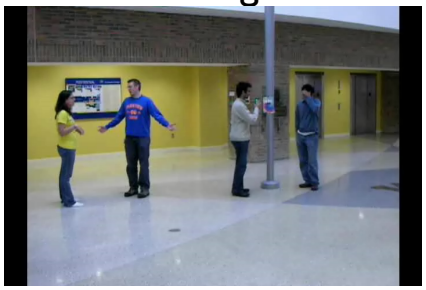
Waiting



Queuing



Walking

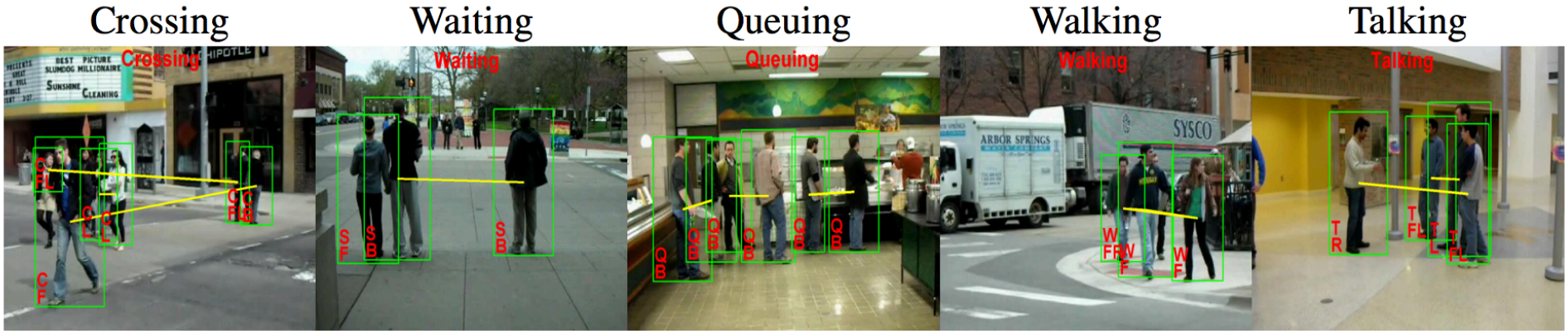


Talking

Choi et al, 2009

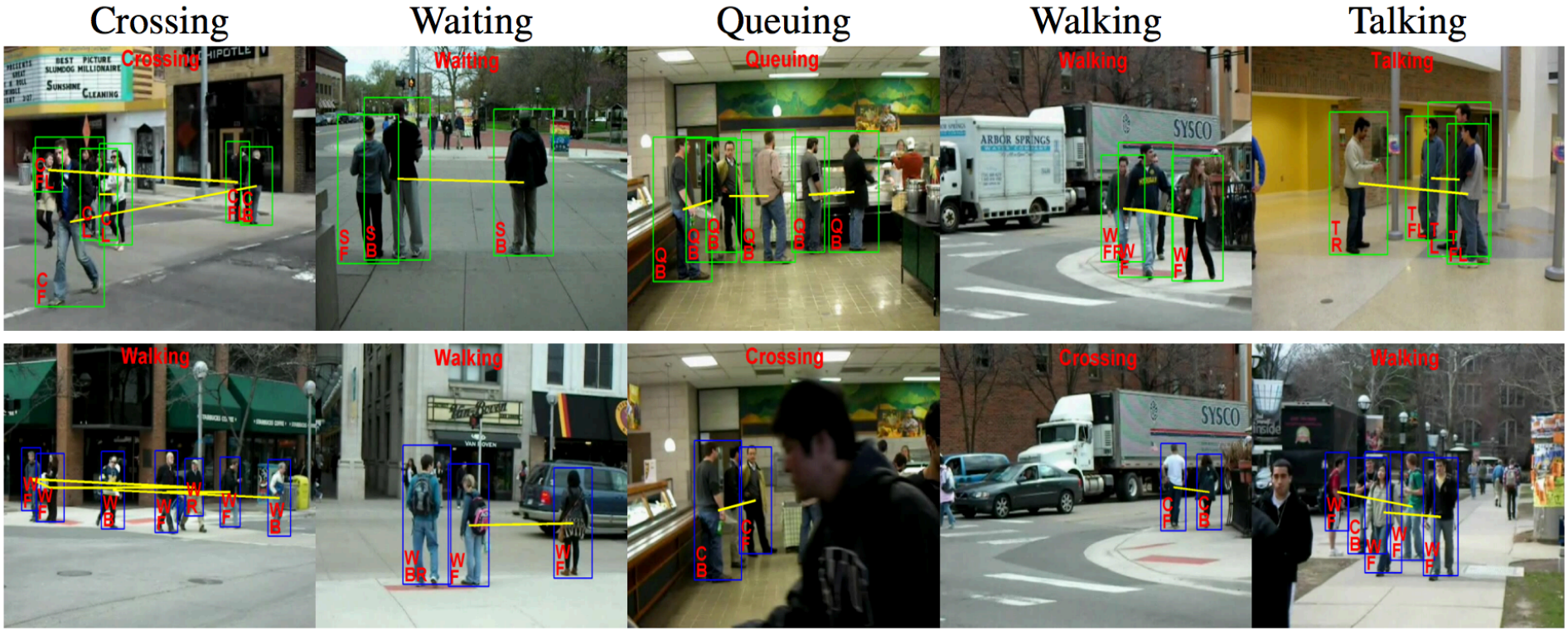
[Lan NIPS10]

Qualitative Examples



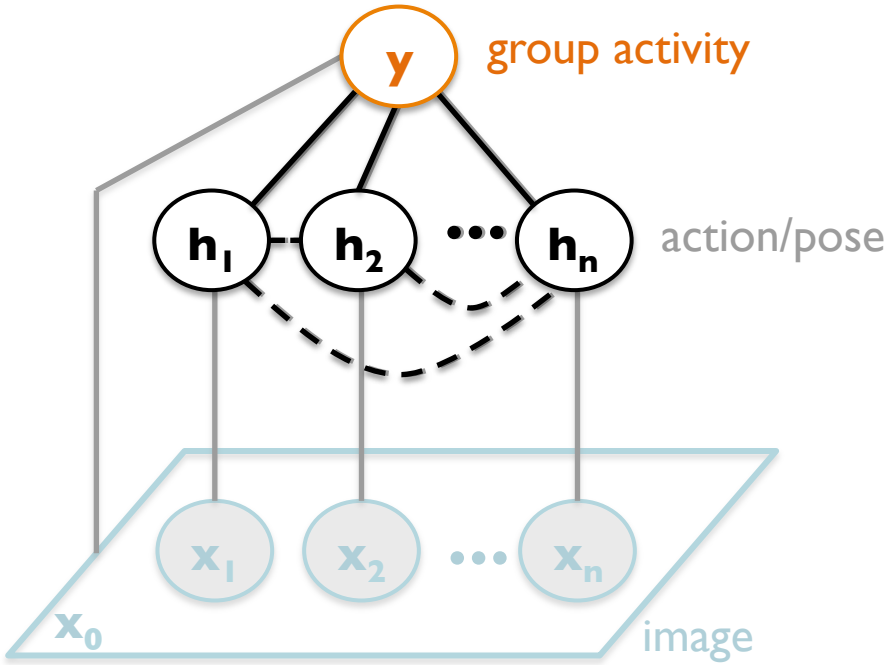
[Lan NIPS10]

Qualitative Examples



[Lan NIPS10]

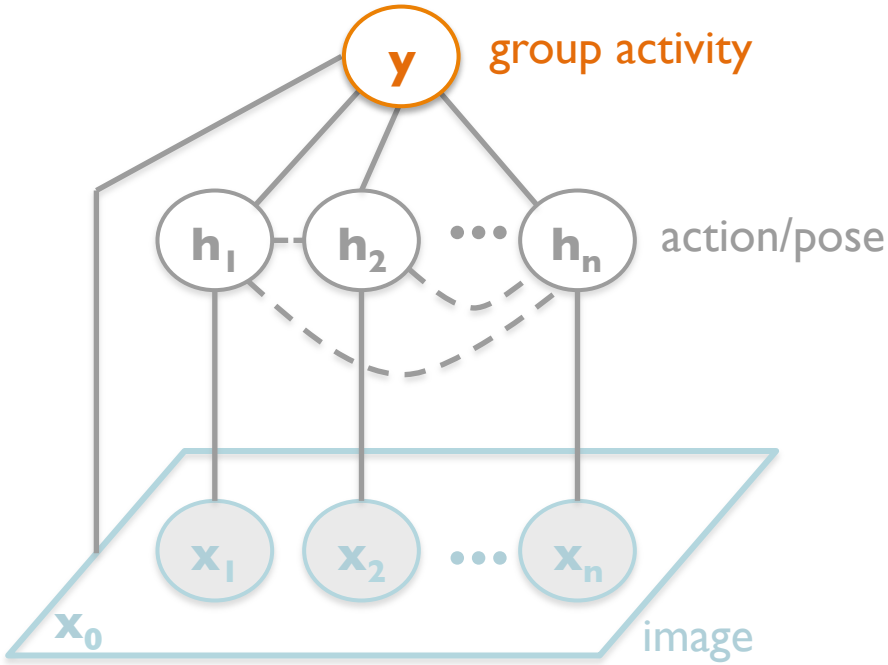
Limitations



No spatial relationship

[Lan NIPS10]

Limitations



No spatial relationship

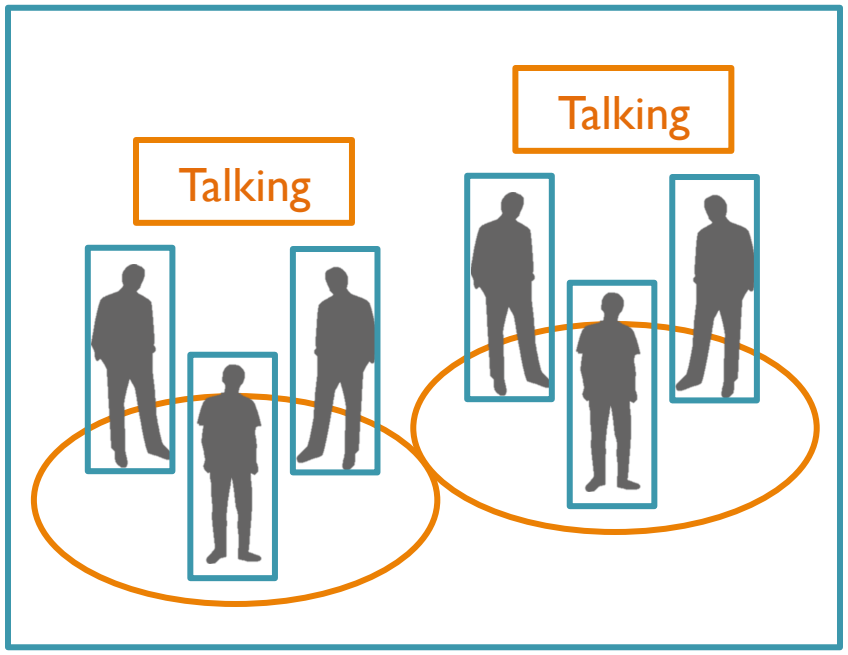
One group assumption

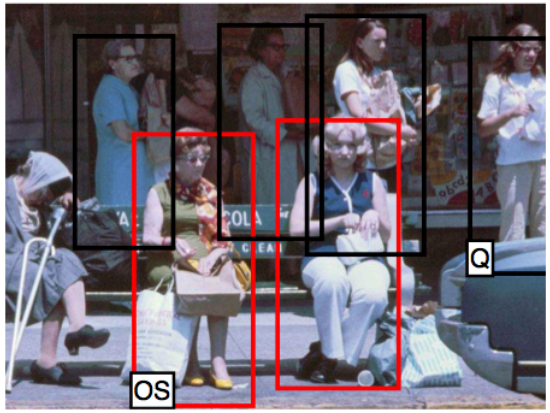
[Choi ECCV14]

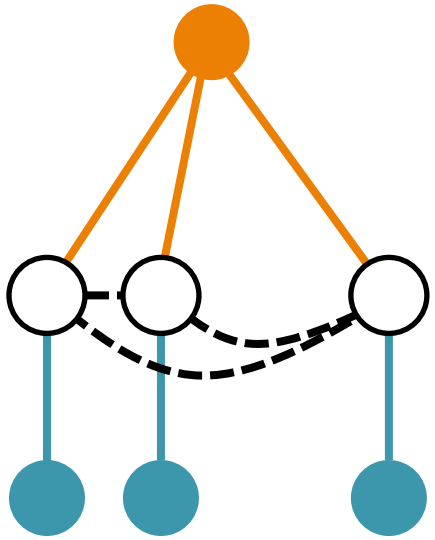
Group Discovery (Choi ECCV14)

Input: image and bounding boxes

Output: clustered groups with activity label



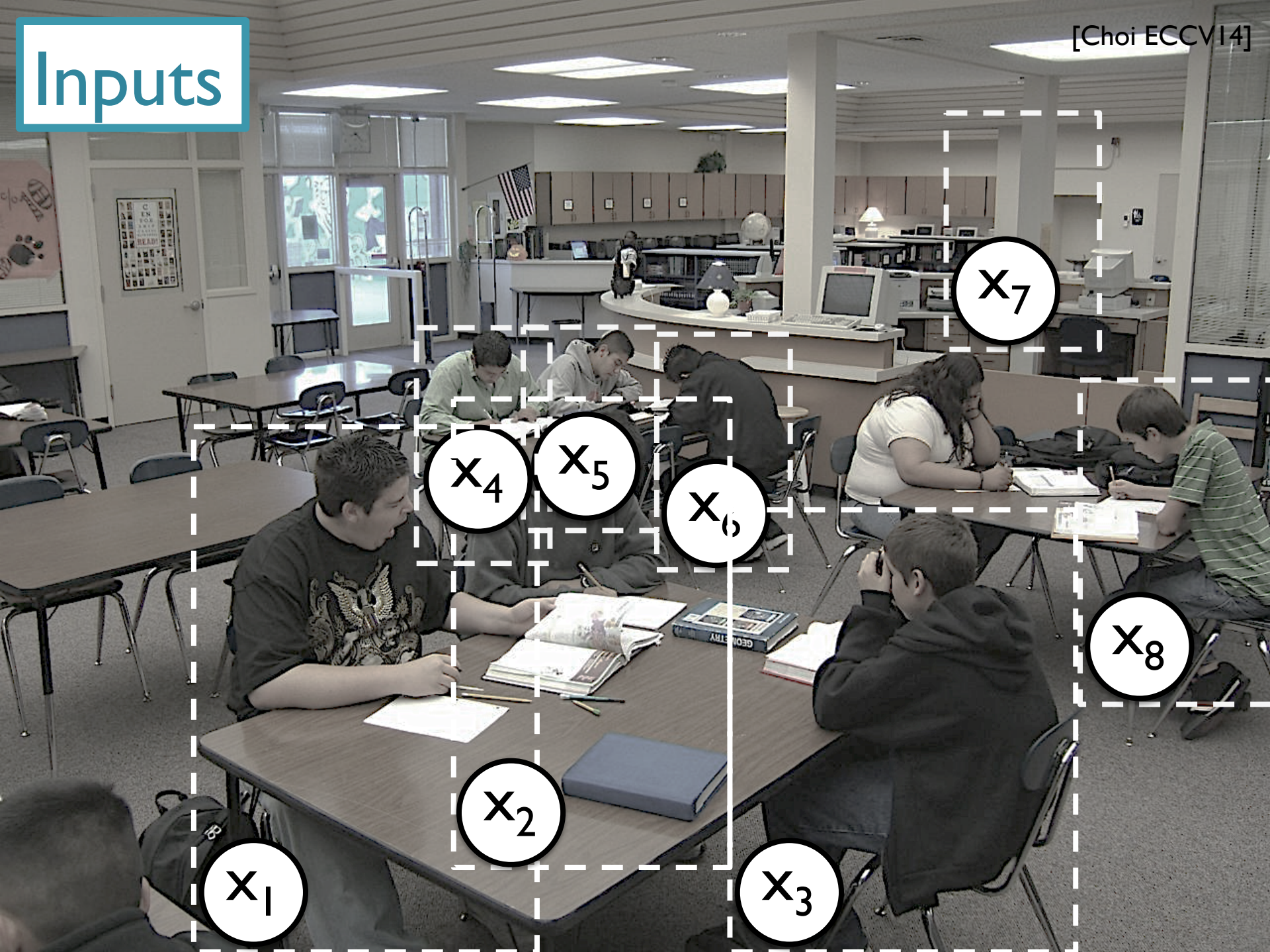





[Choi ECCV14]
GEOMETRY

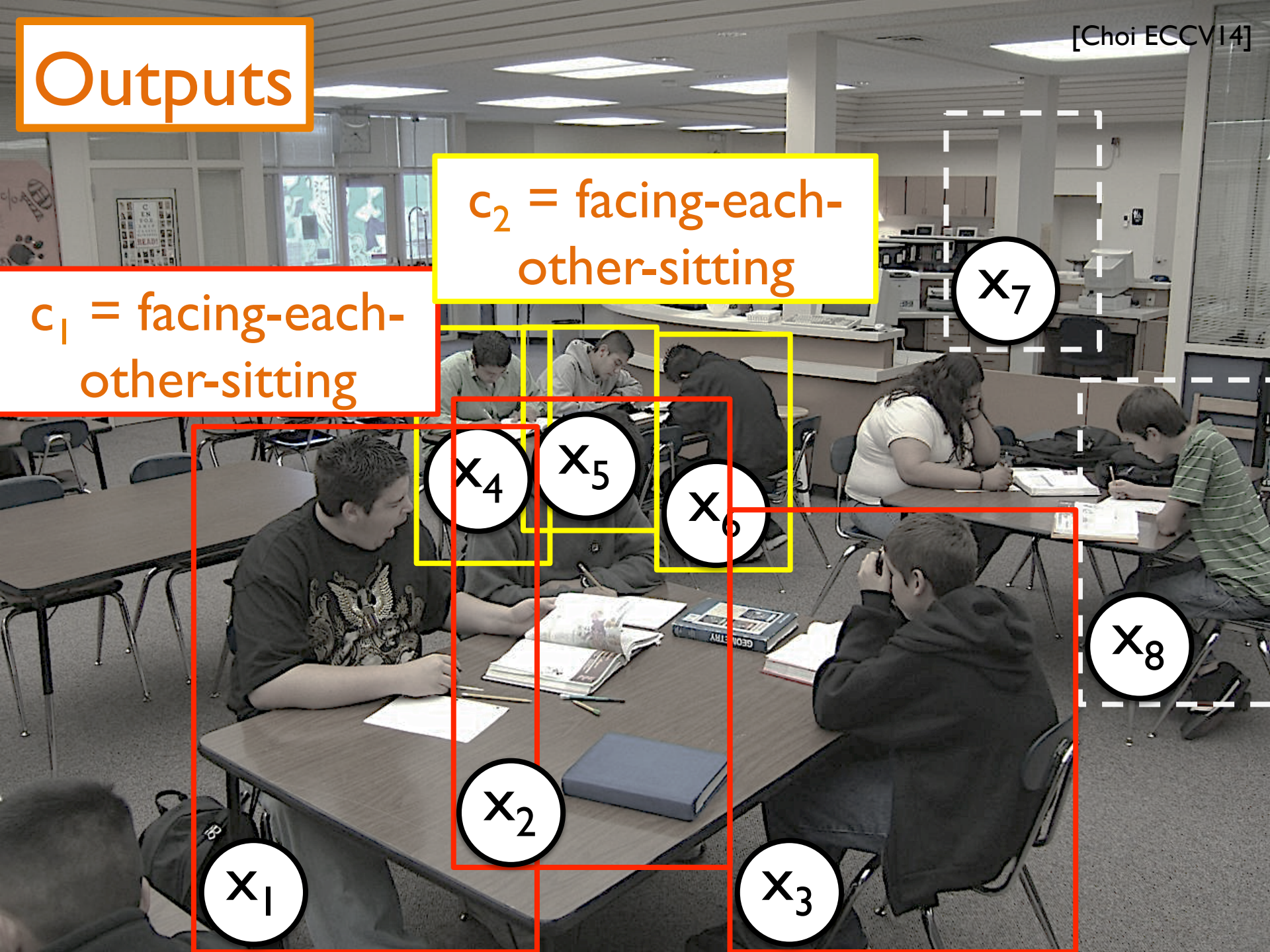
Inputs

[Choi ECCV14]



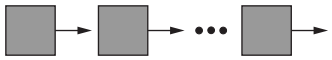
Outputs

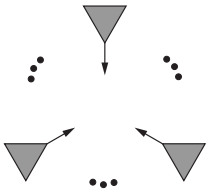
[Choi ECCV14]

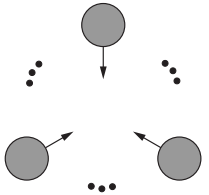


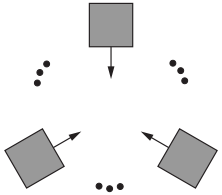
[Choi ECCV14]

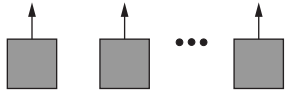
Structured Groups

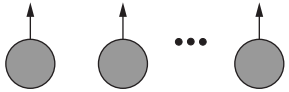


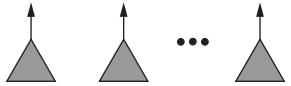






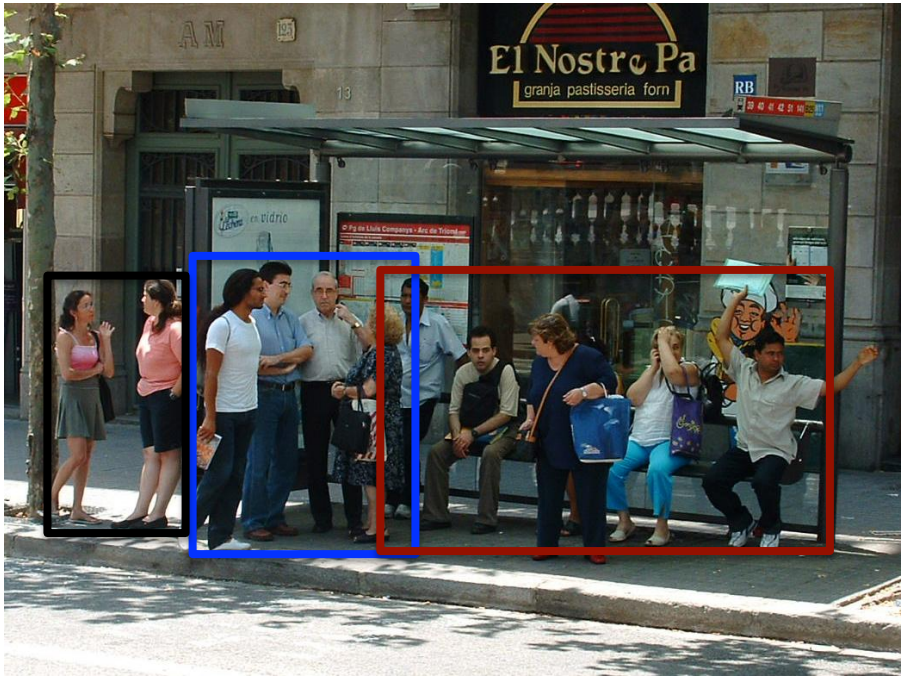






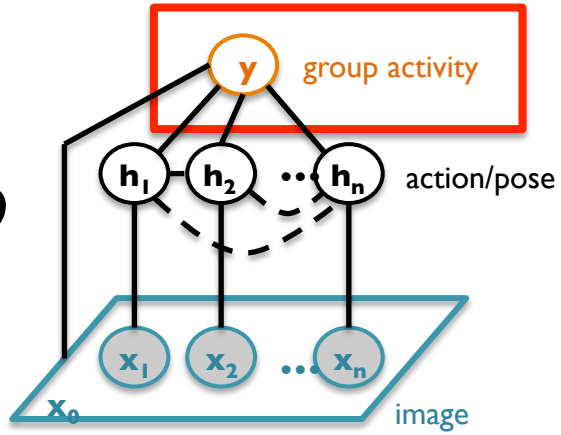
[Choi ECCV14]

Challenges



- Unknown number of groups.

?



[Choi ECCV14]

Challenges

- Unknown number of groups.
- Large intra-class variation.

[Choi ECCV14]

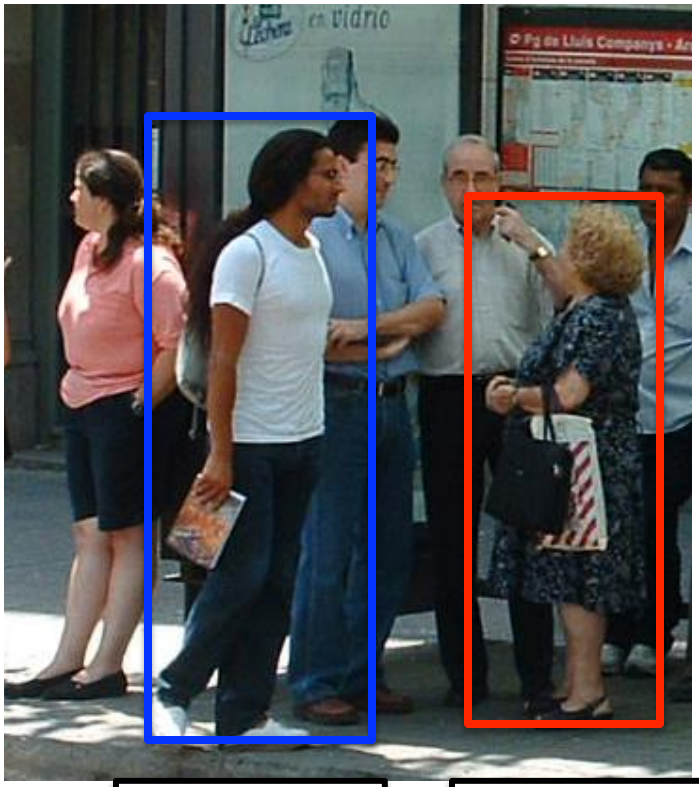
Interactions as Key Social Signal

Standing
Facing-right

- Individual Pose
  - Weak social signal.

[Choi ECCV14]

Interactions as Key Social Signal



- Individual Pose
  - Weak social signal.
- Pair Interaction
  - Strong social signal within a group.

[Choi ECCV14]

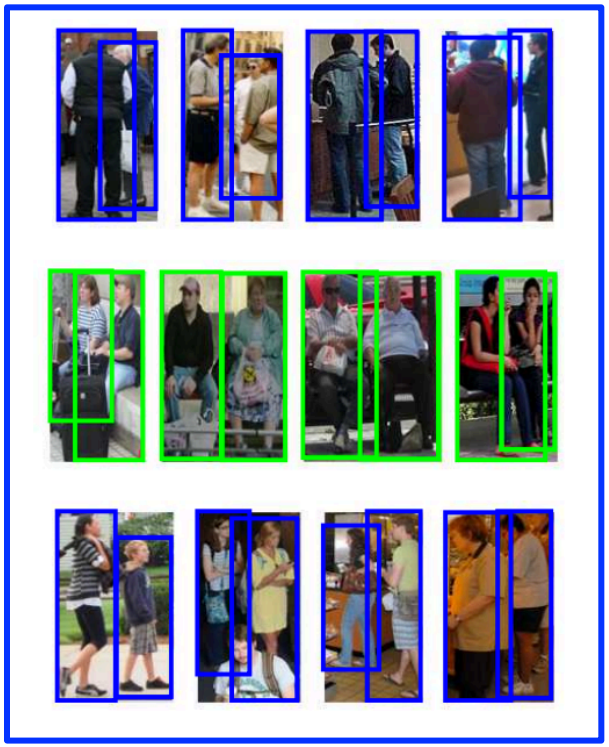
Interactions as Key Social Signal

- Individual Pose
  - Weak social signal.
- Pair Interaction
  - Strong social signal within a group.
  - Repulsive signal in different groups.

[Choi ECCV14]

Interactions as Key Social Signal

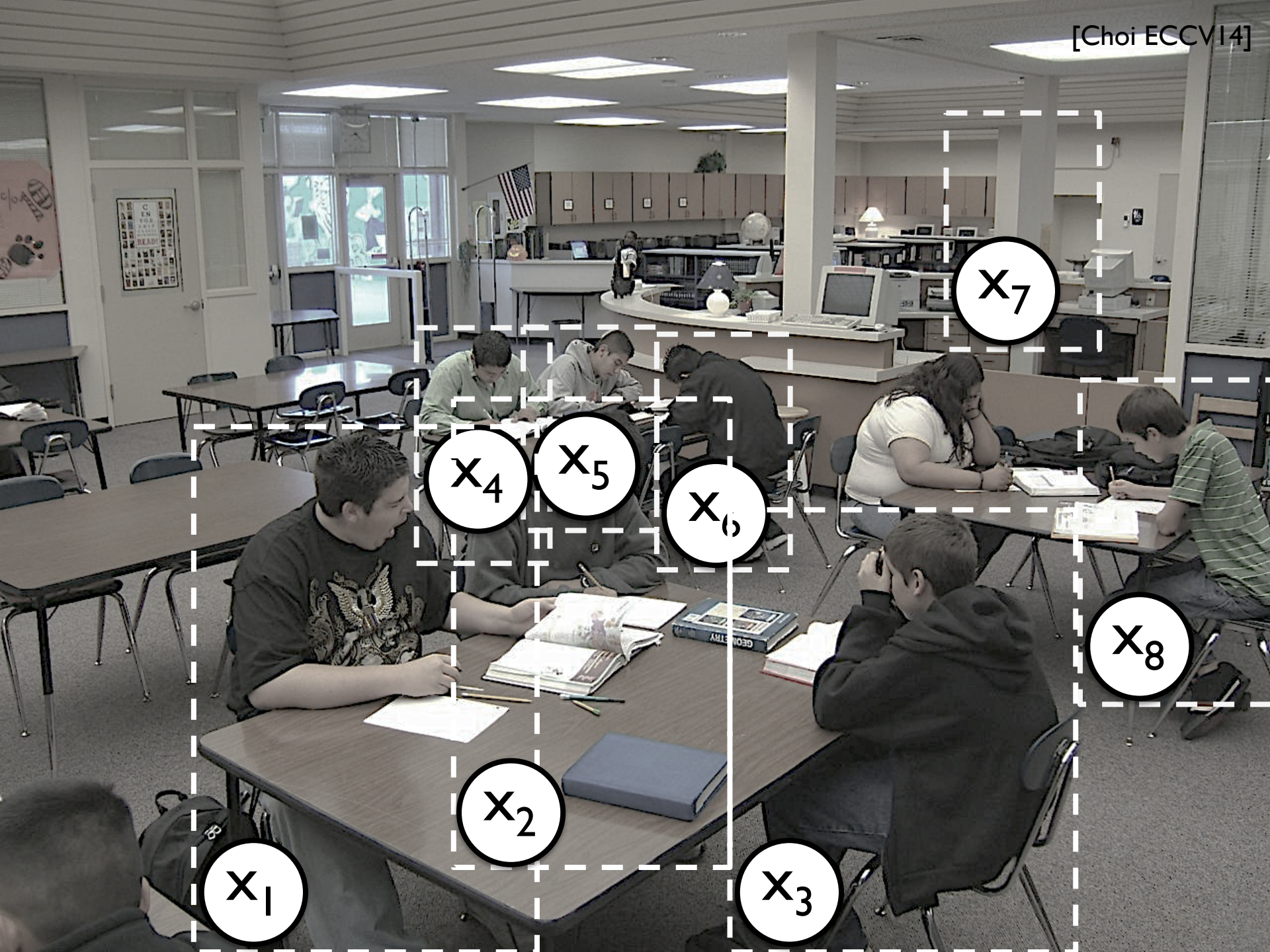
Intra-group Interactions



Attractive Potential for a Group

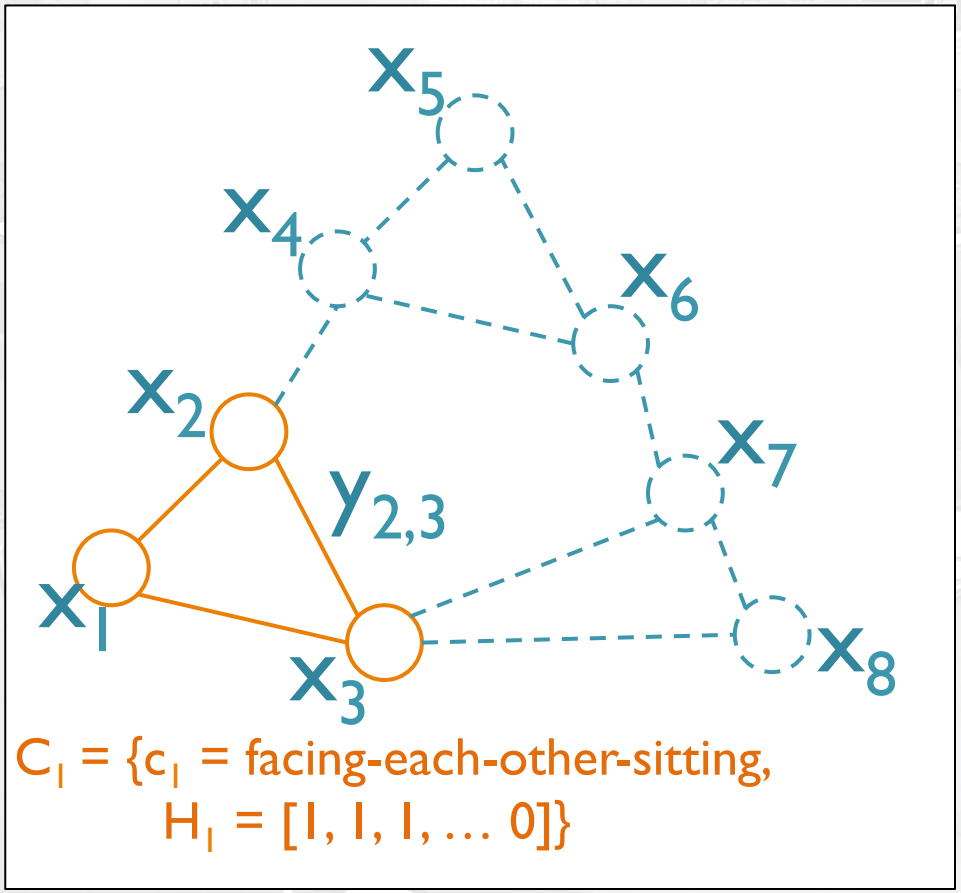
Inter-group Interactions

Repulsive Potential for a Group


[Choi ECCV14]
x7
x4
x5
x6
x8
x2
x1
x3

[Choi ECCV14]

Group Discovery Model



Nodes

- Individual likelihood to belong in a group C.

Edges

- Likelihood of interaction in a group C.

Solid lines

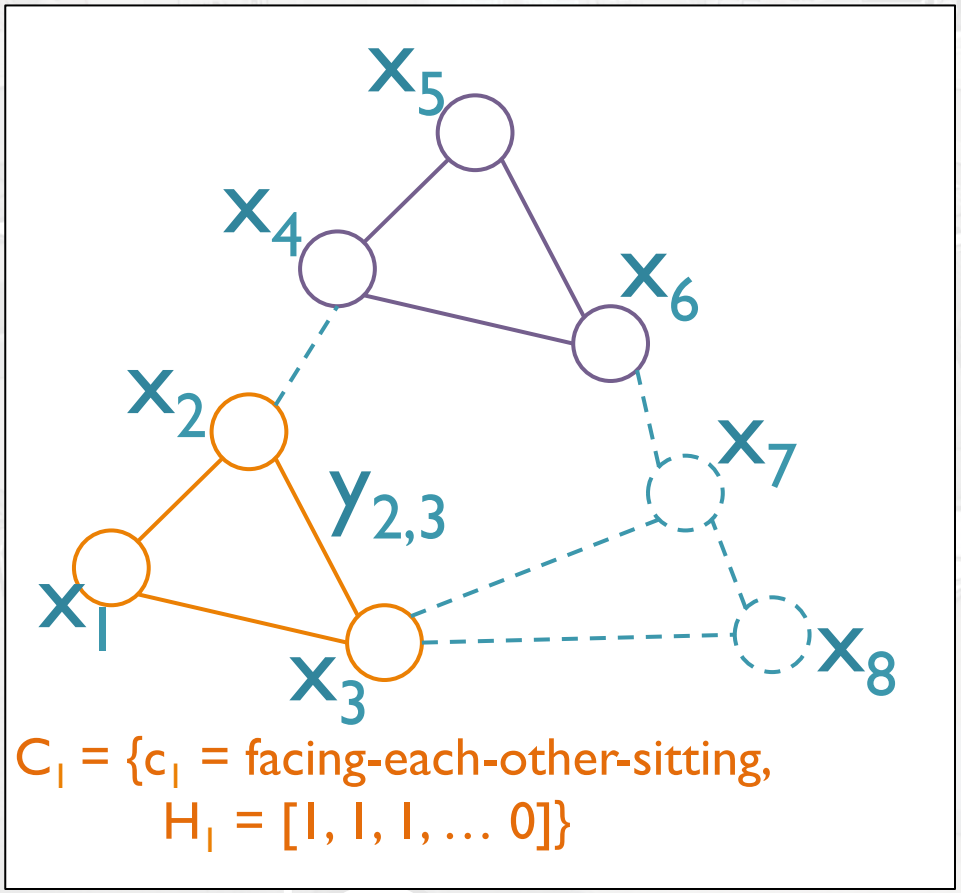
- Members belonging to the group.

Dashed lines

- Not belonging to the group.

[Choi ECCV14]

Group Discovery Model



Nodes

- Individual likelihood to belong in a group C.

Edges

- Likelihood of interaction in a group C.

Solid lines

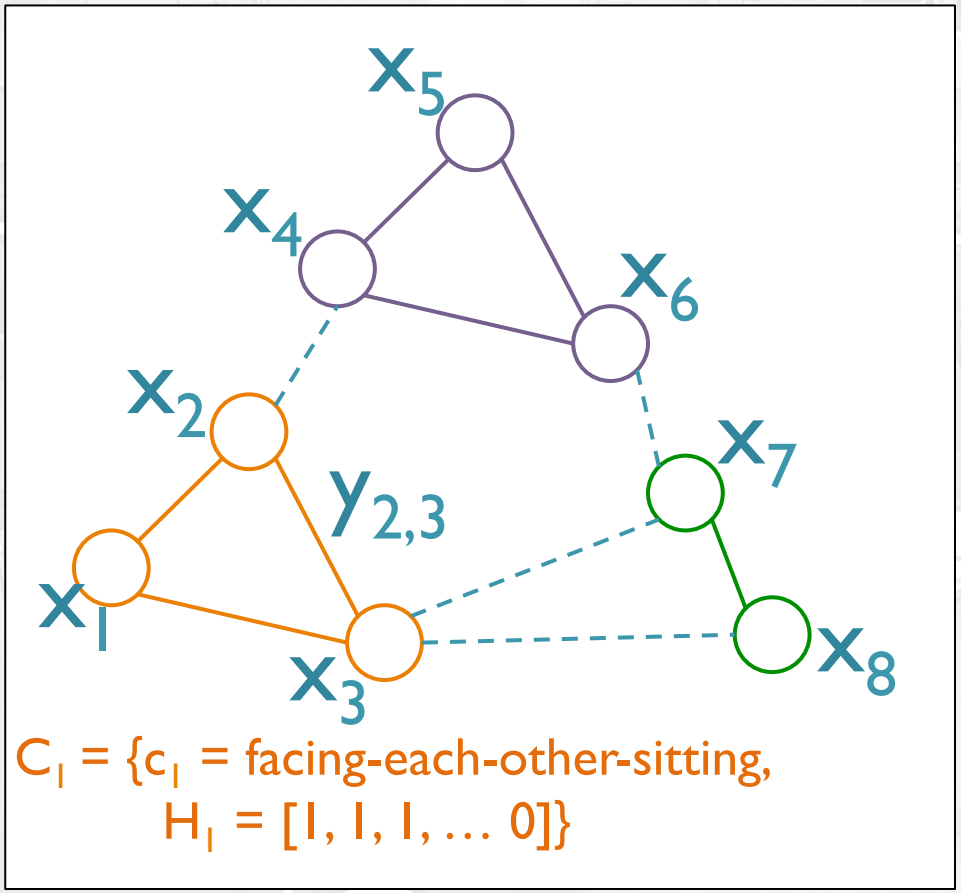
- Members belonging to the group.

Dashed lines

- Not belonging to the group.

[Choi ECCV14]

Group Discovery Model



Nodes

- Individual likelihood to belong in a group C.

Edges

- Likelihood of interaction in a group C.

Solid lines

- Members belonging to the group.

Dashed lines

- Not belonging to the group.

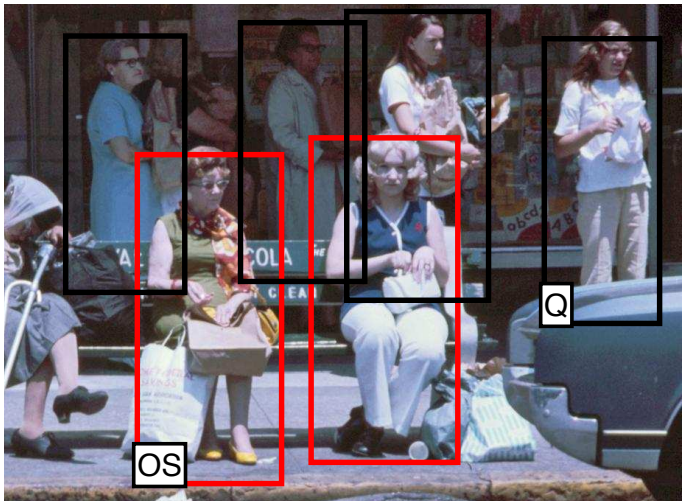
[Choi ECCV14]

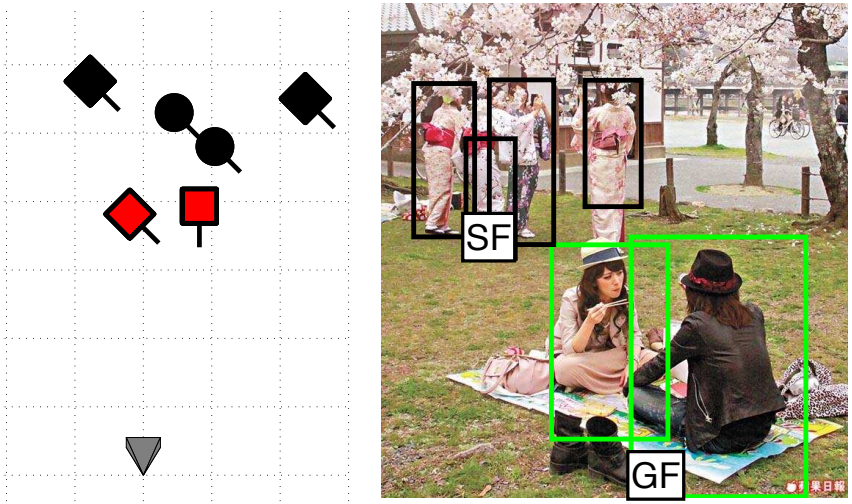
Datasets

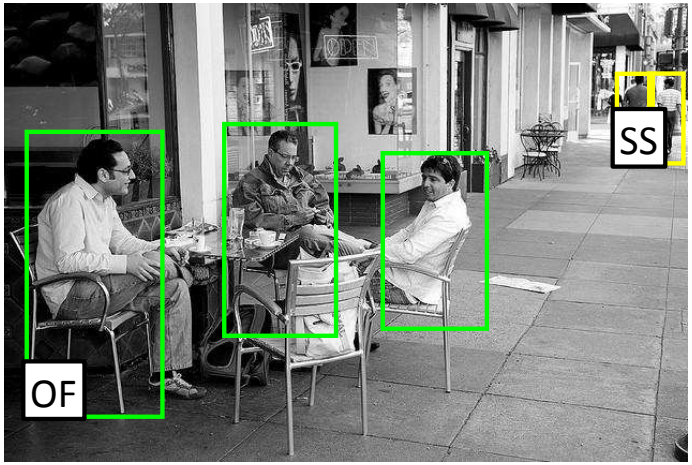
- Structured Group Dataset
  - Available at http://cvgl.stanford.edu/projects/groupdiscovery/
  - 588 images with 5,415 individuals and 1,719 groups.
  - 7 structured groups:
    - Queue, Standing-Facing, Sitting-Object-Facing, Sitting-Ground-Facing, Standing-Side, Sitting-Object-Side, Sitting-Ground-Side.

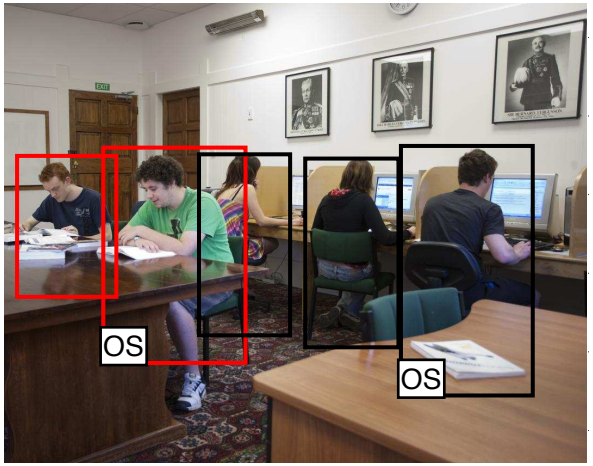
[Choi ECCV14]

Group Discovery Results

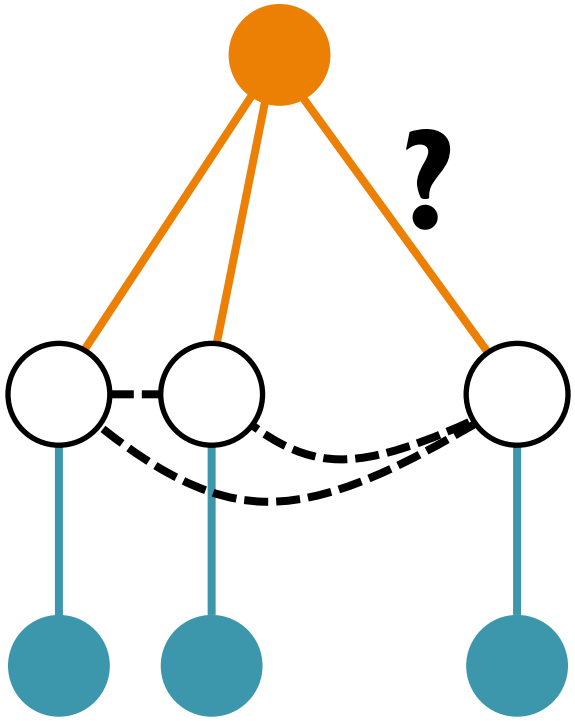








Learning the Models



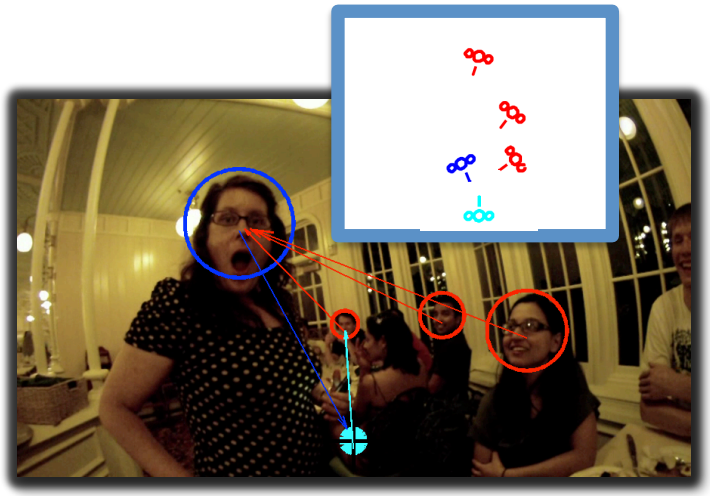
Model weights can be learned in a Max-Margin framework using Structural SVM.

$$\min_{\mathbf{w},\boldsymbol{\xi}} \quad \frac{1}{2}\|\mathbf{w}\|^2 + \frac{C}{n}\sum_{i=1}^{n}\xi_i, \;\; \text{s.t.} \;\; \forall i, \xi_i \geq 0$$

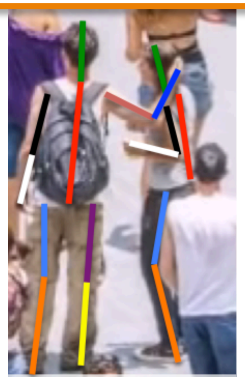
$$\forall i, \; \forall \mathbf{y} \in \mathcal{Y} \setminus \mathbf{y}_i : \; \langle \mathbf{w}, \delta\Psi_i(\mathbf{y}) \rangle \geq \triangle(\mathbf{y}_i, \mathbf{y}) - \xi_i$$

Tsochantaridis et al, 2004

Signals for Social Statics



The left touches the right's head.



Conversation