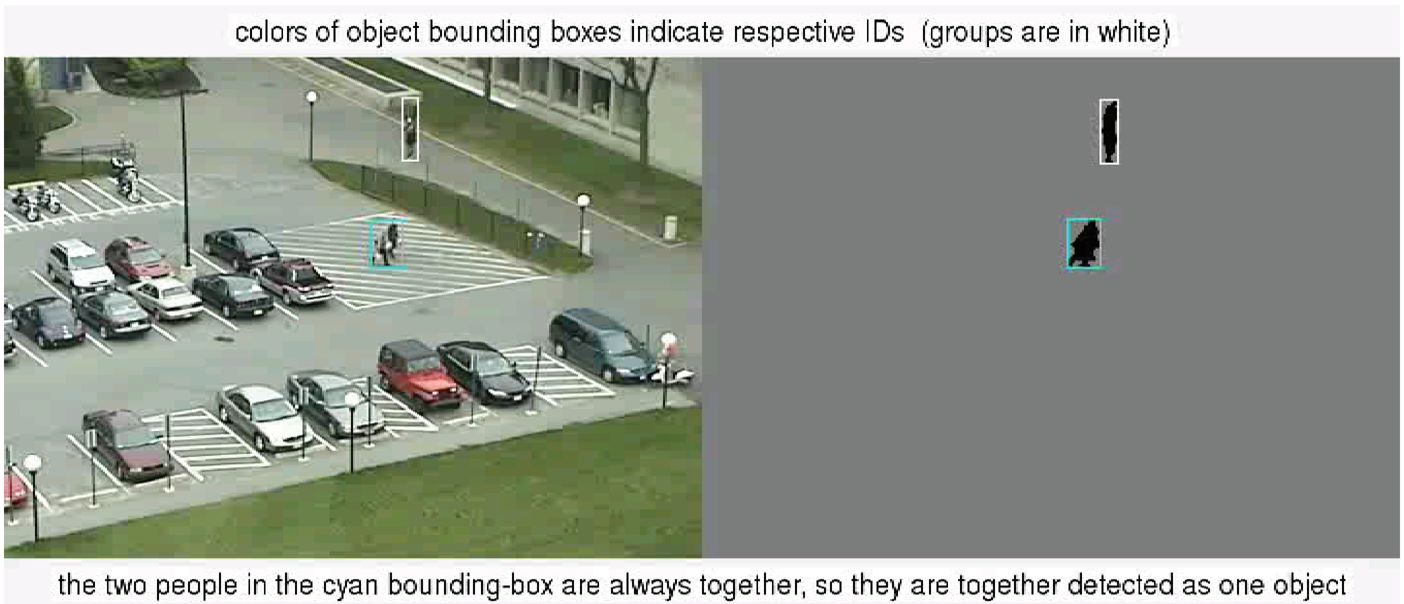
Behavior Analysis in Crowded Environments

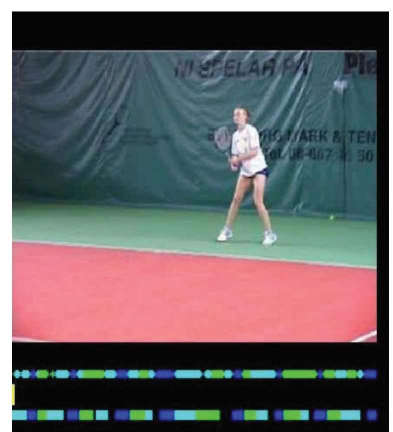
Xiaogang Wang
Department of Electronic Engineering
The Chinese University of Hong Kong

June 25, 2011

Behavior Analysis in Sparse Scenes

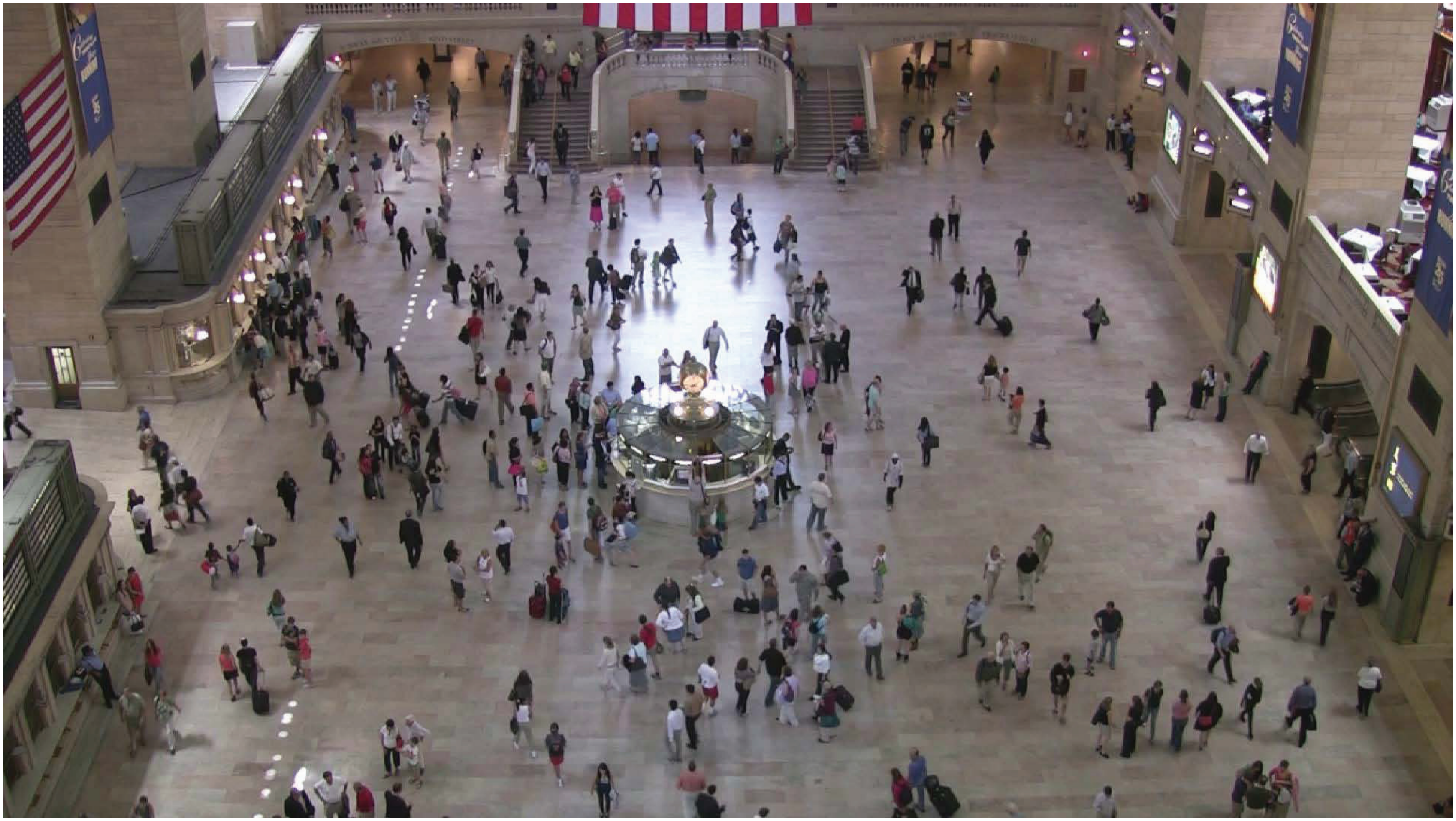


Zelnik-Manor & Irani CVPR'04



Crowded Environments

Crowded Environments

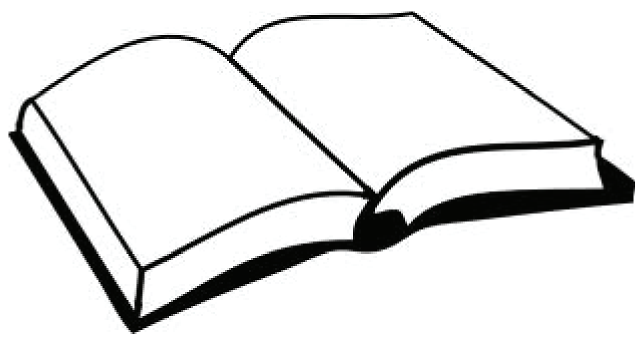


Outline

- Introduction
  - Why is behavior analysis in crowded environment interesting?
  - Major challenges
- Behavior analysis under hierarchical Bayesian models
  - Based local motions
  - Based noisy tracklets
- Other works
- Conclusions and future work

Crowd

"The crowd, an agglomeration of people, presents new characteristics very different from those of the individuals composing it, the sentiments and ideas of all the persons in the gathering take one and the same direction, and their conscious personality vanishes." -- by Le Bon (1841~1931) in "The Crowd: A Study of the Popular Mind"



Why is behavior analysis in crowded environment interesting?

- Many places of security interest are crowded

Train station

Shopping mall

Airport

Street intersection

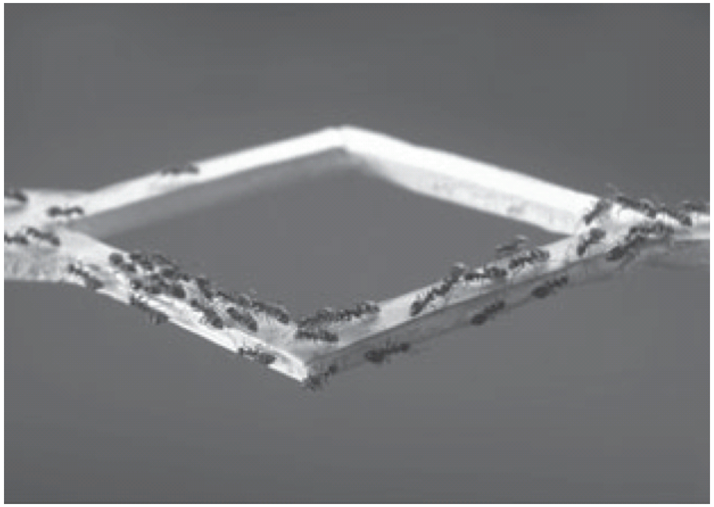
- Crowd control
- Providing guidelines for planning and designing crowded areas

Why is behavior analysis in crowded environment interesting?

- The study of pedestrian crowds is an interesting subject of social research
  - Pedestrian behaviors in crowd present new characteristics than individual personalities
  - Self-organization of collective behavior patterns due to nonlinear interactions among pedestrians
  - Variables of pedestrian motions are measurable: leading to a deep insight of other social processes, such opinion formation

Why is behavior analysis in crowded environment interesting?

- The self-organization phenomena are observed in other fields
  - At medium and high pedestrian densities, the motion of pedestrians shows striking analogies with the motion of gases and fluids [Helbing *Statistical Mechanics*'97]
  - The self-organization of collective behaviors are also observed in animal groups [Moussaid et al. *Topics in Cognitive Science*'09]



Interdisciplinary subject

- Statistical physics: understanding the fundamental mechanism of forming collective behaviors from interactions of individuals
- Computer graphics: simulating crowd behaviors
- Computer vision:
  - Learning and detecting collective motion patterns
  - Temporally segmenting the video sequences into different global crowd behaviors
  - Detecting abnormal behaviors

Formation of collective motion

- Externally planned or organized
- Self-organization from nonlinear interactions of pedestrians

Major challenges

- Crowded: it is difficult to do detection and tracking due to frequent occlusions and scene clutters
- Complex: many different types of behaviors happen together

Crowded and complex

Crowded but relatively simple

Outline

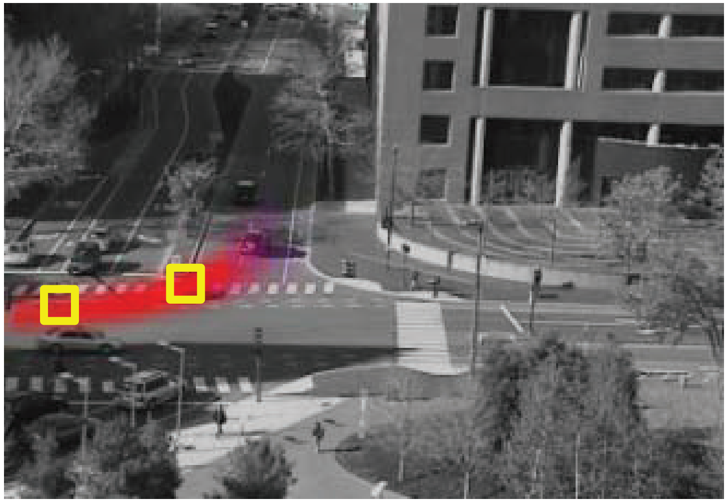
- Introduction
  - ❖ Why is behavior analysis in crowded environment interesting?
  - ❖ Major challenges
- Behavior analysis under hierarchical Bayesian models
  - ❖ Based local motions
  - ❖ Based noisy tracklets
- Other works
- Conclusions

X. Wang, X. Ma, and E. Grimson, "Unsupervised Activity Perception in Crowded and Complicated Scenes Using Hierarchical Bayesian Models," *IEEE Trans. on PAMI*, Vol. 31, 539-555, 2009.

X. Wang, X. Ma, and E. Grimson, "Unsupervised Activity Perception by Hierarchical Bayesian Models," *CVPR 2007*.

Learning motion patterns without tracking

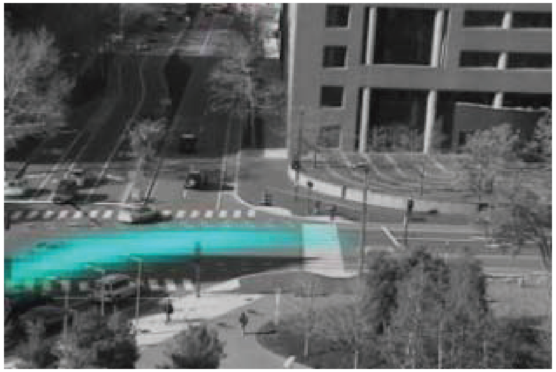
- Motion patterns are the pathways of moving objects
- Learning motions patterns from the temporal co-occurrence of moving pixels



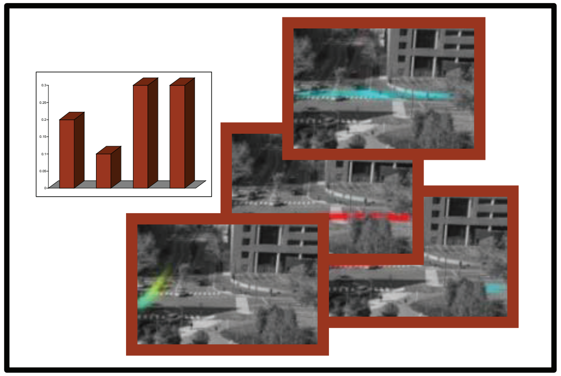
Our tasks

- Cluster moving pixels into atomic activities and segment a video sequence into global behaviors in crowded scenes

Example of data



Model of an atomic activity

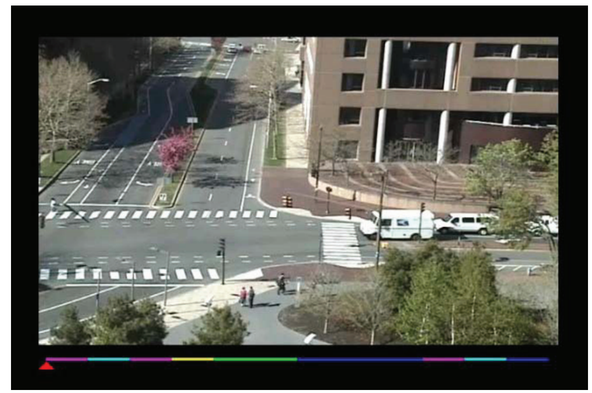


Model of an global behavior

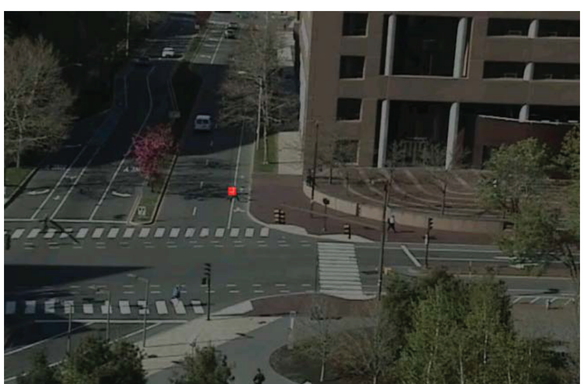
Our tasks

- Cluster moving pixels into atomic activities and segment a video sequence into global behaviors in crowded scenes

Motion segmentation



Video segmentation



Abnormality detection

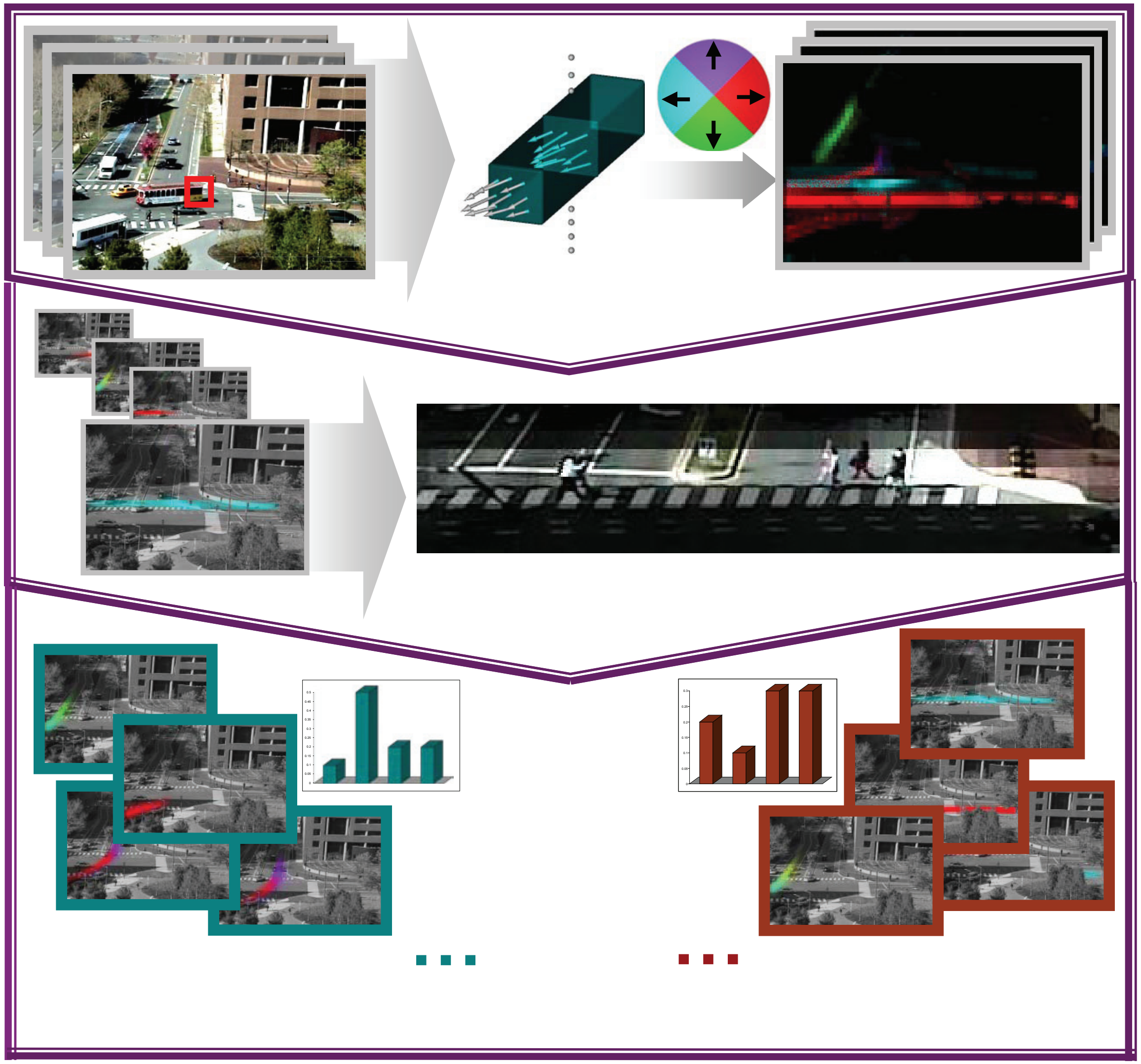
Interaction query

High level picture of our approach

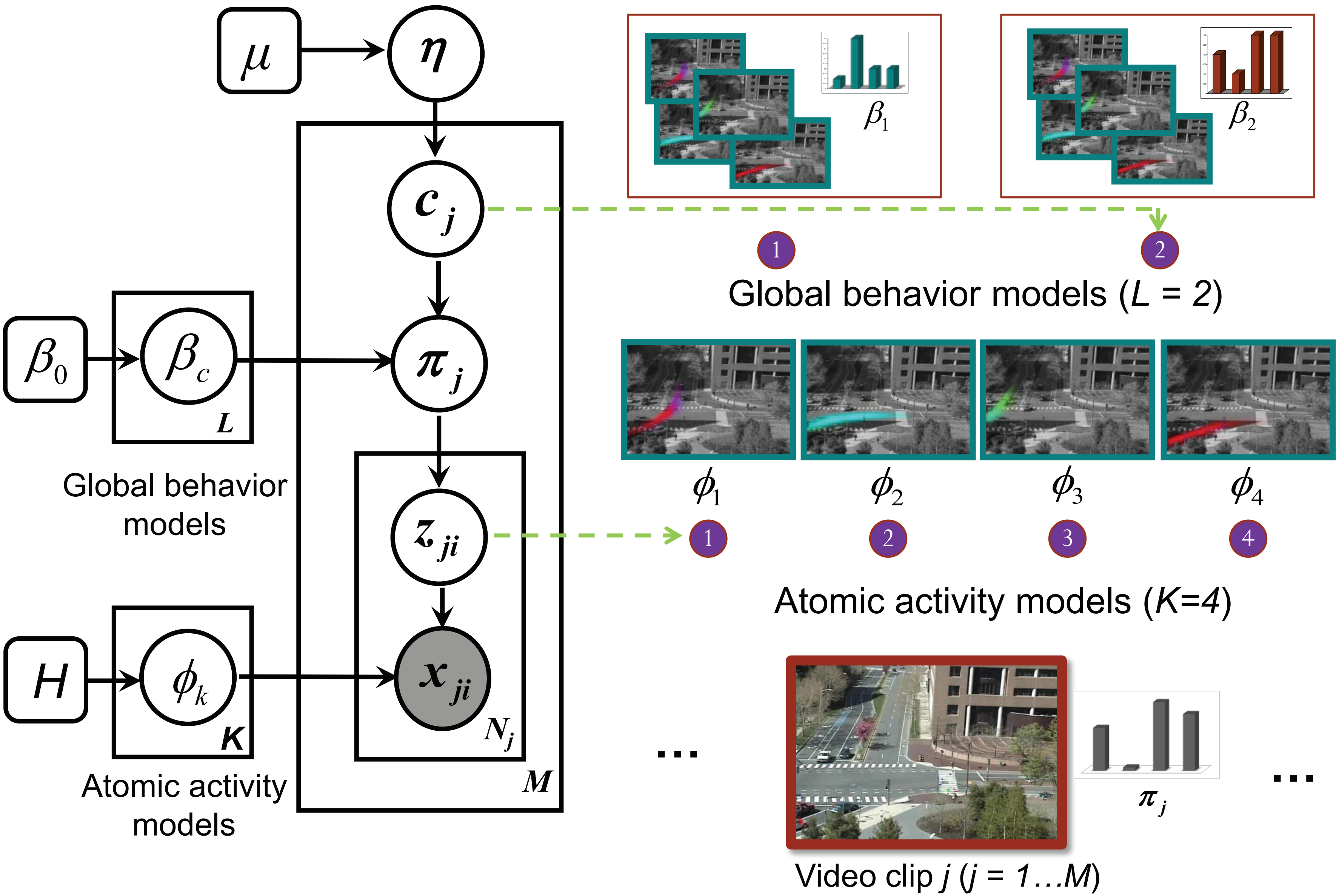
Motion Features
(a)

Atomic activities
modeled as distributions over the feature codebook
(b)

Global behaviors
modeled as distributions over atomic activities
(c)

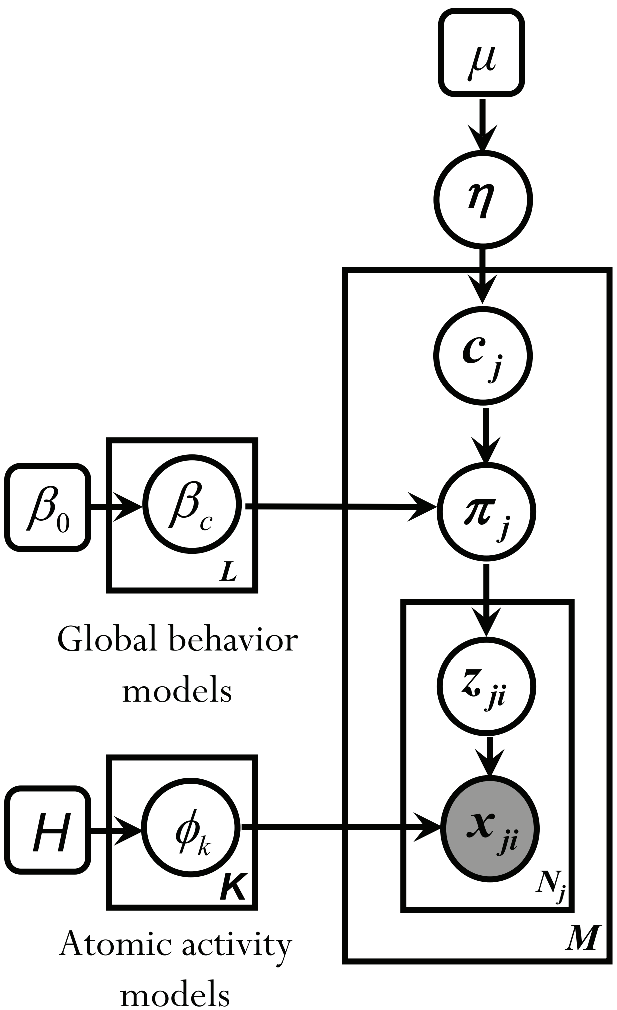


Parametric hierarchical Bayesian model

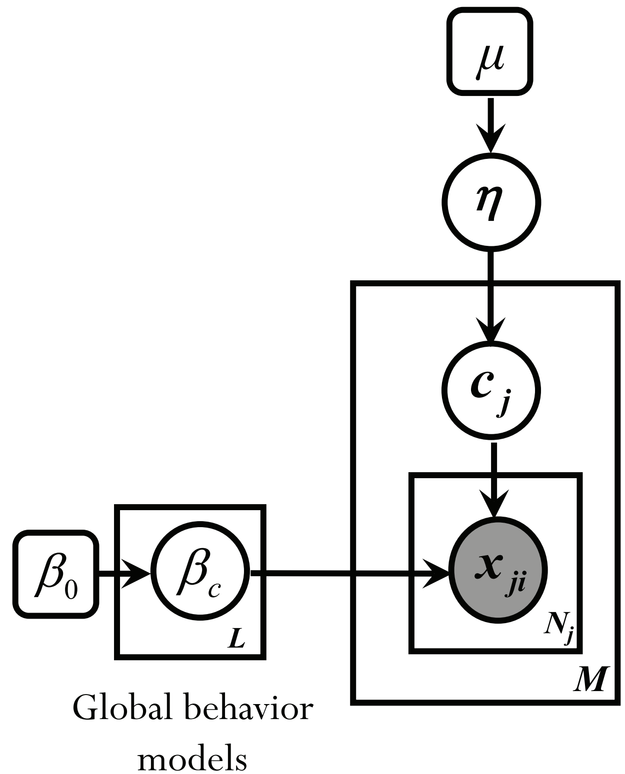


Advantages of this hierarchical Bayesian model

- More compact representation of video clips on the top of atomic activities
  - Number of atomic activity categories (29) versus size of feature codebook (13,824)



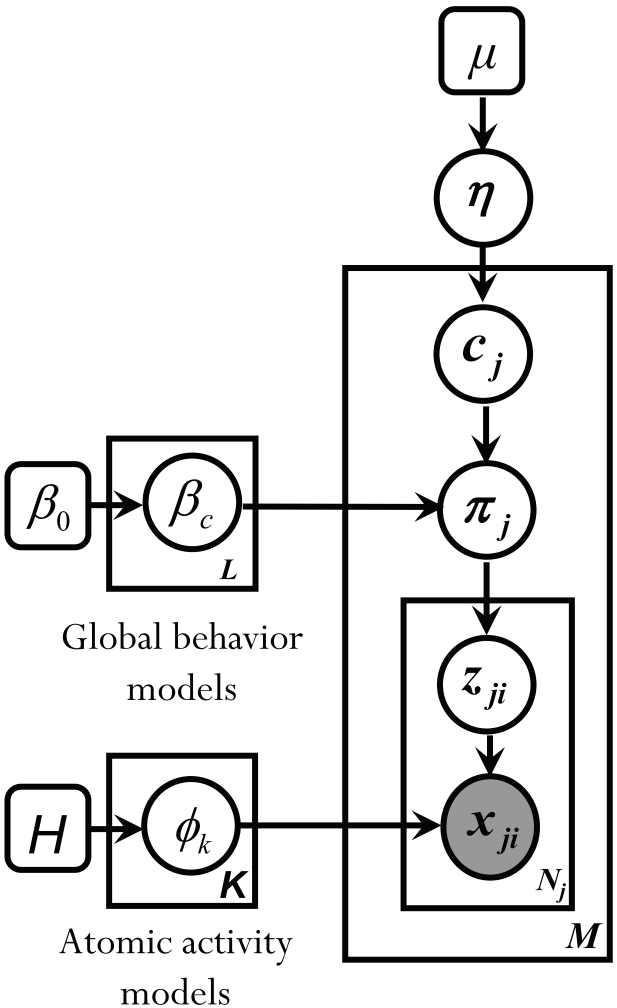
Our model



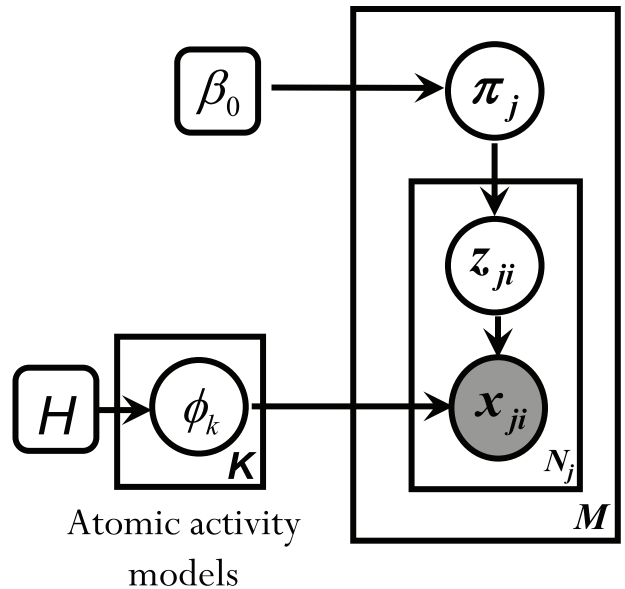
Cluster video clips directly using motion feature vectors without atomic activities

Advantages of this hierarchical Bayesian model (cont)

- Priors of global behaviors help to cluster moving pixels



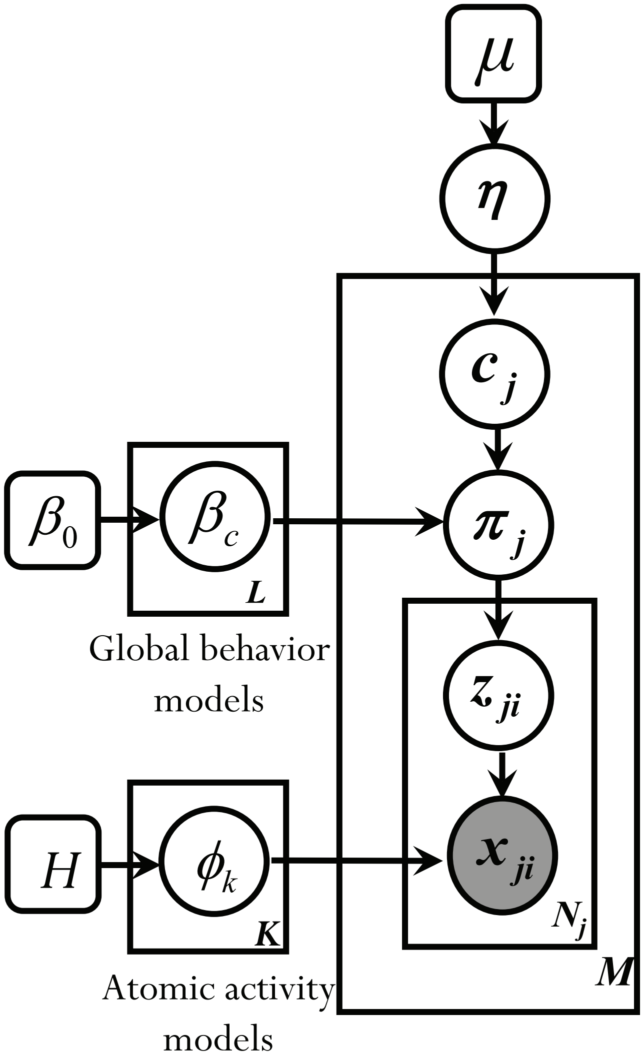
Our model



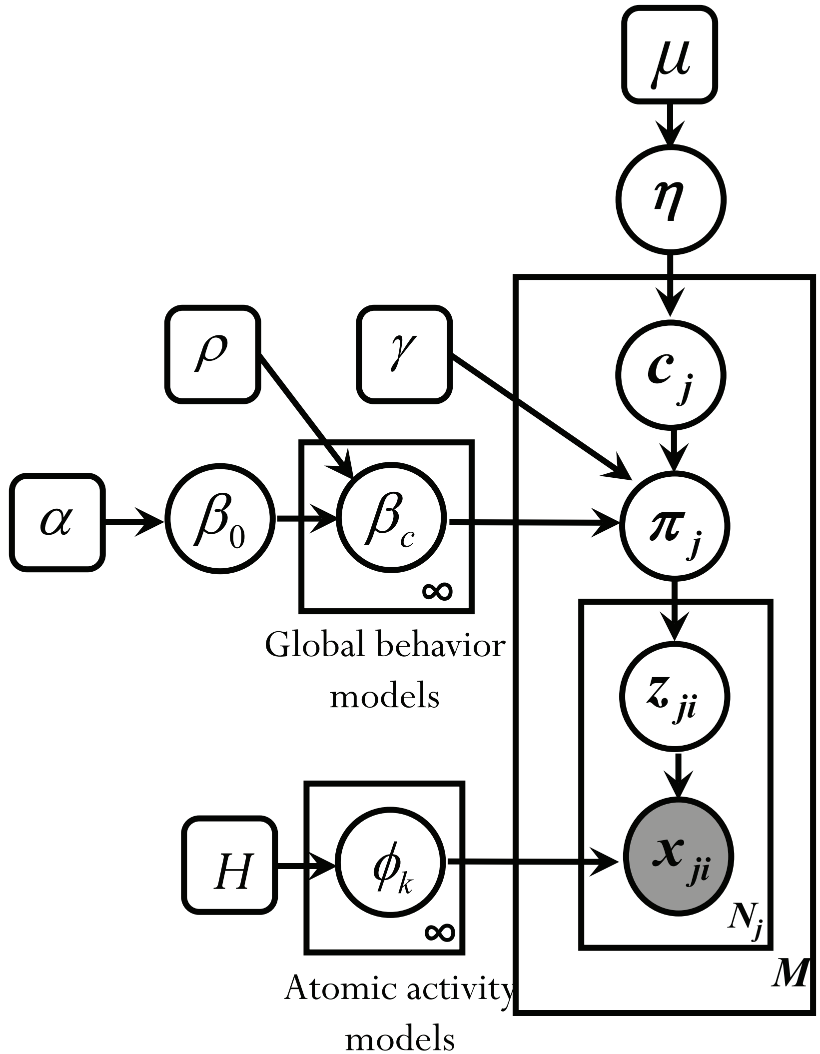
Cluster moving pixels without modeling interactions
Latent Dirichlet Allocation (LDA) [Blei et al. *JASA'03*]

Two atomic models have ambiguity in yellow circle area

Dual Hierarchical Dirichlet Processes (Dual-HDP)



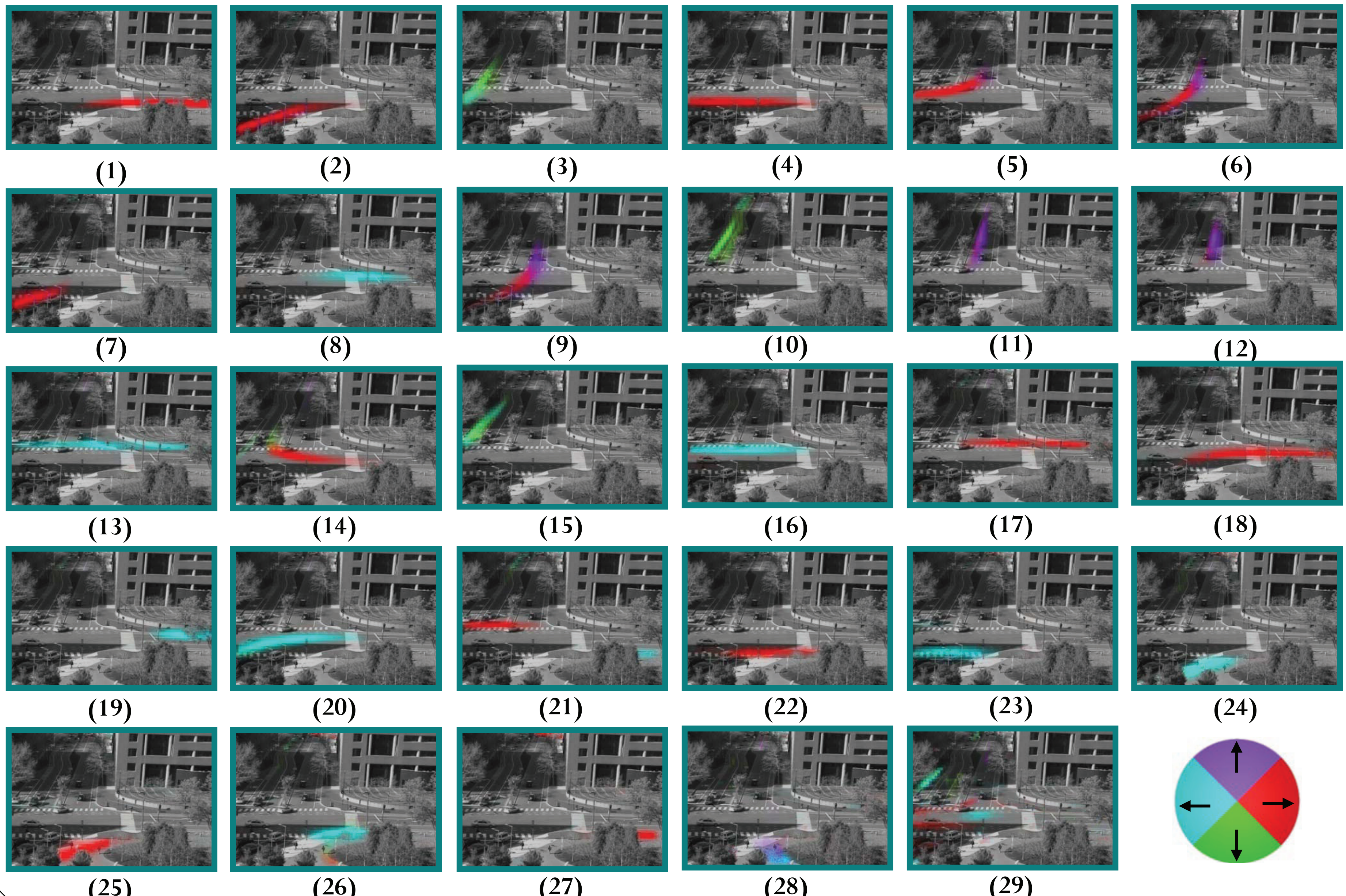
Parametric hierarchical Bayesian model

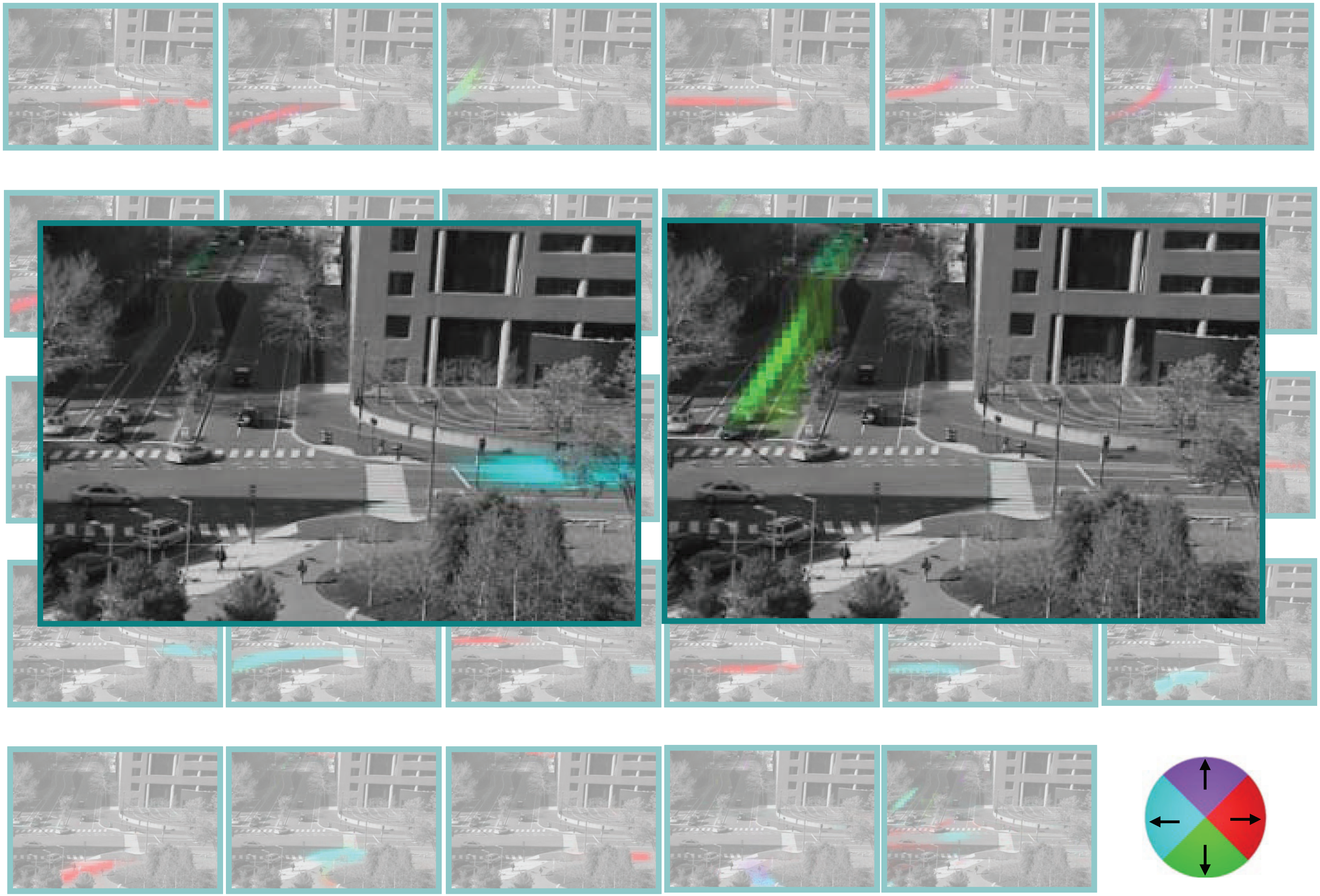


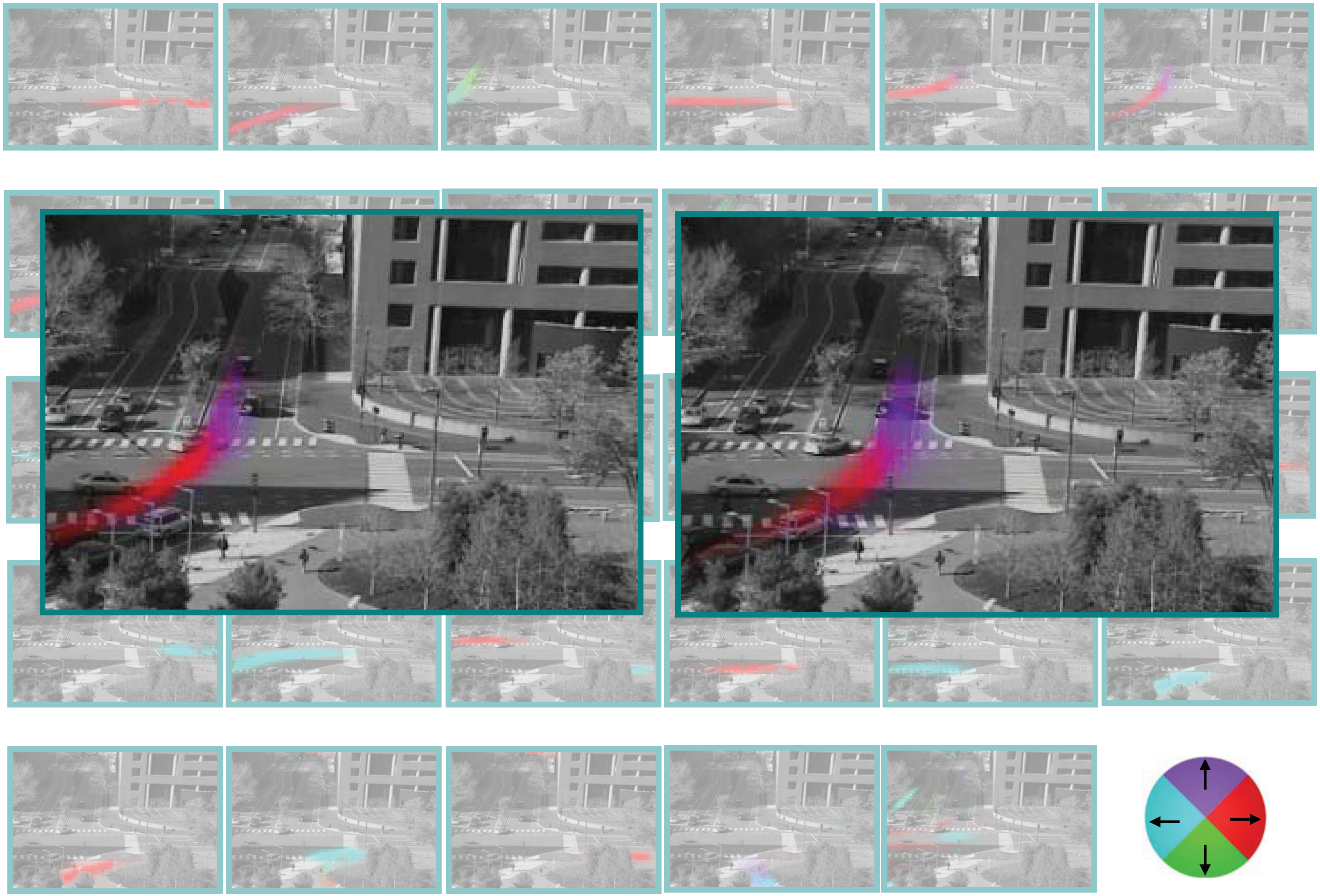
Dual-HDP

Related to the nested HDP [Rodriguez et al. 2006], Transformed HDP [Sudderth et. al IJCV'07], HDP [Teh JASA'04]

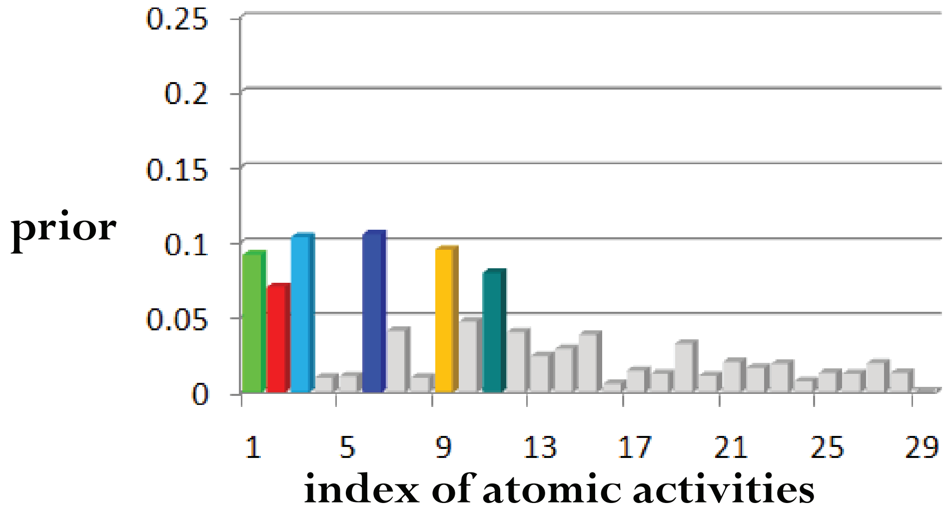
Learned atomic activities from a traffic scene

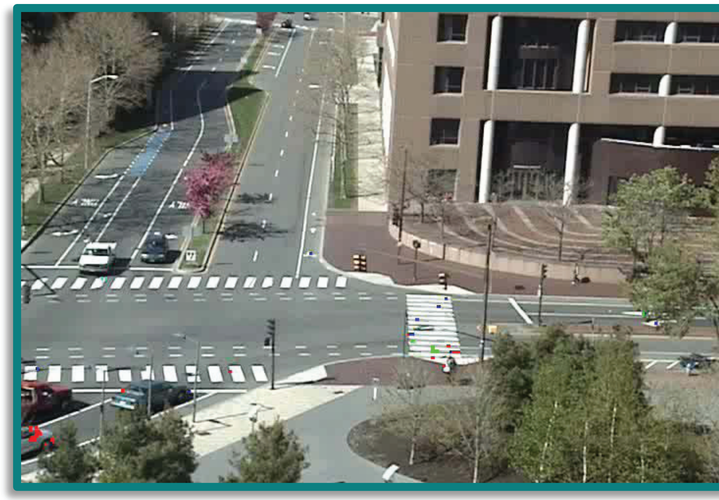






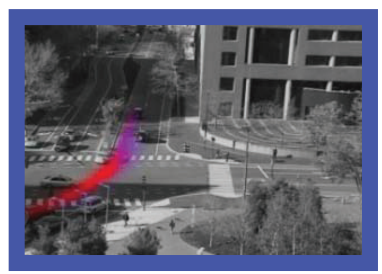
Global behavior I: green light for south/north traffic

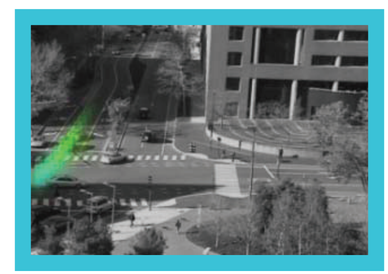


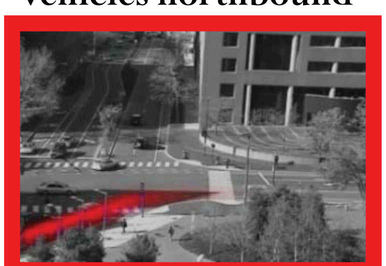


Top six atomic activities


vehicles northbound


vehicles northbound

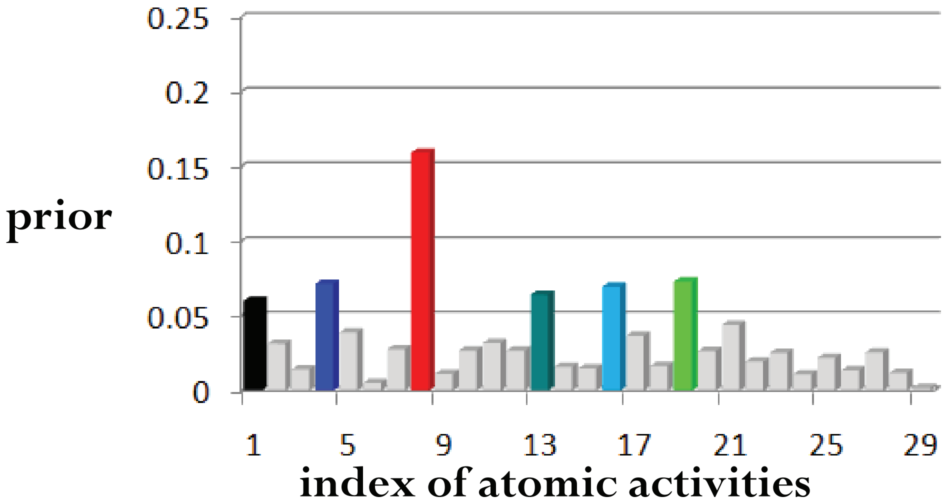

vehicles southbound


vehicles incoming northbound


vehicles incoming southbound


vehicles outgoing eastbound

Global behavior II: green light for east/west traffic



Top six atomic activities

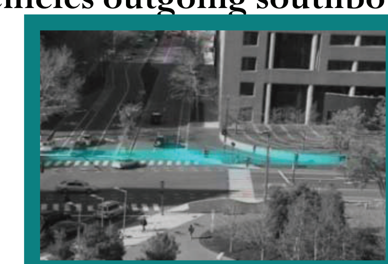
vehicles incoming westbound

vehicles outgoing westbound

vehicles outgoing southbound

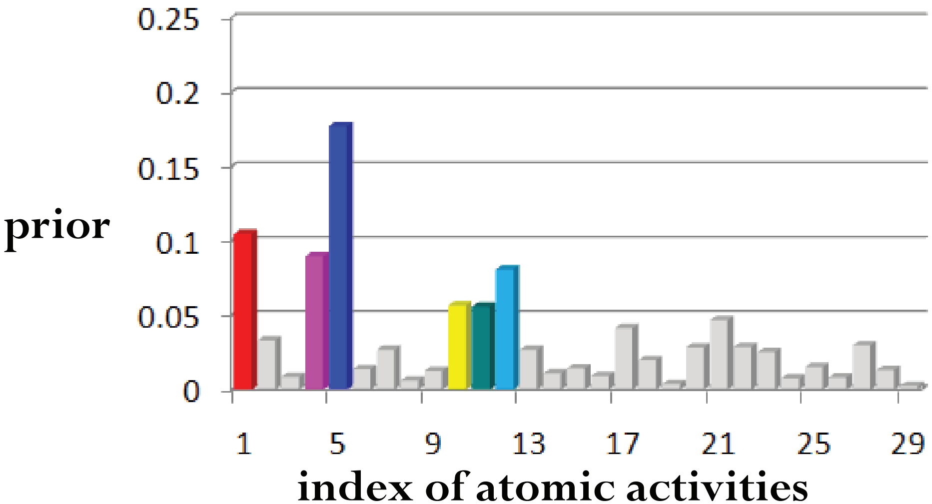
vehicles incoming eastbound

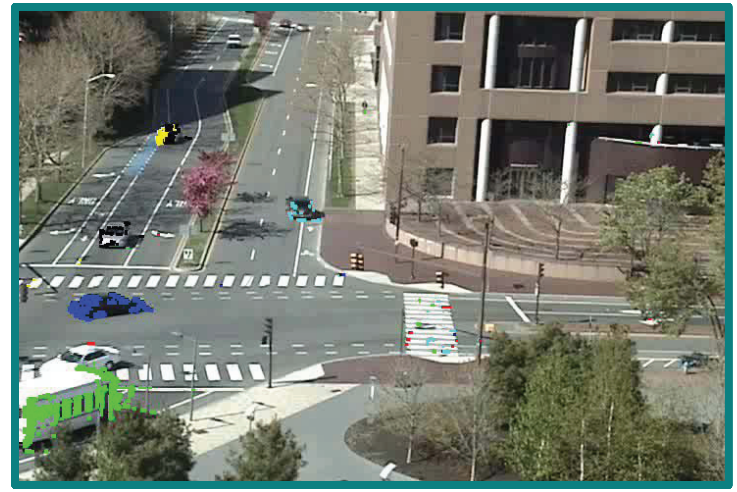
vehicles outgoing eastbound



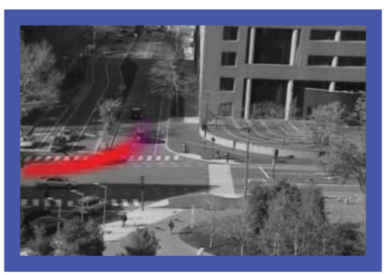
pedestrians westbound

Global behavior III: left turn signal for east/west traffic





Top six atomic activities



vehicles turning left eastbound

vehicles outgoing northbound

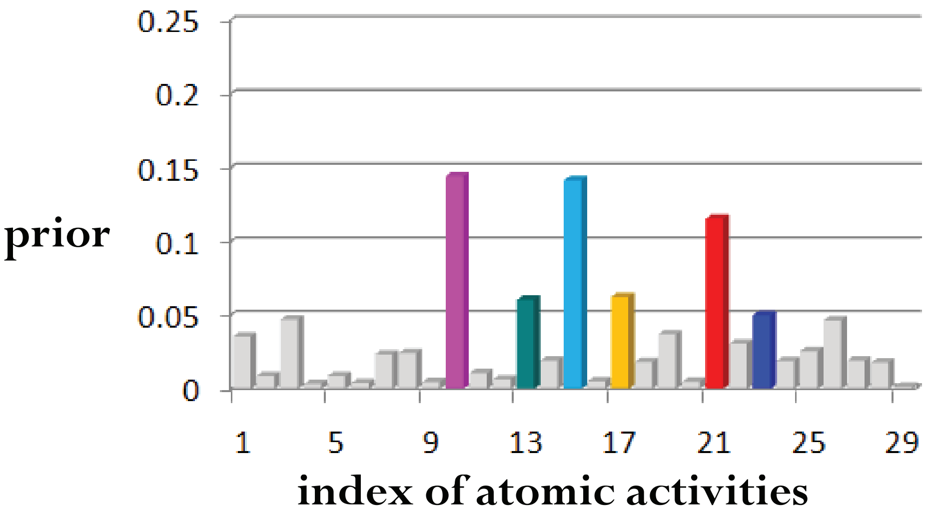
vehicles outgoing northbound

vehicles incoming eastbound

vehicles outgoing eastbound

vehicles stopping southbound

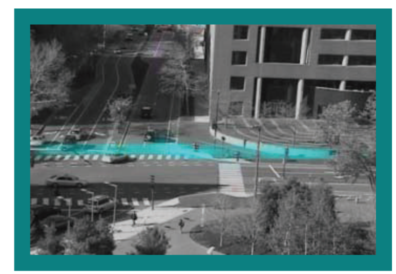
Global behavior IV: walk sign



Top six atomic activities

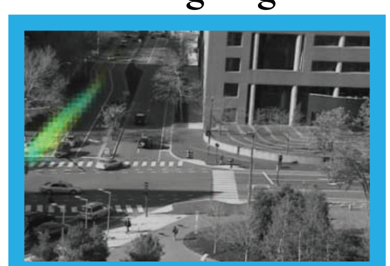
pedestrians incoming eastbound

pedestrians outgoing eastbound

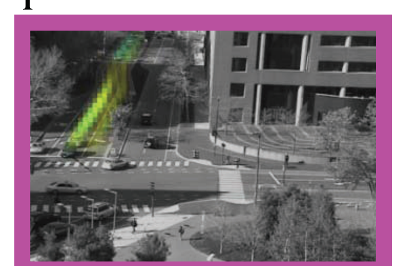


pedestrians westbound

pedestrians westbound

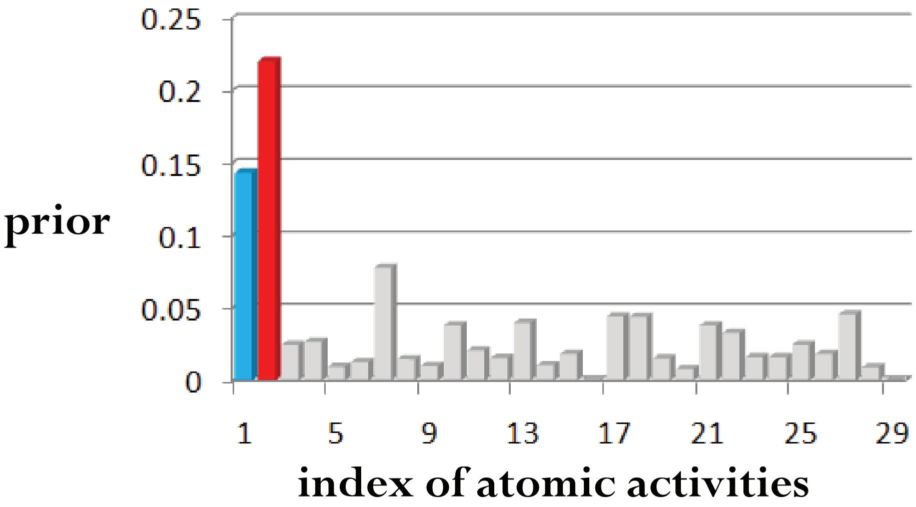


vehicles stopping



vehicles stopping

Global behavior V: northbound right turns



Top two atomic activities

vehicles incoming northbound

vehicles outgoing eastbound

Temporal video segmentation

- green light for east/west traffic
- green light for south/north traffic
- left turn signal for east/west traffic
- walk sign
- northbound right turns

Confusion matrix of video segmentation

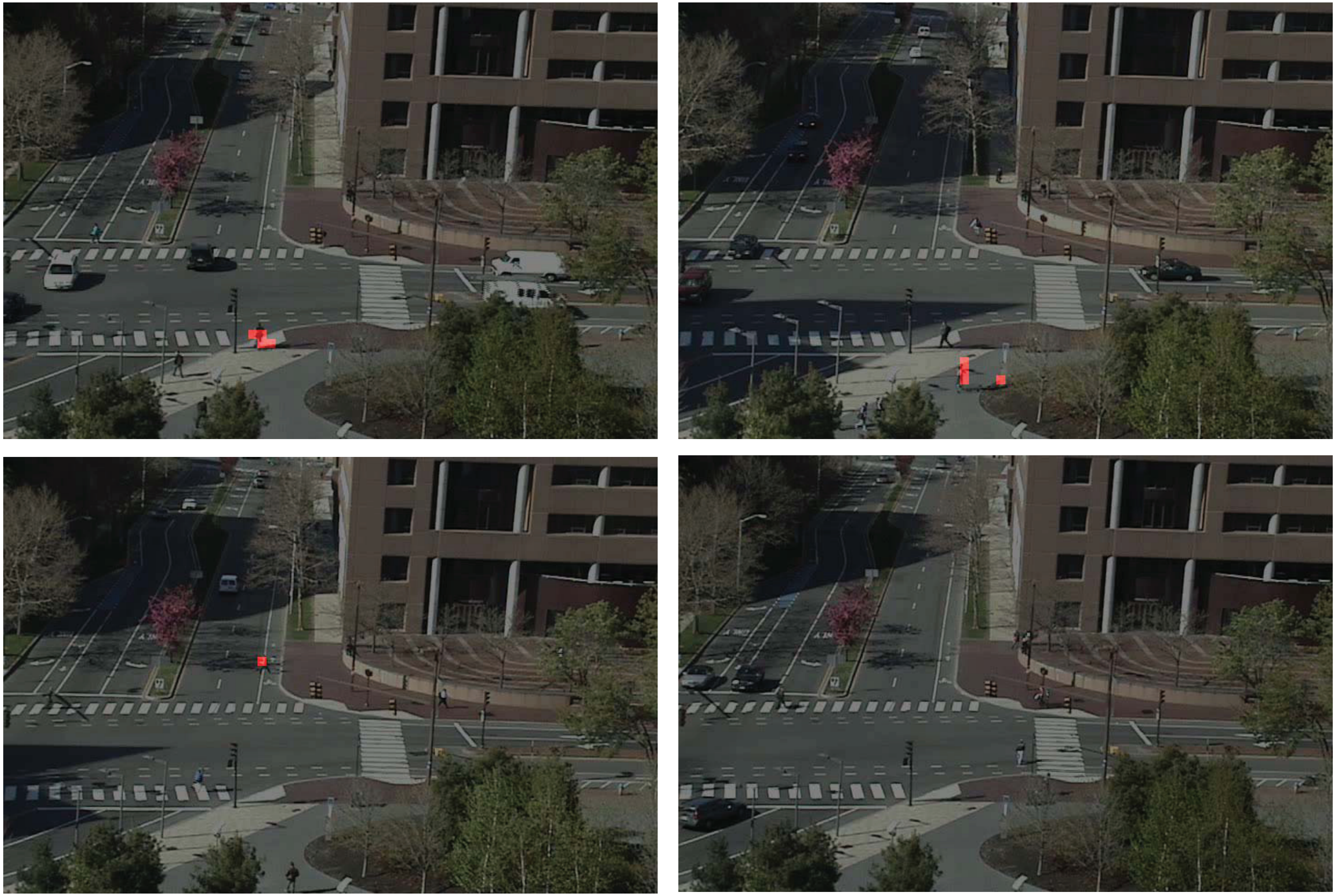
Clustering result

Manual label

149	0	2	0	0
8	74	4	2	11
10	3	60	1	2
4	0	2	88	11
4	2	6	5	92

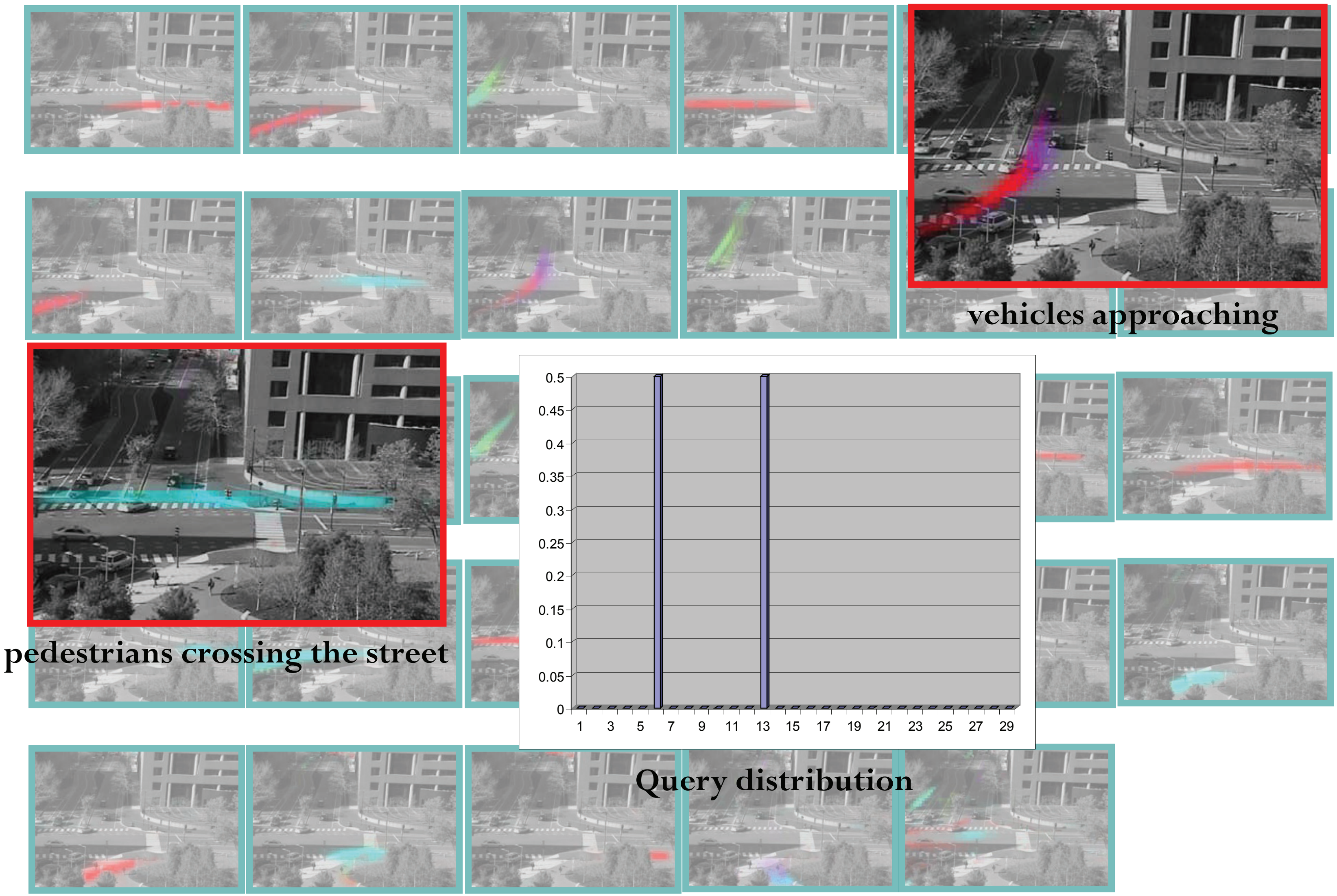
- The average accuracy is **85.74%** using our approach.
- The average accuracy is 65.6% when modeling atomic activities and global behaviors in two separate steps.
- The approaches of using a motion feature vector to represent a video clip perform poorly on this data.

Abnormality detection results



Top four abnormal video clips

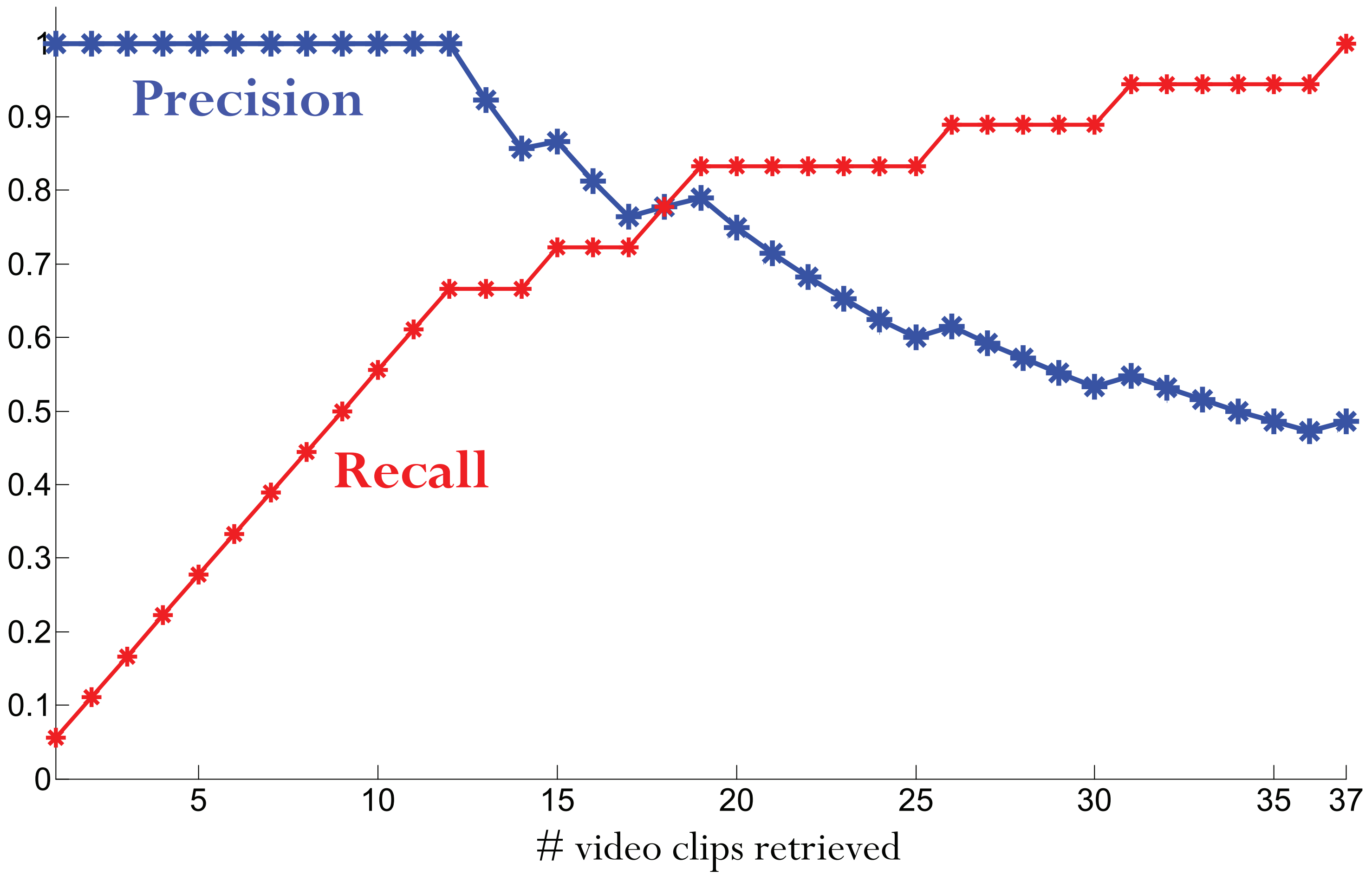
Interaction query



Top four retrieved jay-walking examples




Precision and recall of jay-walking retrieval



There are totally 18 instances, all found among the top 37 video clips out of 540 video clips.

Train Station Scene

Atomic activities

A scene where it fails…

Learned models of anatomic activities

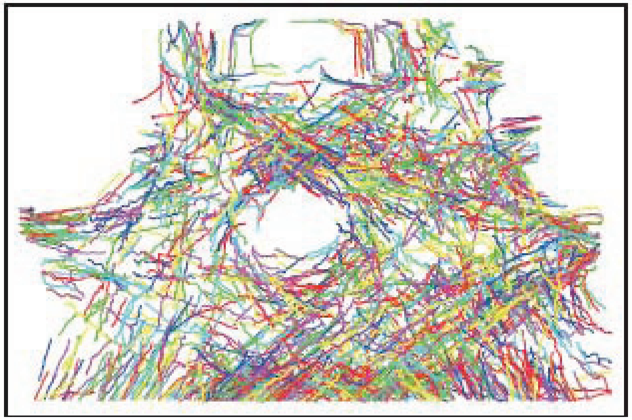
Outline

- Introduction
  - ❖ Why is behavior analysis in crowded environment interesting?
  - ❖ Major challenges
- Behavior analysis under hierarchical Bayesian models
  - ❖ Based local motions
  - ❖ Based noisy tracklets
- Other works
- Conclusions

B. Zhou, X. Wang, and X. Tang, "Random Field Topic for Semantic Region Analysis," *CVPR* 2011.

Tracklets

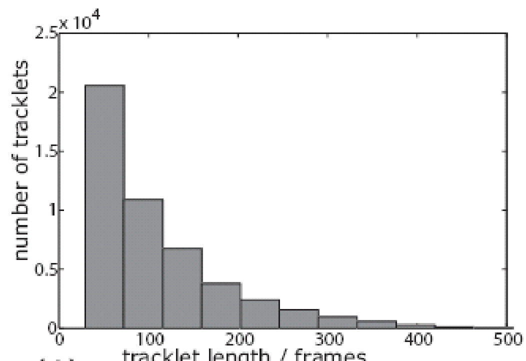
- Fragments of trajectories obtained by weak trackers. They are short and very noisy.



Statistics of 47,866 Tracklets

- The size of the scene is 1080 x 1920
- Sources: regions where objects appear
- Sinks: regions where objects disappear

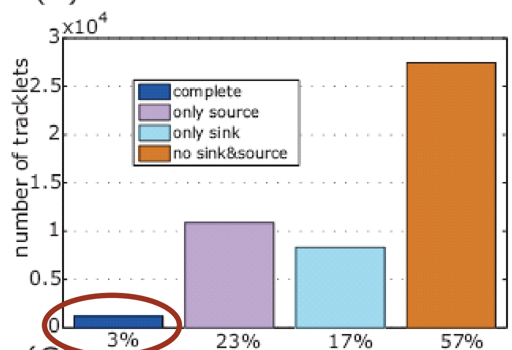
Histogram of the lengths of tracklets



(A)

(B)

Sources and sinks of the scene



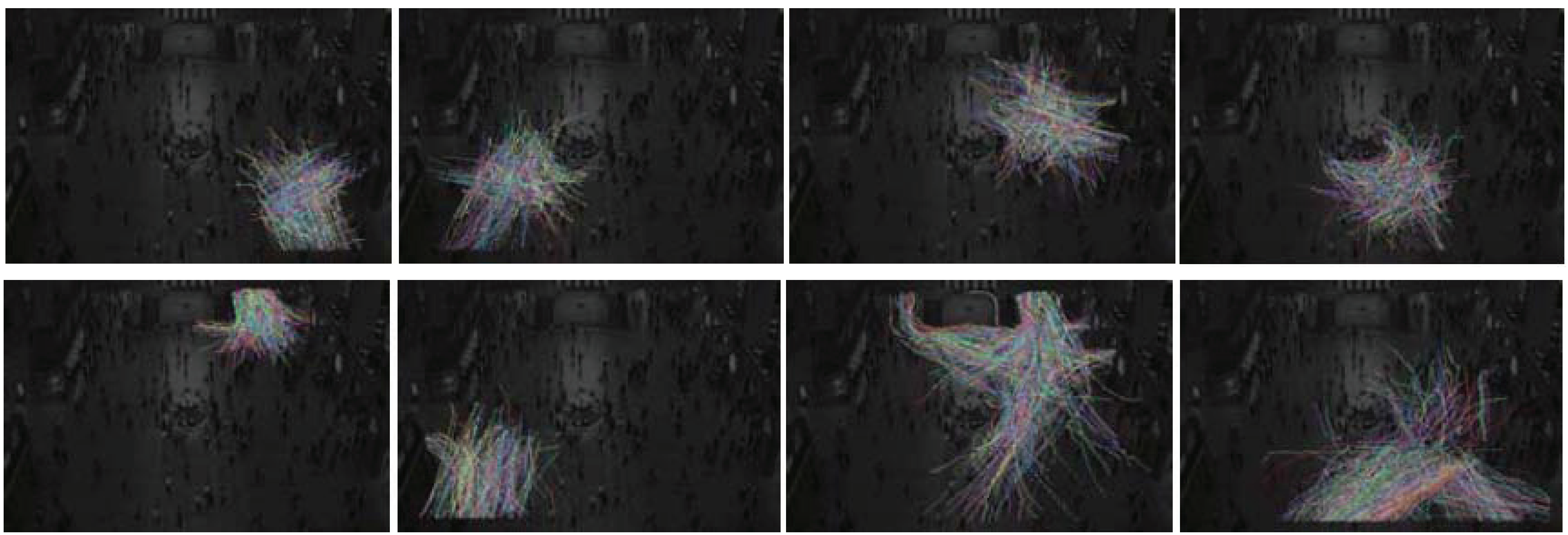
(C)

Source \ Sink	1	2	3	4	5	6	7
1	0	1	0	0	1	0	0
2	1	0	0	0	2	4	5
3	1	0	0	1	33	8	3
4	0	0	6	0	4	4	0
5	0	0	0	0	0	3	0
6	29	13	6	3	672	0	49
7	66	19	6	6	113	17	0

(D)

Numbers of tracklets connecting sources and sinks

Directly Clustering Tracklets

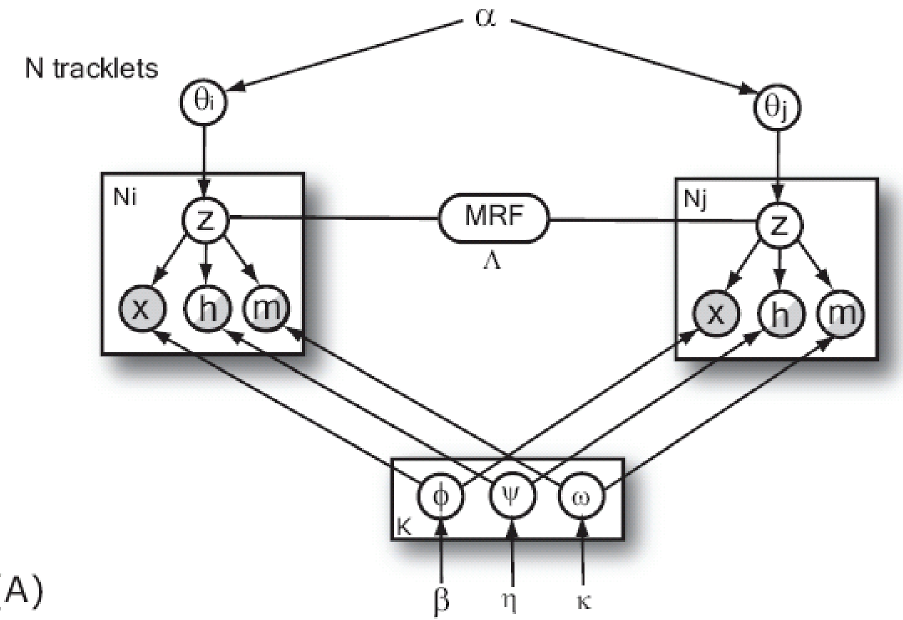


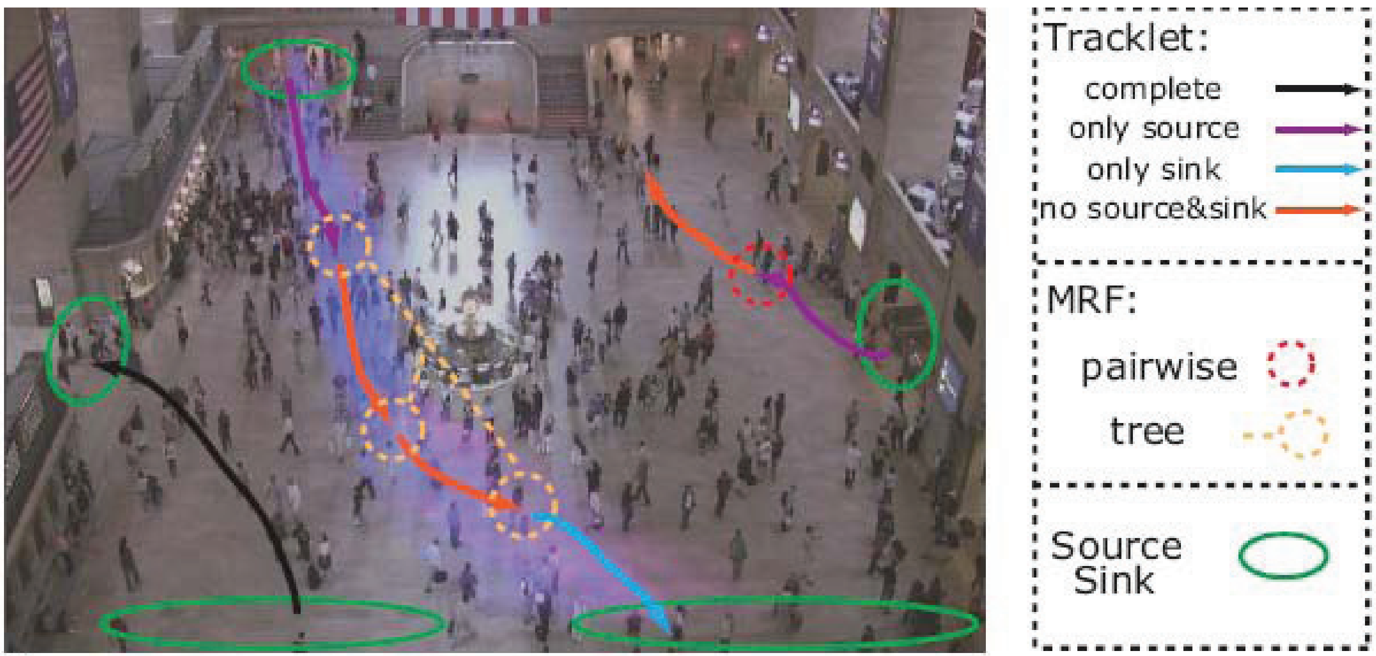
Spectral clustering + Hausdorff distance: X. Wang et al ECCV'06

Random Field Topic Models

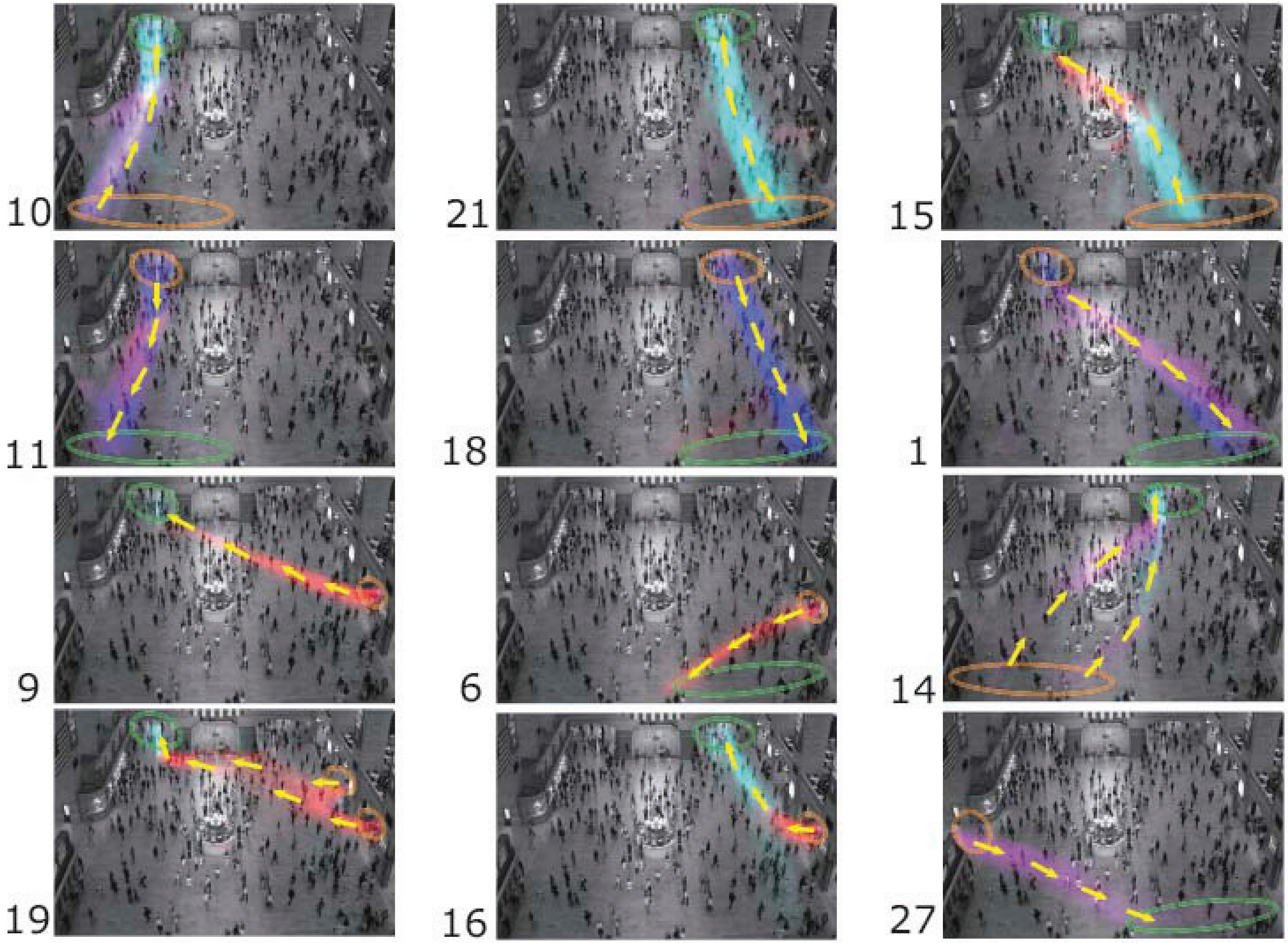
- MRF models the dependency between tracklets based on their spatial and temporal consistency and velocity similarity
- Model the sources and sinks

B. Zhou, X. Wang, and X. Tang, "Random Field Topic for Semantic Region Analysis," *CVPR* 2011.

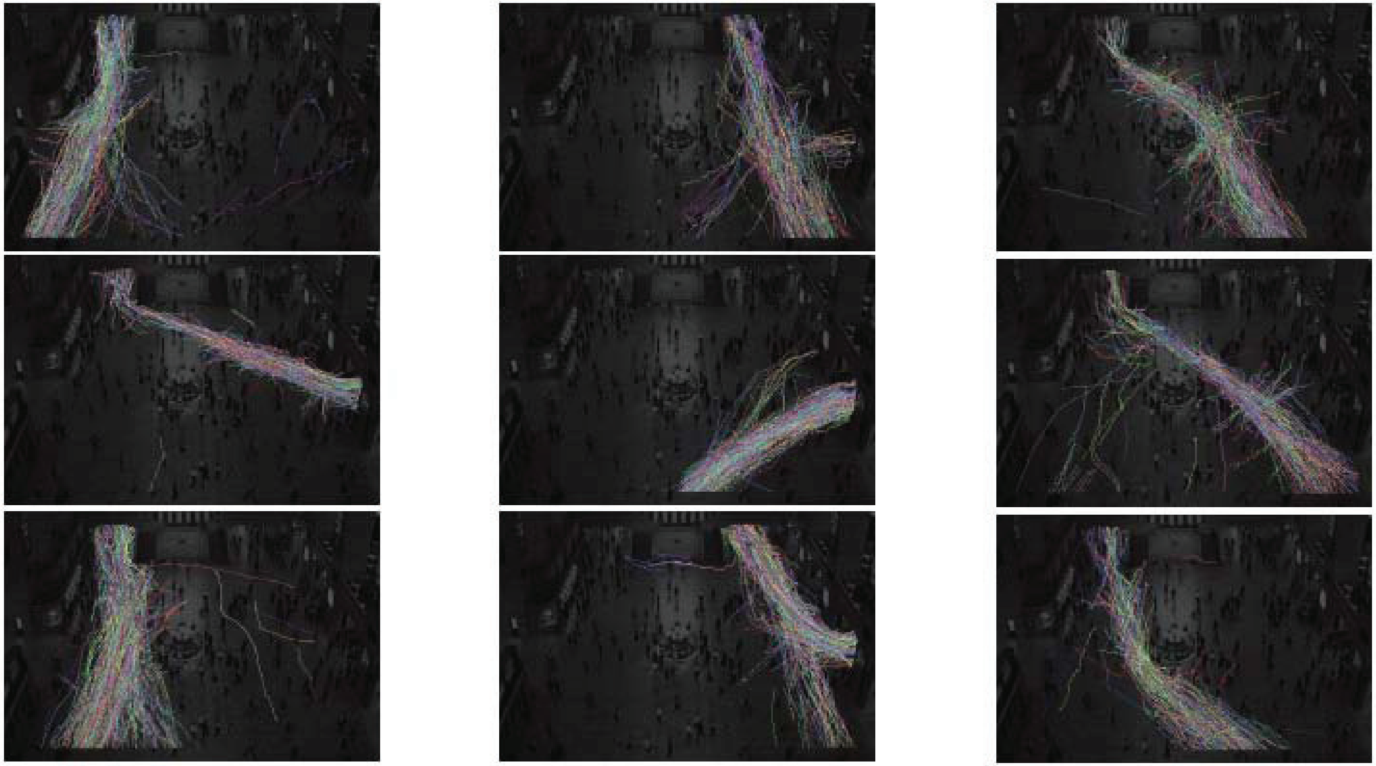

α
N tracklets
θi
θj
Ni
Nj
z
z
MRF
Λ
x
h
m
x
h
m
ϕ
ψ
ω
K
β
η
κ
(A)


Tracklet:
complete
only source
only sink
no source&sink
MRF:
pairwise
tree
Source
Sink
(B)

Learned Models of Paths

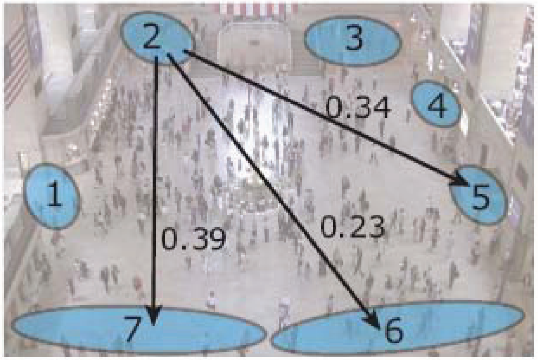


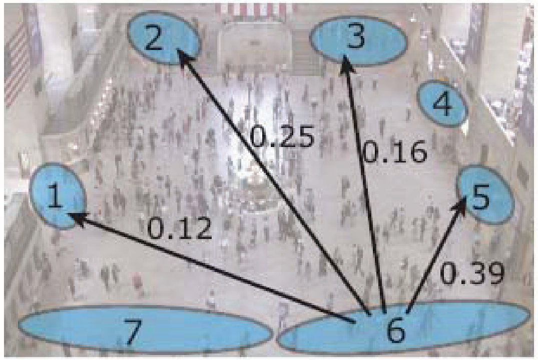
Tracklet Clustering Results



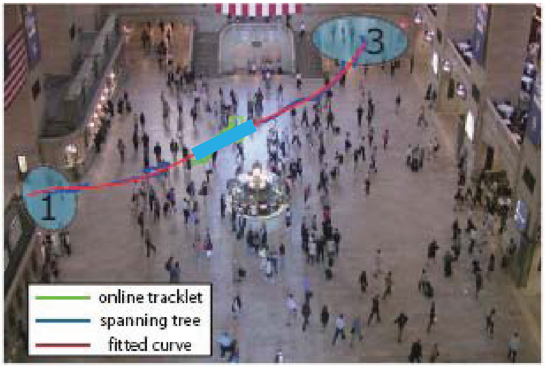
Potential Applications

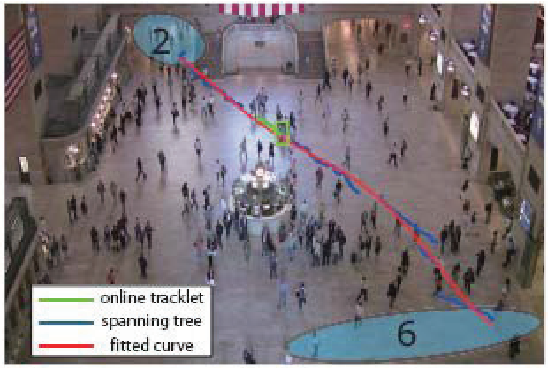
- Estimating the transition probabilities and traffic flows between sources and sinks





- Predicting the behaviors of pedestrians



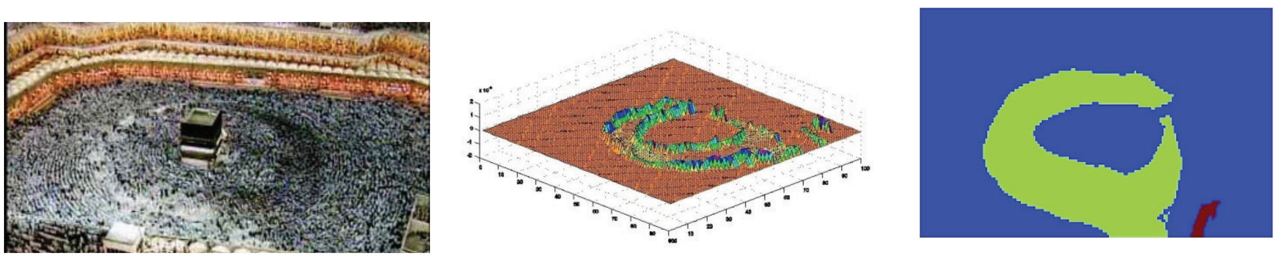


Outline

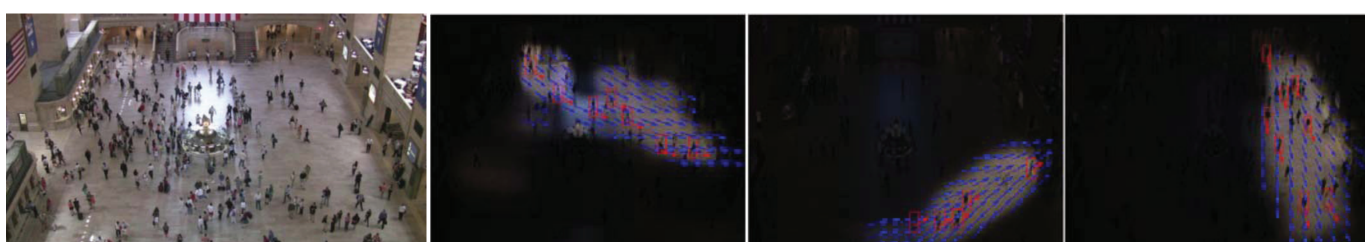
- Introduction
  - ❖Why is behavior analysis in crowded environment interesting?
  - ❖Major challenges
- Behavior analysis under hierarchical Bayesian models
  - ❖ Based local motions
  - ❖Based noisy tracklets
- Other works
- Conclusions

Other Works

- Crowd flow segmentation from Lagrangian coherent structures
  - S. Ali and M. Shah, “A Lagrangian particle dynamic approach for crowd flow segmentation and stability analysis,” CVPR 2007



- Flow field segmentation using a Lie algebraic approach
  - D. Lin, J. Fisher, E. Grimson, “Learning visual flows: A Lie algebraic approach,” CVPR 2009.



Other Works

- Abnormal crowd behavior detection using social force model
  - ➢ Mehran et al. CVPR’09
- Modeling social behavior for multi-target tracking
  - ➢ Pellegrini et al. ICCV’09
- Crowd behavior analysis across multiple camera views
  - ➢ Loy et al. CVPR 2009

Conclusions

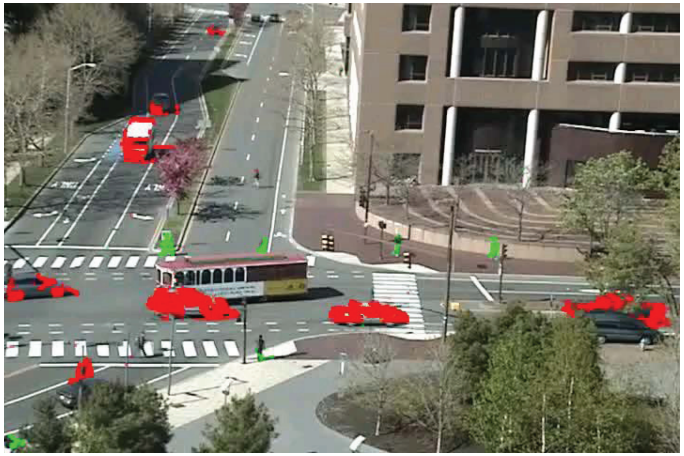
- Behavior analysis in crowded environments receives a lot attentions from different fields
- It is challenging problem from the computer vision point of view
- Propose an nonparametric hierarchical Bayesian model to learn behavior models in crowded environments from local motions
- It models single-agent activities, multi-agent interactions and global behaviors at different hierarchical levels
- Propose a random field topic model to learn behavior models from tracklets

Future Work

- Integrating the macroscopic models and the microscopic models
- Modeling the dynamic variations of the crowd behaviors
- Predicting the behaviors of individuals and the crowds
- Using the behavior models to improve detection and tracking
  - M. Wang and X. Wang, "Automatic Adaptation of a Generic Pedestrian Detector to a Specific Traffic Scene," *CVPR* 2011

Atomic activities related to vehicles

Atomic activities related to pedestrians



1st IEEE Workshop on Modeling, Simulation and Visual Analysis of Large Crowds

in conjunction with 13th International Conference on Computer Vision (ICCV)
6-13 November, 2011, Barcelona, Spain

Thank you!