Bolei Zhou[1] , Xiaogang Wang[2,3] and Xiaoou Tang[1,3]

[1] Department of Information Engineering, [2] Department of Electronic Engineering, The Chinese University of Hong Kong
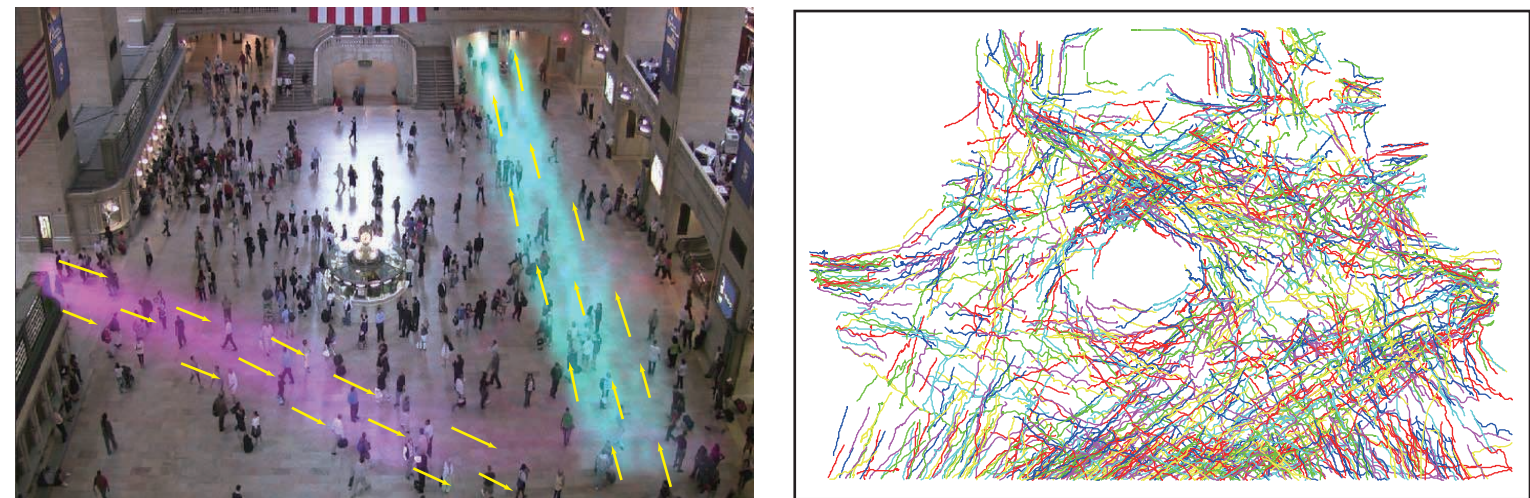[3] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

CVPR 11
20 COLORADO 11

Introduction

Goal: Identifying semantic regions of crowded scenes from tracklets

☐ Semantic regions: pathways commonly taken by objects in scene.
application: tracking, detection, activity analysis, etc
☐ Tracklets: fragments of trajectory by weak keypoint tracker, such as KLT.
caused by scene clutter, very noisy.
☐ Challenge: how to learn semantic regions from such noisy trajectories

Two key components in modeling

Tracklet:
complete
only source
only sink
no source&sink

MRF:
pairwise
tree

Source
Sink

Modeling correlations between tracklets

☐ Pairwise correlation: dependencies between two tracklets
I. temporal overlap; II. spatial overlap. III. velocity overlap
Markov Random Field of tracklets

☐ High-level correlation: dependencies among several tracklets
Spanning tree on MRF of tracklets

Modeling source and sink of the scene

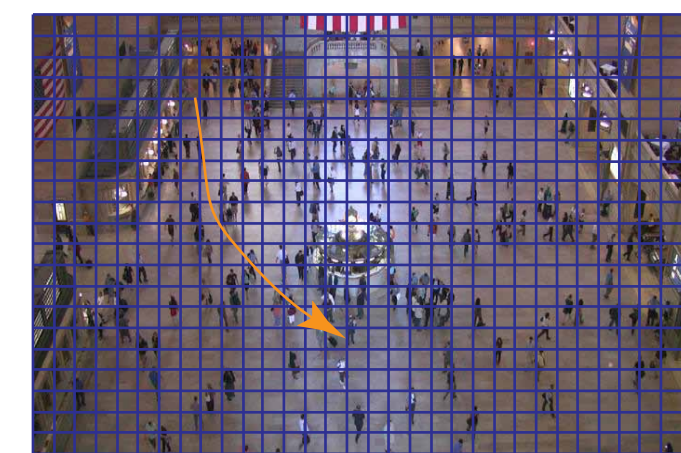Region of Interests: Entry and exit locations of the scene, initial position and ending position of people.

Graphical Model

standard LDA topic model

Random Field Topic model
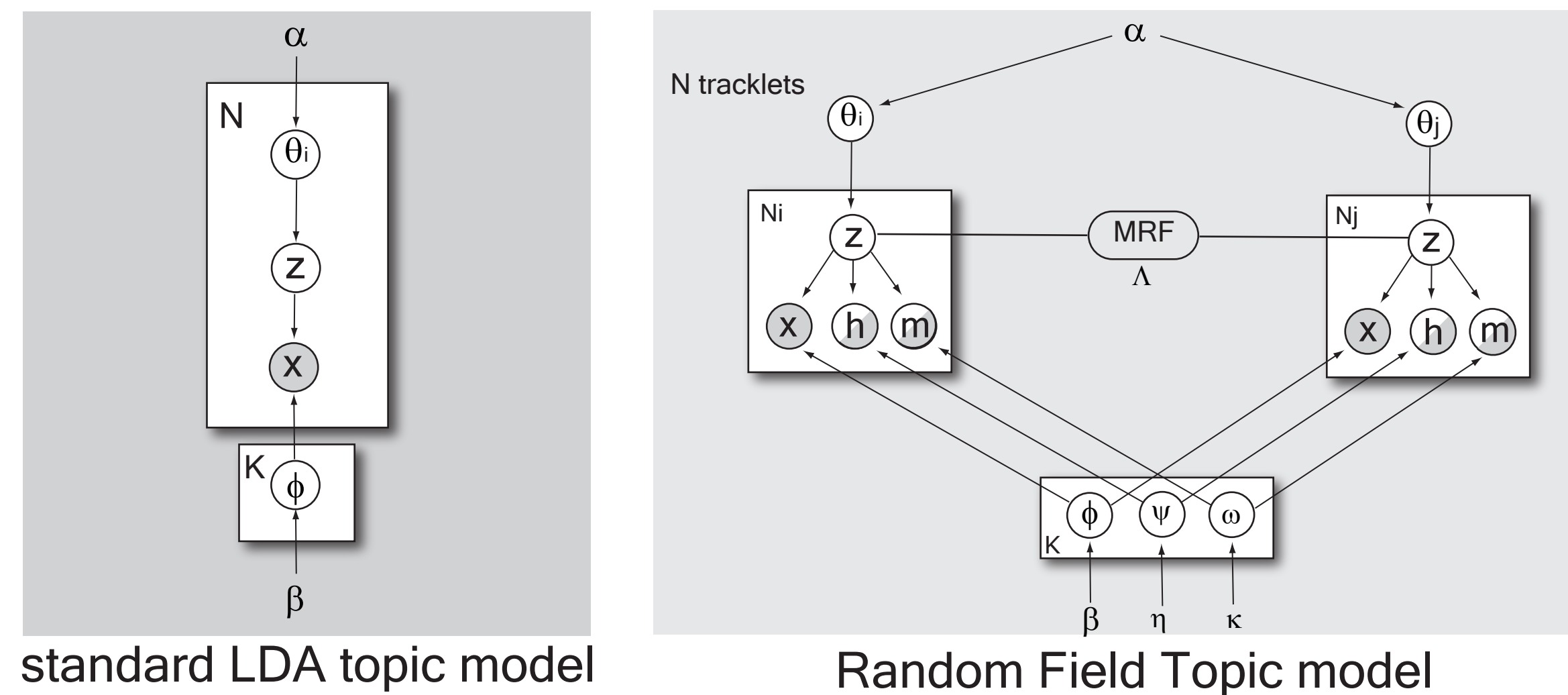
Topic modeling:

points of tracklet -> visual words in cell
tracklet -> document of words
semantic region -> topic

divide scene into cells

MRF conditions:

pairwise MRF: $I.\ t_i^e < t_j^s < t_i^e + T,\ II.\ |x_i^e - x_j^s| + |y_i^e - y_j^s| < S,\ III.\ \frac{v_i^e \bullet v_j^s}{\|v_i^e\|\|v_j^s\|} > C.$

$$\Lambda\,(z_{in_1},\ z_{jn_2}) = \exp\left(\frac{v_i^e \bullet v_j^s}{\|v_i^e\|\|v_j^s\|} - 1\right)$$

$$p\,(Z\,|\,\theta) \propto \exp\left(\sum_i \log\theta_i + \sum_{j \in \varepsilon(i)} \sum_{n_1,\,n_2} \Lambda\,(z_{in_1},\ z_{jn_2})\right)$$

tree MRF: spanning search on pairwise MRF
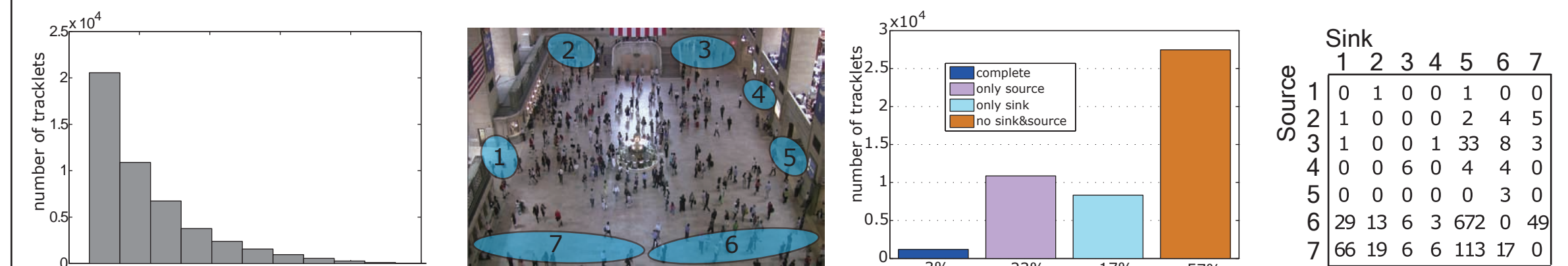with influence of source and sink

Model inference: Gibbs sampling

$p(z_{in} = k | X, Z_{\backslash in}, H, M)$

$$\propto \frac{n_{k,\backslash in}^{(w)} + \beta}{\sum_{w=1}^{W}(n_{k,\backslash in}^{(w)} + \beta)} \frac{n_{k,\backslash in}^{(p)} + \eta}{\sum_{p=1}^{P}(n_{k,\backslash in}^{(p)} + \eta)} \frac{n_{k,\backslash in}^{(q)} + \kappa}{\sum_{q=1}^{Q}(n_{k,\backslash in}^{(q)} + \kappa)} \frac{n_{i,\backslash n}^{(k)} + \alpha}{\sum_{k=1}^{K}(n_{i,\backslash n}^{(k)} + \alpha)} \left(\exp\left(\sum_{j \in \gamma(i)} \sum_{n'} \Lambda\,(z_{in}, z_{jn'})\right)\right)$$
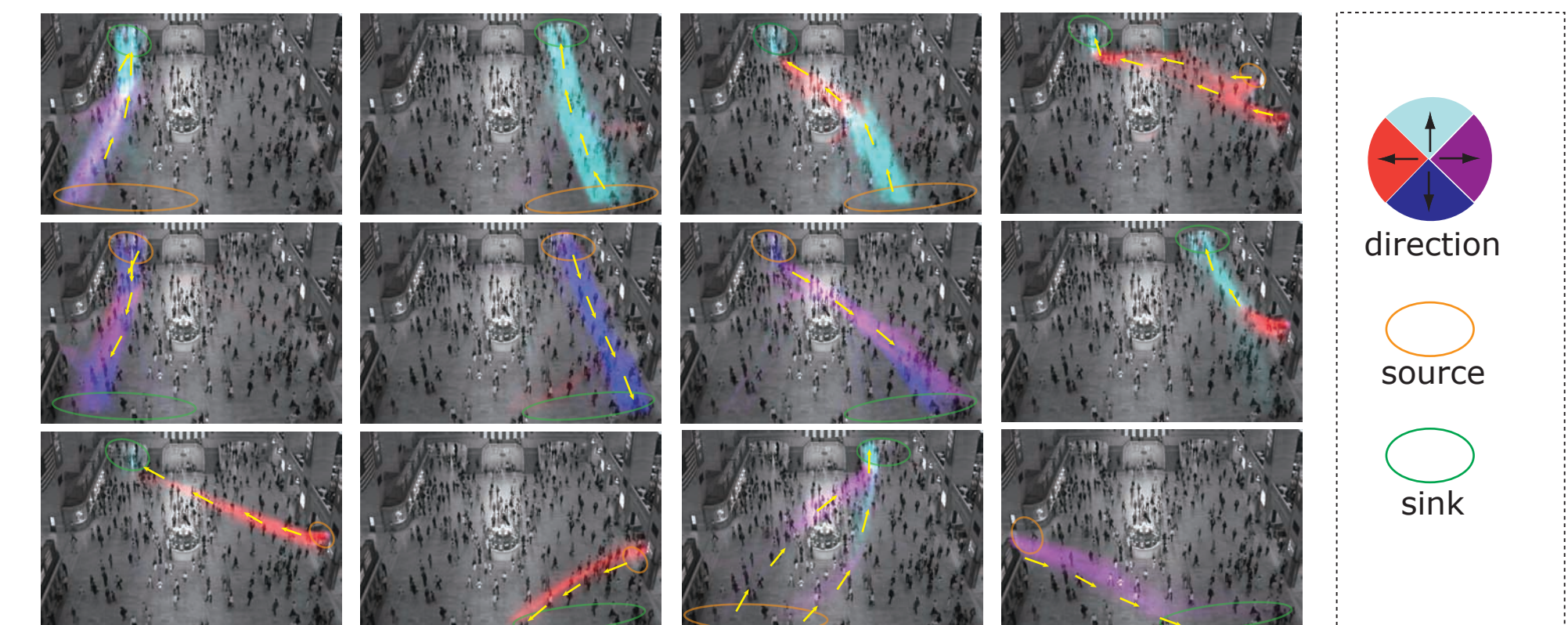
source&sink   source&sink   MRF
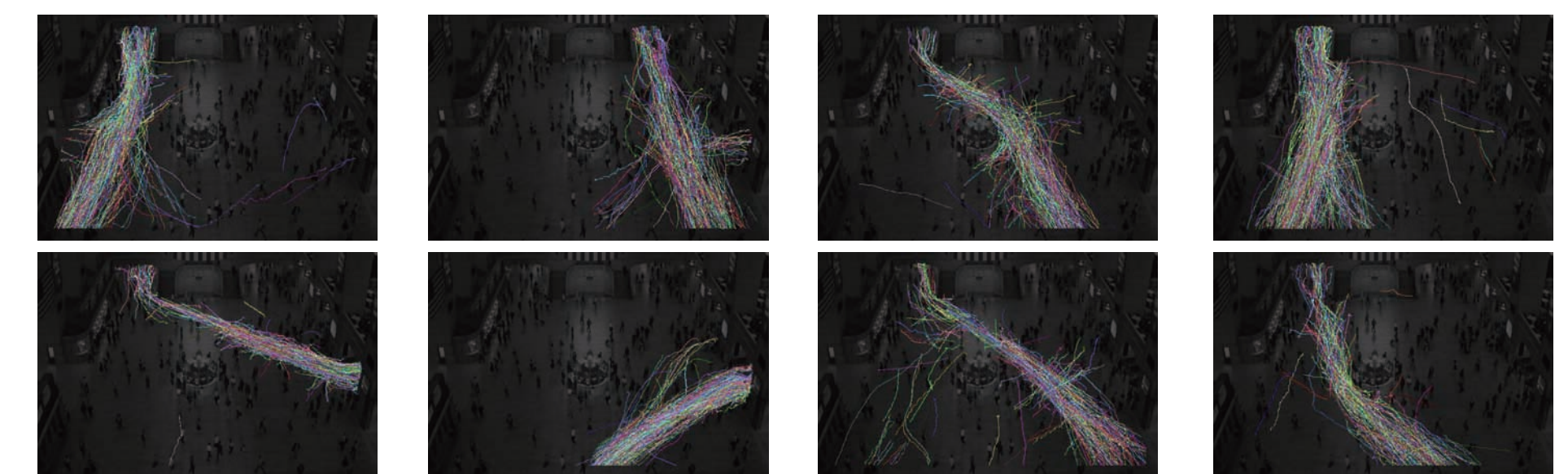
Experimental Results & Applications

Statistics of the tracklets captured from the scene

	Sink						
	1	2	3	4	5	6	7
Source 1	0	1	0	0	1	0	0
2	1	0	0	0	2	4	5
3	1	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

Learned semantic regions
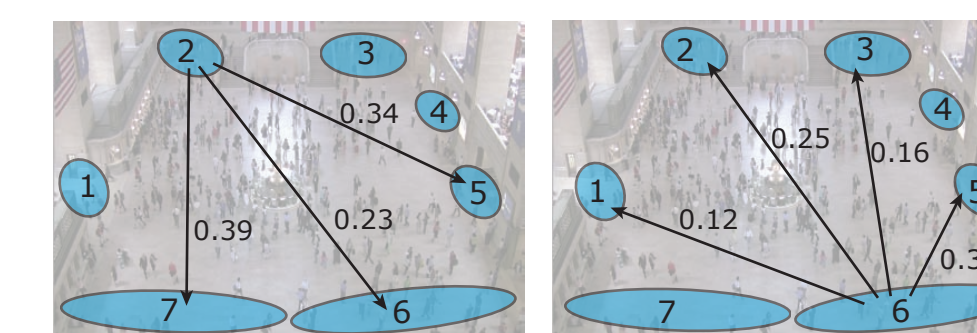

direction
source
sink

Trajectory clustering

Potential applications

people flow transition ratios   people prediction: future path