Unusual Event Detection without Object Tracking

László ***Czúni*** / Ákos Utasi

University of Pannonia, Veszprém

Faculty of Informatics

Outline

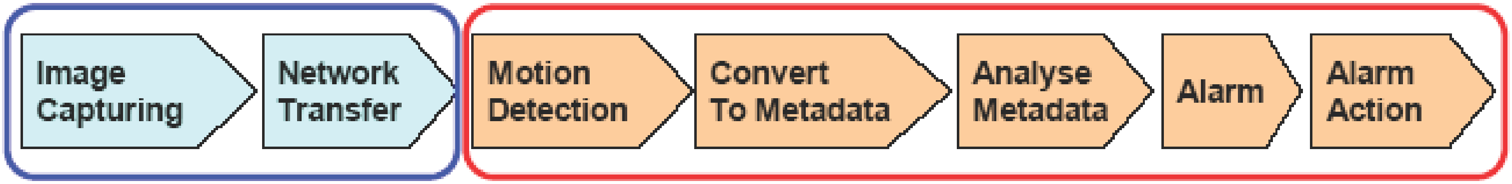
- State of visual surveillance
- Tasks, problems, previous models
- Hidden Markov Model approach:
  - About HMMs
  - Preprocessing steps
  - Generating observations
  - Model training (and problems)
  - Detection, analysis
- Hierarchical HMM approach
- Results, demonstration

Application Areas of Visual Surveillance

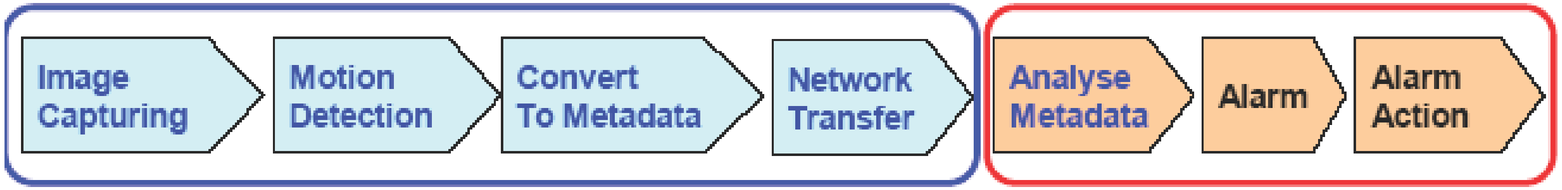
- **Transportation**: traffic counting by lanes, speed estimation, numberplate recognition, forbidden motion detection, forbidden areas
- **Trade (shops, banks) and public organizations (schools, hospitals, offices):** running human detection, lost/stolen object detection, path discovering, queue detection, crowd detection
- **Industry:** process analysis, unusual event detetion, quality monitoring
- and a lot more...

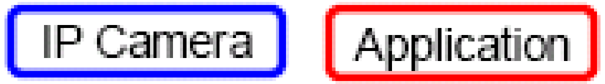
Distributed Data Processing

- Old digital (IP) approach:

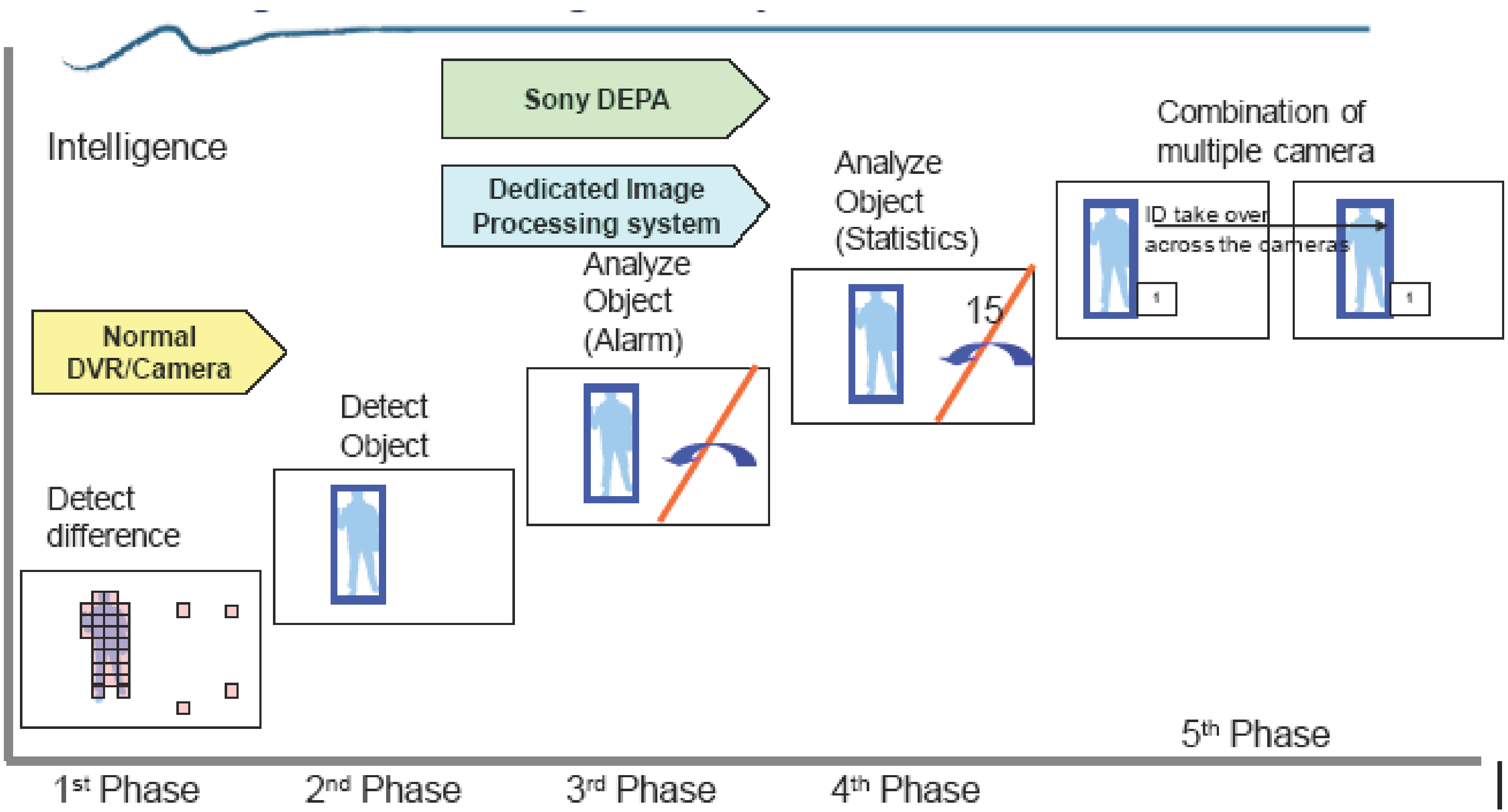


- Distributed processing:



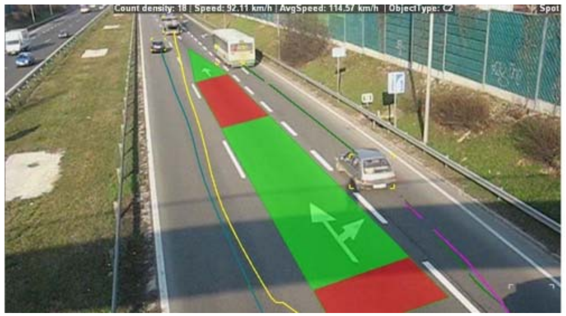


Sony's Distributed Enhanced Processing Architecture

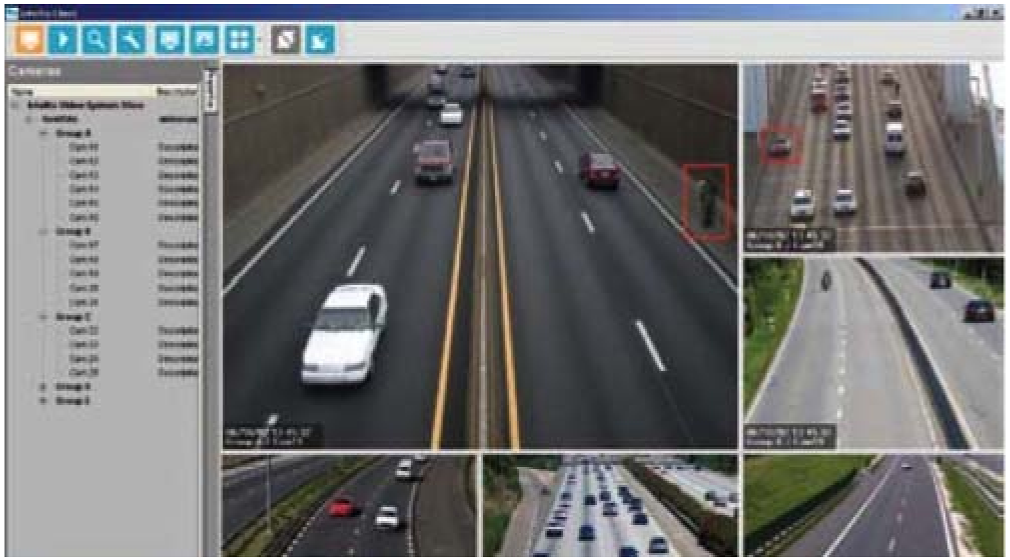


Example for distributed processing: Intellio product line for traffic monitoring

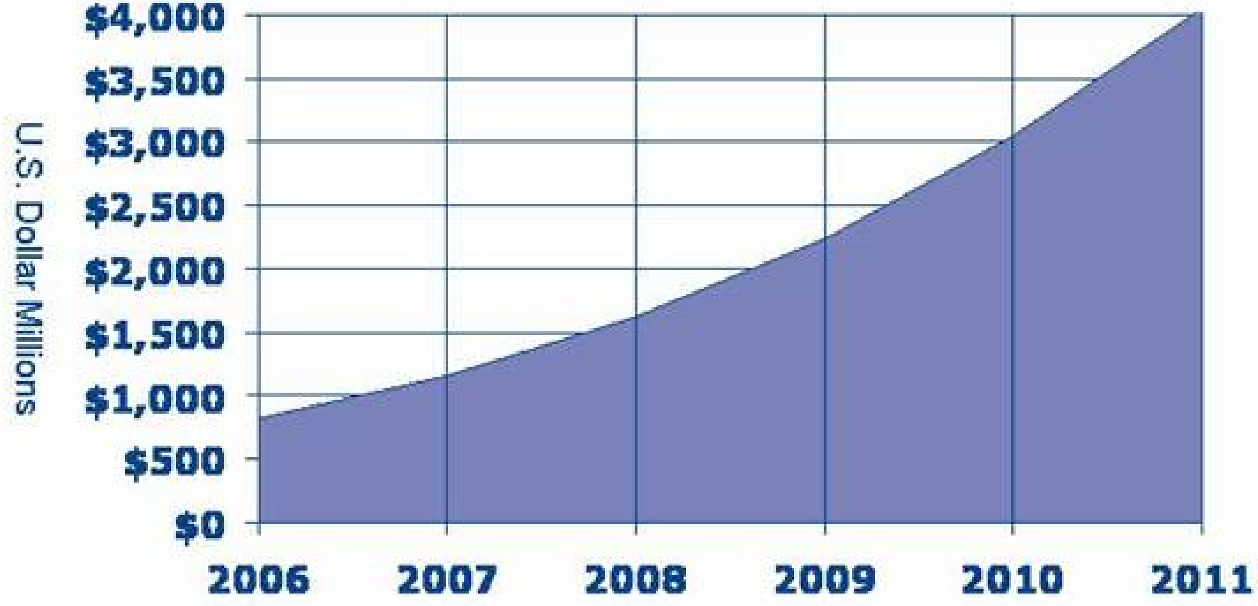
- M0 ring around Budapest is equipped with Intellio's Intelligent cameras



- Distributed system can estimate:
  - Speed of vehicles
  - Motion at forbidden areas
  - Speed Dome control for high resolution images
  - Emergency alarms and accident prevention
  - Integration with loop detectors



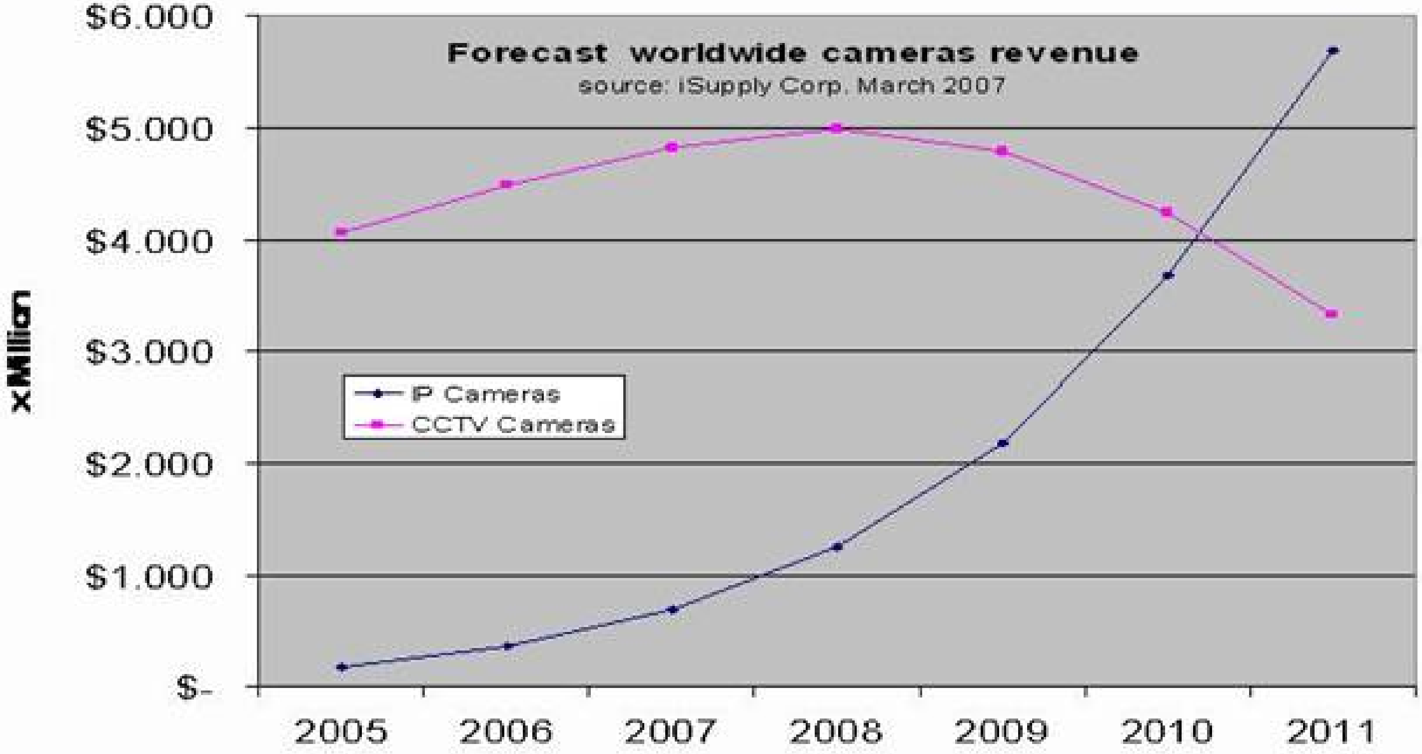
Market Trends: IP Video Surveillance



- IP video surveillance market shows very high growth
- **Forecast annual growth rate of 38% (IMS Study)**

Optelecom-nkf

Market Trends: IP Cameras Overtake Video over Fiber



IP video will become the dominant technology

Optelecom-nkf

Worldwide Research Activity

- B. T. Moeslund and E. Granum. A survey of advances in vision-based human motion capture. *Computer Vision and Image Understanding*, 81(3):231-268, 2001. **155 papers**
- T. B. Moeslund, A. Hilton, and V. KrÄuger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2-3):90-126, 2006. **424 papers 2000-2006**
- Niels Haering, Péter L. Venetianer, Alan Lipton. The evolution of video surveillance: an overview, Machine Vision and Applications (2008) 19:279-290

Recent works at the University of Pannonia

- Camera calibration for omnivision systems: generating undistorted perspective image from annular image
- Improved motion detection: reducing the foreground aperture problem
- Unusual event detection
- Surveillance video segmentation

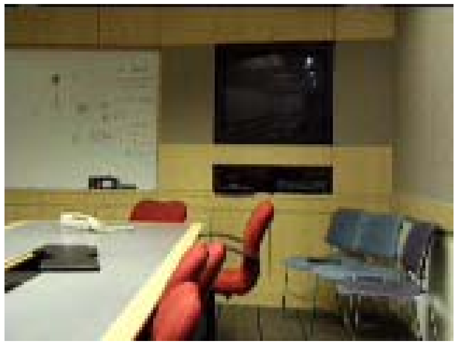
Improved Motion Detection

Problem: ***foreground aperture problem*** *(some moving areas are not detected in homogenous regions)****.***

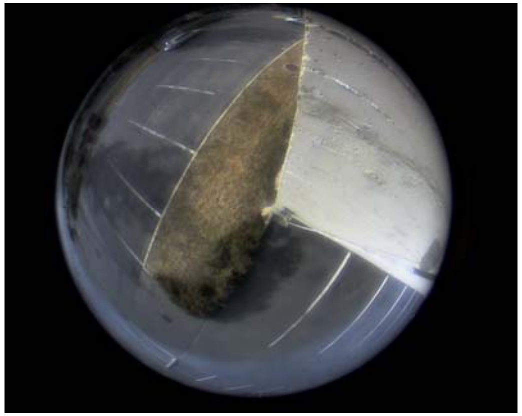
Solution: improved Mixture of Gaussians method.



Input Test Video | Original MOG | **Improved MOG**



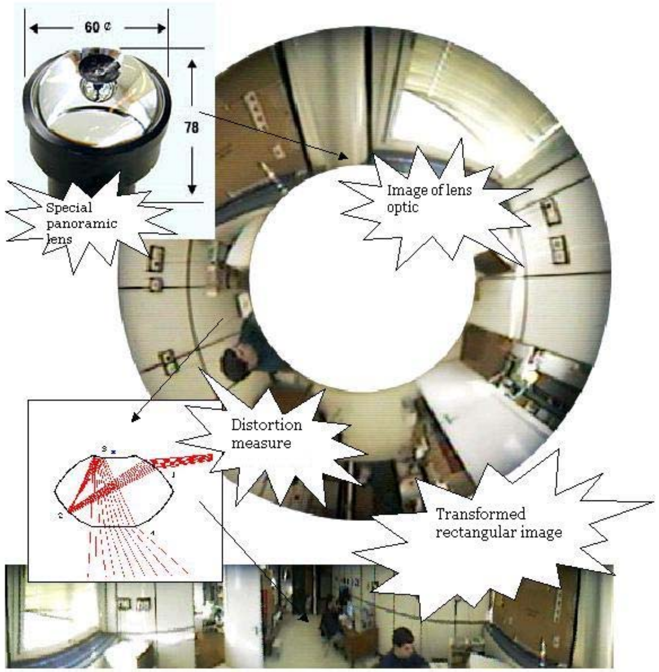
Omnivision for Security



Camera Image

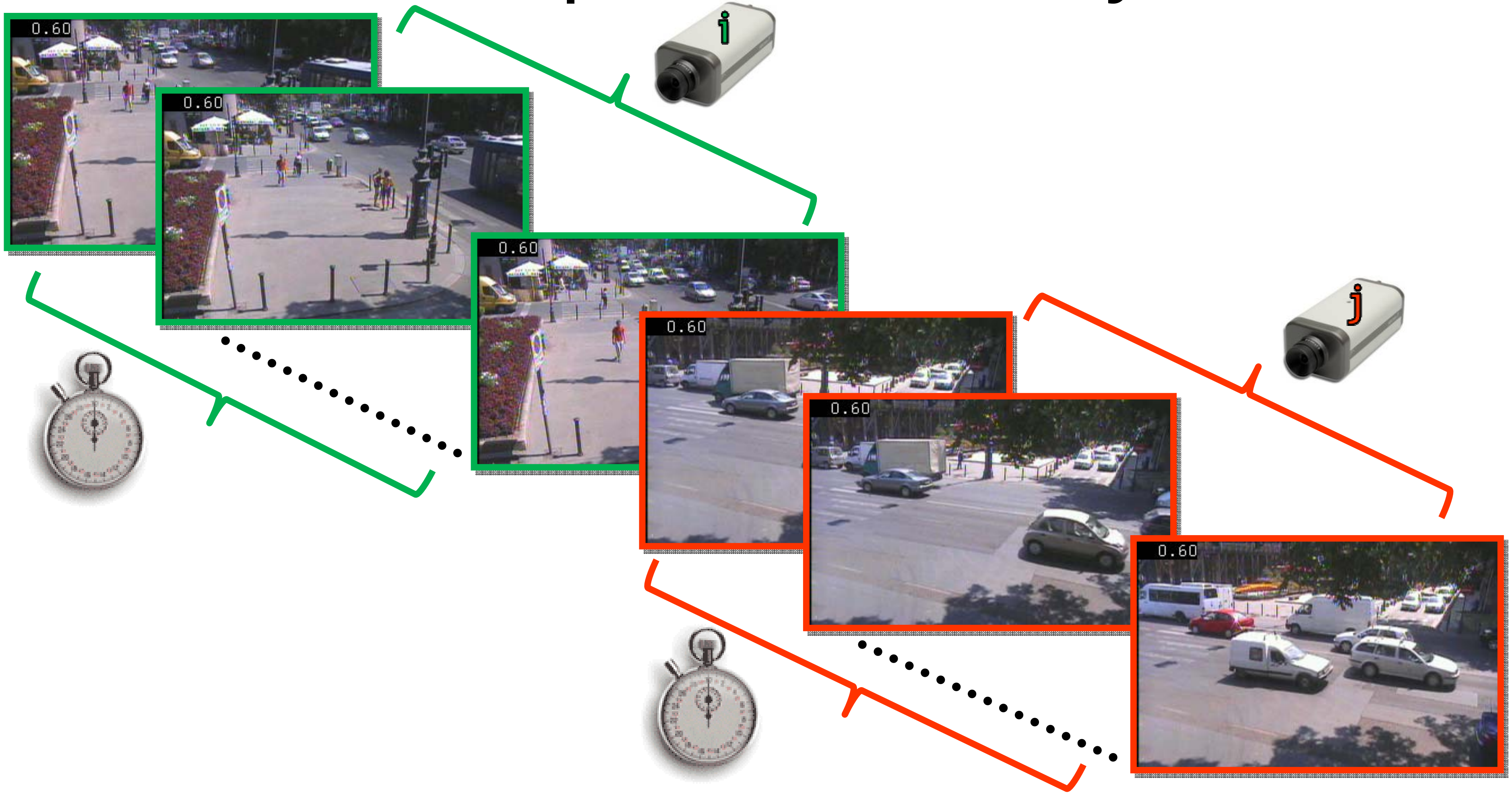
360 degree squared image

Virtual prespective images



Camera calibration

Hidden Semi Markov Models for Temporal Video Segmentation of Time-multiplexed Security Videos



Our Motivation

- Built up surveillance systems in cities:
  – Low-cost camera networks (hybrid)
  – Monitoring outdoor traffic
- Process camera images to detect anomaly:
  – Modeling aspects:
    - Learn the fluctuation of traffic
    - Unsupervised learning
    - No apriori knowledge
    - Robust (noise)
  – Anomaly detection: real-time processing

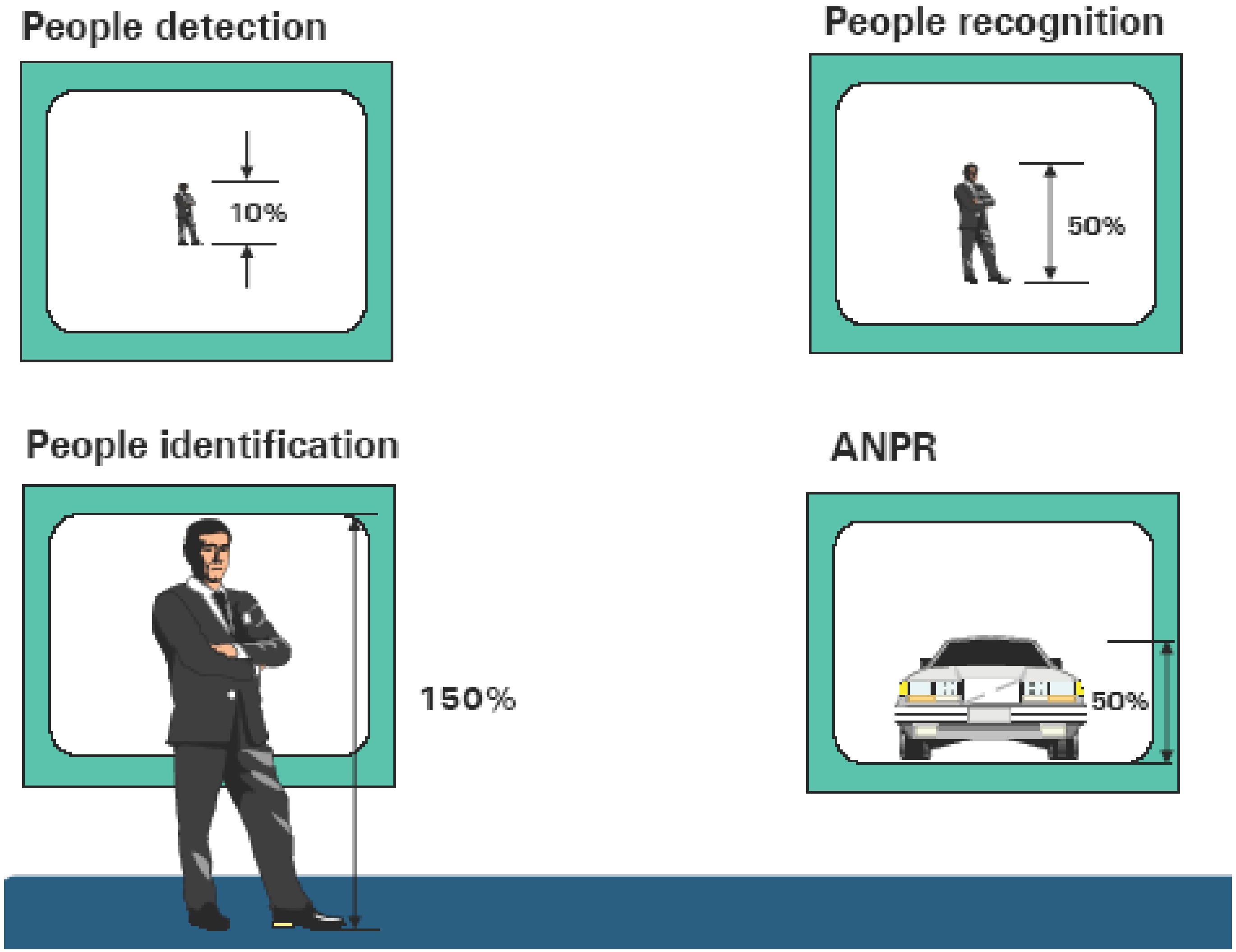
Our Motivation

"After 12 minutes of continuous video monitoring, an operator will often miss up to 45% of onsite activity. After 22 minutes of viewing, up to 95% is overlooked."

IMS Research

Typical video quality

What do we want to see?



Problems

- From the device:
  - Electronic noise
  - Optical distortion
  - Flicker
  - Auto whitebalance
  - Aliasing errors
  - Framedrop
  - etc.
- From the scene:
  - Weather conditions (rain, wind etc.)
  - Light conditions (flare, head lights, etc.)
  - Occlusion
  - Shadows
  - etc.

Conventional object tracking unreliable!

Problems

Device noise

Framedrop

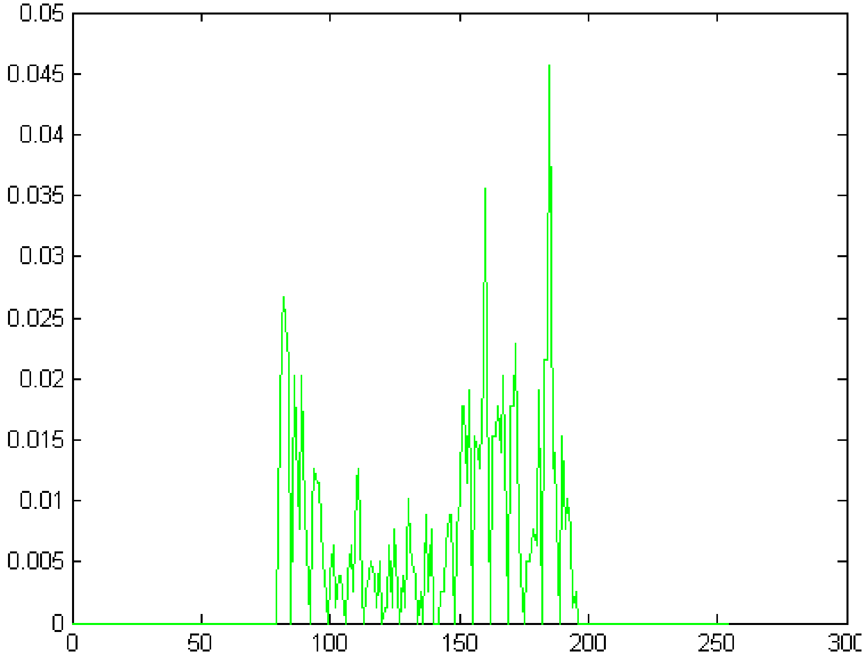
Occlusion, shadow

Object tracking?

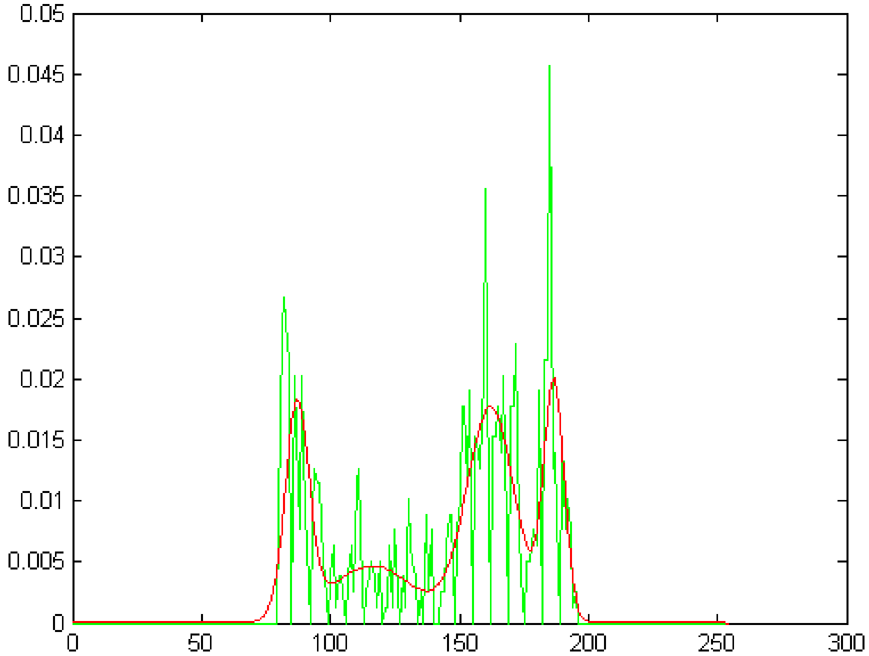
Occlusion/disocclusion… Noise… Ragged object masks… Shadows…

Basic concepts

- Mixture of Gaussians (MOG, GMM)



Histogram

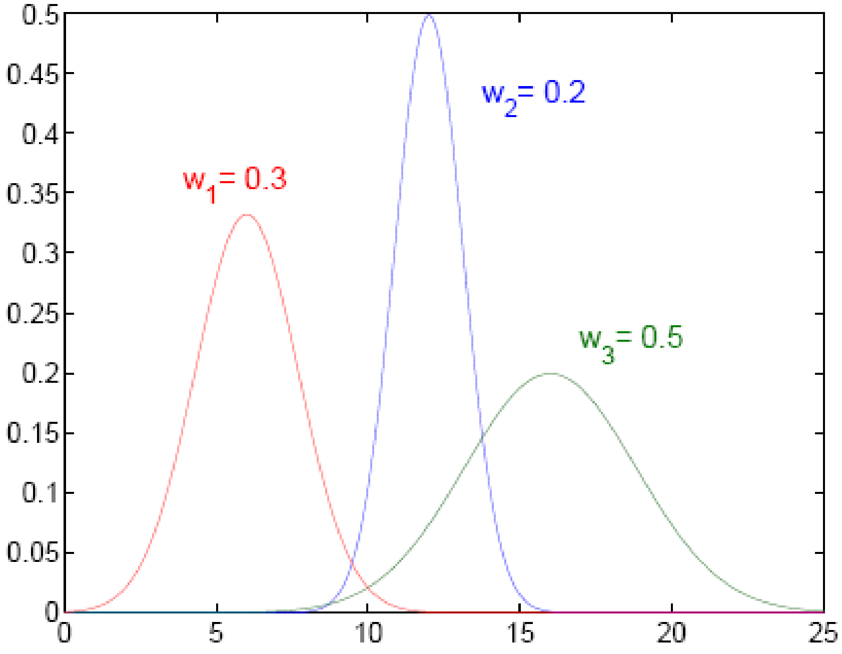


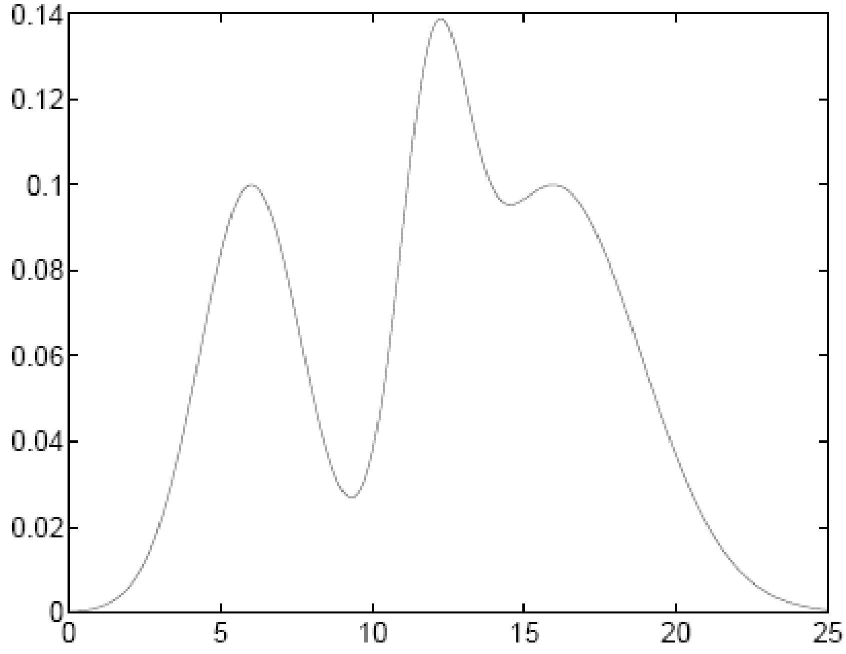
MOG (red)

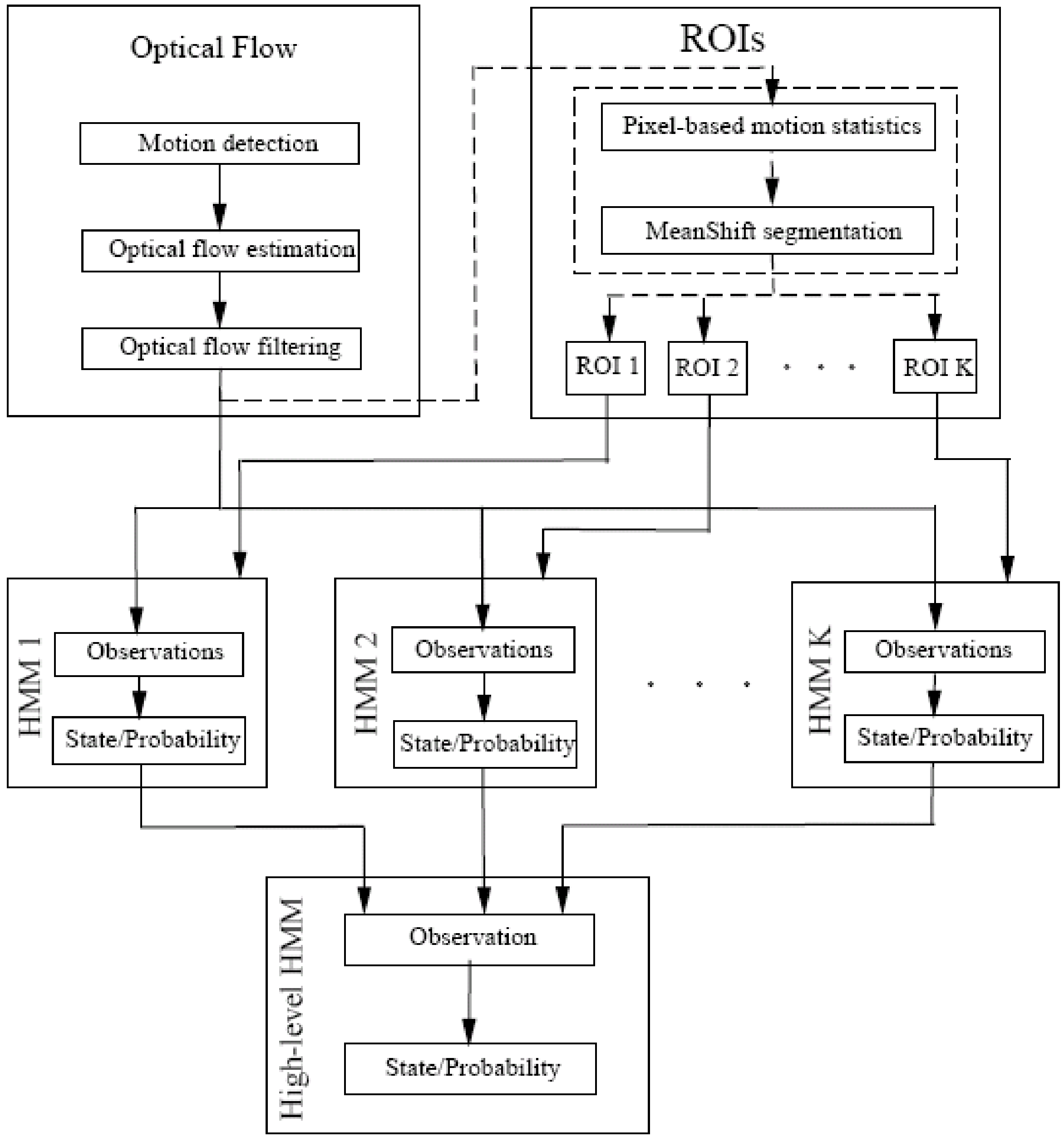
$$P_{MOG}(x) = \sum_{k=1}^{K} w_k \mathcal{N}(x|\mu_k, \Sigma_k)$$

Basic concepts

- Fitting Mixture of Gaussians
  - Expectation-Maximization algorithm (iterative)
  - Accurate but slow





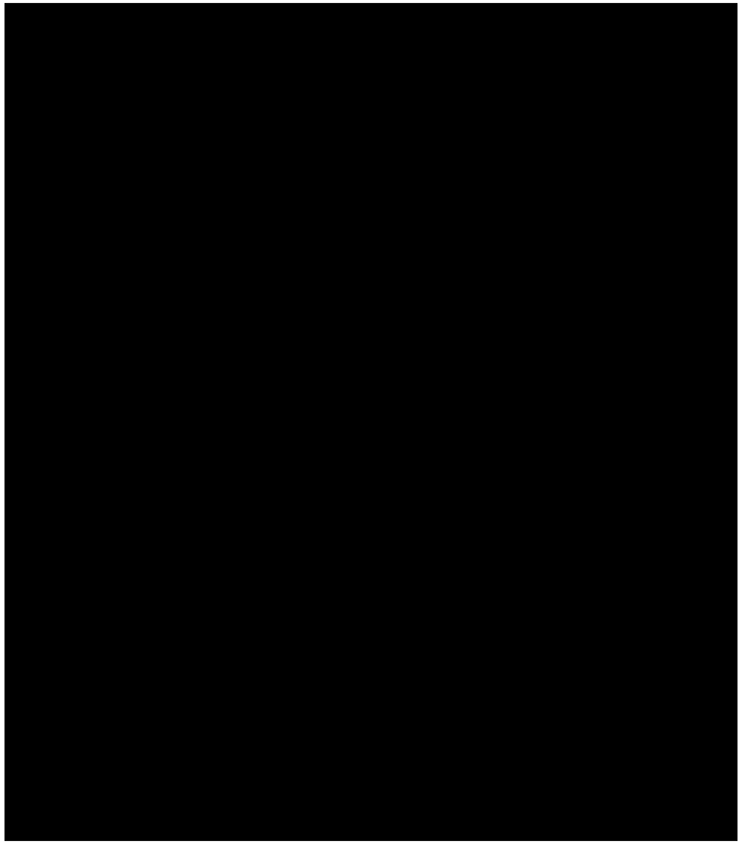

Optical Flow
Motion detection
Optical flow estimation
Optical flow filtering
ROIs
Pixel-based motion statistics
MeanShift segmentation
ROI 1
ROI 2
ROI K
HMM 1
Observations
State/Probability
HMM 2
Observations
State/Probability
HMM K
Observations
State/Probability
High-level HMM
Observation
State/Probability

Preprocessing

- Background-foreground separation:
  – Robust method of Stauffer and Grimson (MOG)
- Optical flow (e.g. Bergen, Lucas-Kanade):
  – Preferrably only over motion detected areas
  – Some filtering advised: drop very small and very large vectors
  – Noisy output
  – Real-time operation

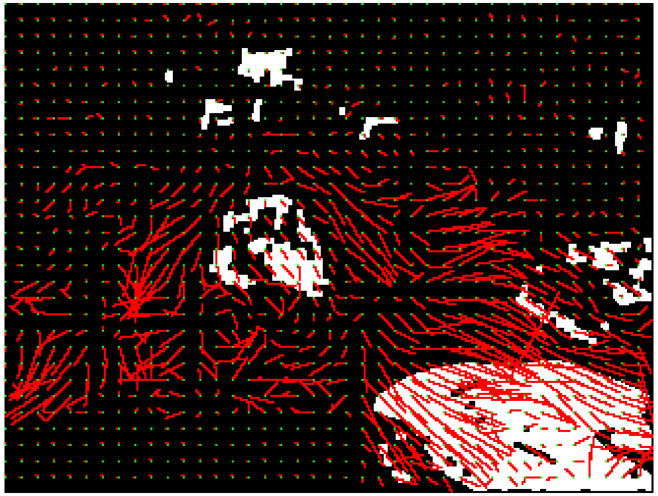
Example for motion detection

- Foreground-Background Segmentation based on

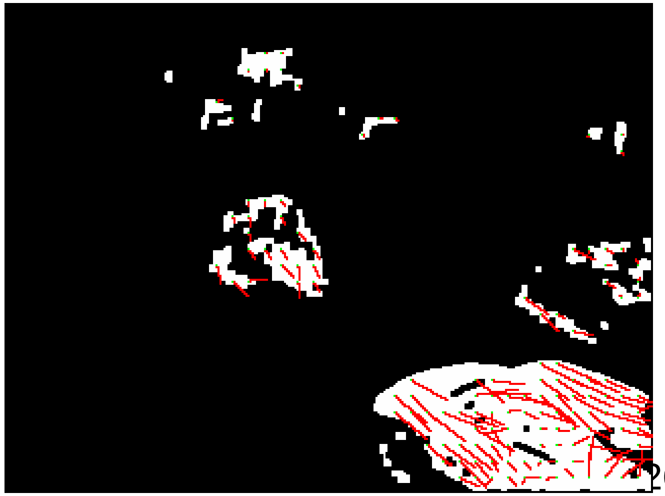


Optical flow

Original



Unfiltered



Filtered

Models without object level analysis

- We define a motion vector observation unusual if its probability is low according to prior observations.

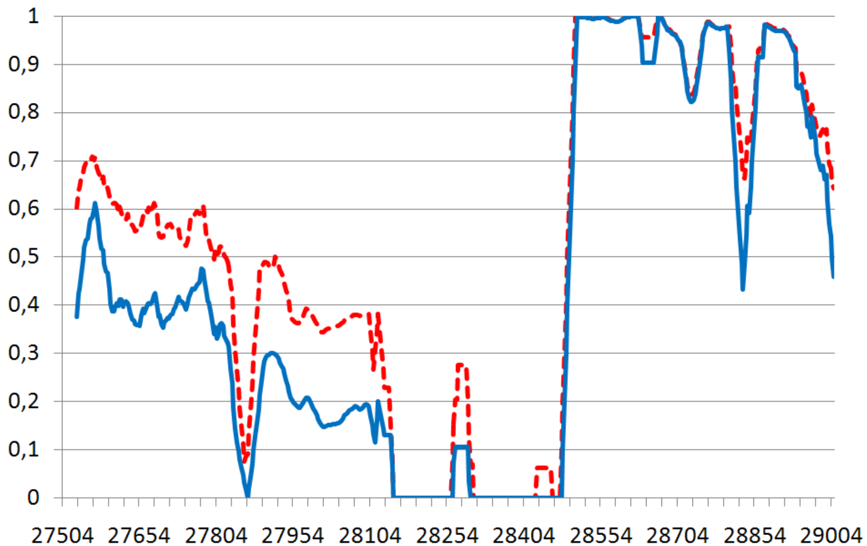
$$P_{Dir} = \left\| O_{Dir} \right\| \Big/ \sum_{Dir} \left\| O_{Dir} \right\|$$

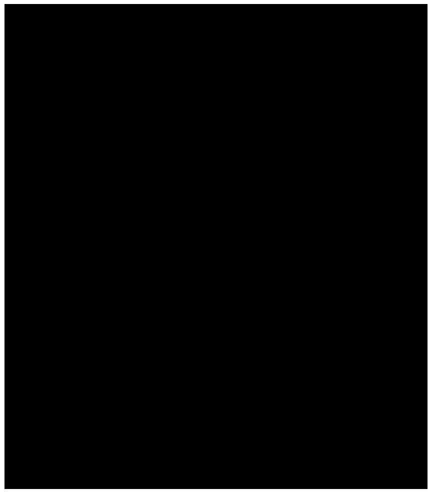
$$P^{(U)}{}_{Dir} = 1 - P_{Dir}$$

- Unsupervised learning.

- To get temporal support we can apply some Markovian assumptions:

$$P^{(U,M)}{}_{x,y,t} = P^{(U)}{}_{Dir,x,y,t} \cdot \max_{x',y' \in R} \{ P^{(U)}{}_{Dir,x',y',t-1} \}$$





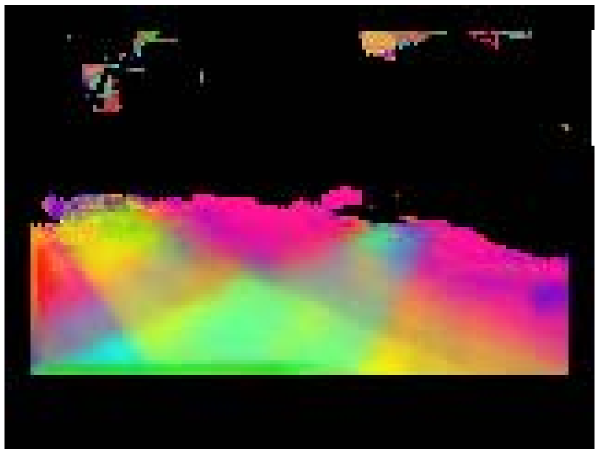
Cyclist in the wrong direction detected.

Preprocessing

- Collect motion direction statistics in pixels:
  - For a motion vector classify it's direction:

$$Dir \in \{N, E, S, W, NE, SE, SW, NW\}$$

  - Create 8-bin motion direction histograms in each pixel
  - Histogram $\longrightarrow$ empirical probability (left, mixing 8 colours)
- Construct regions from statistics:
  - MeanShift: Spatial distance + Histogram distance (right)



$$\overline{RGB}_{x,y} = \sum_{Dir} P_{E,x,y}(Dir)\overline{C}_{Dir}$$

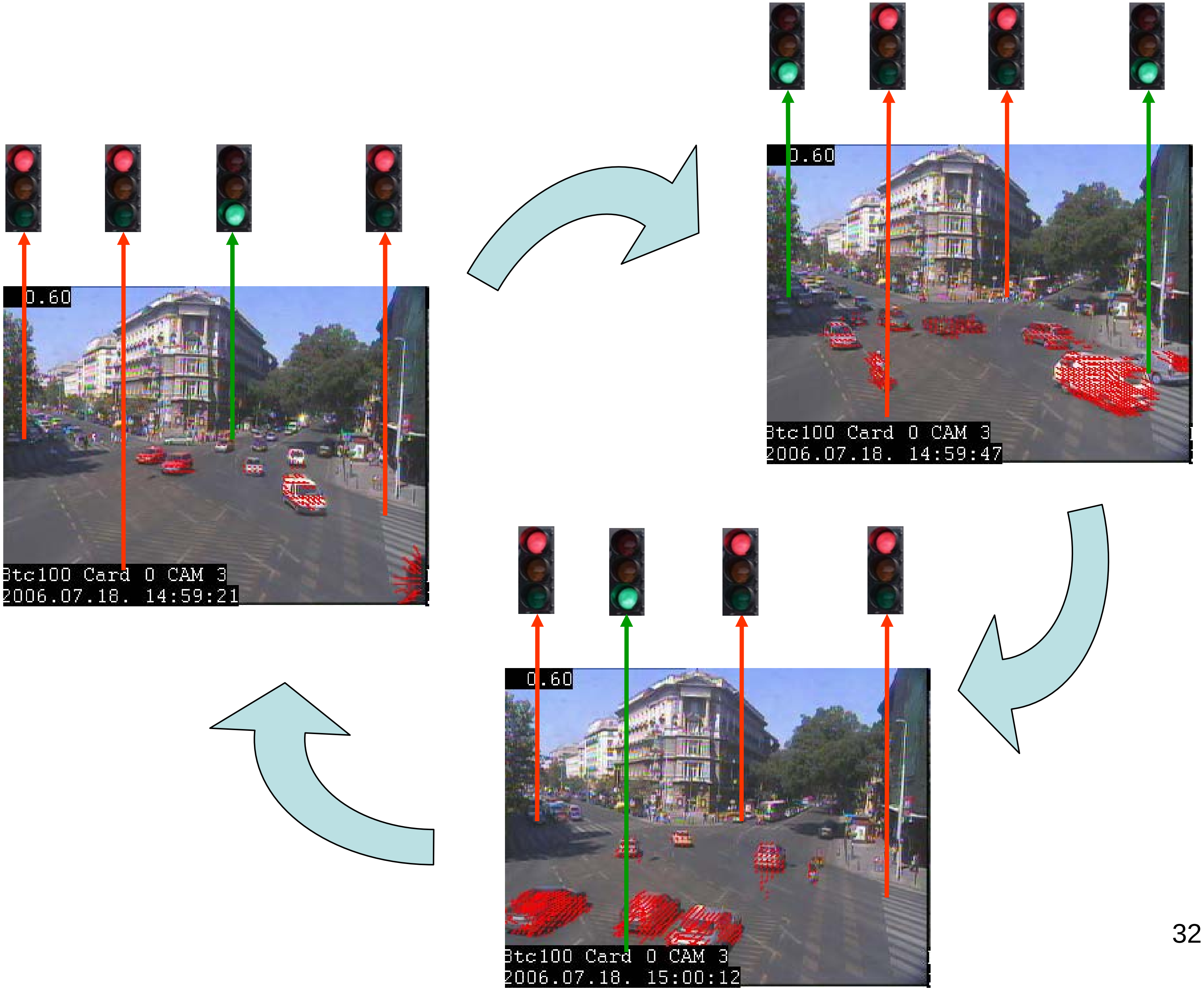
Segmentation $\longrightarrow$

Hidden Markov Model

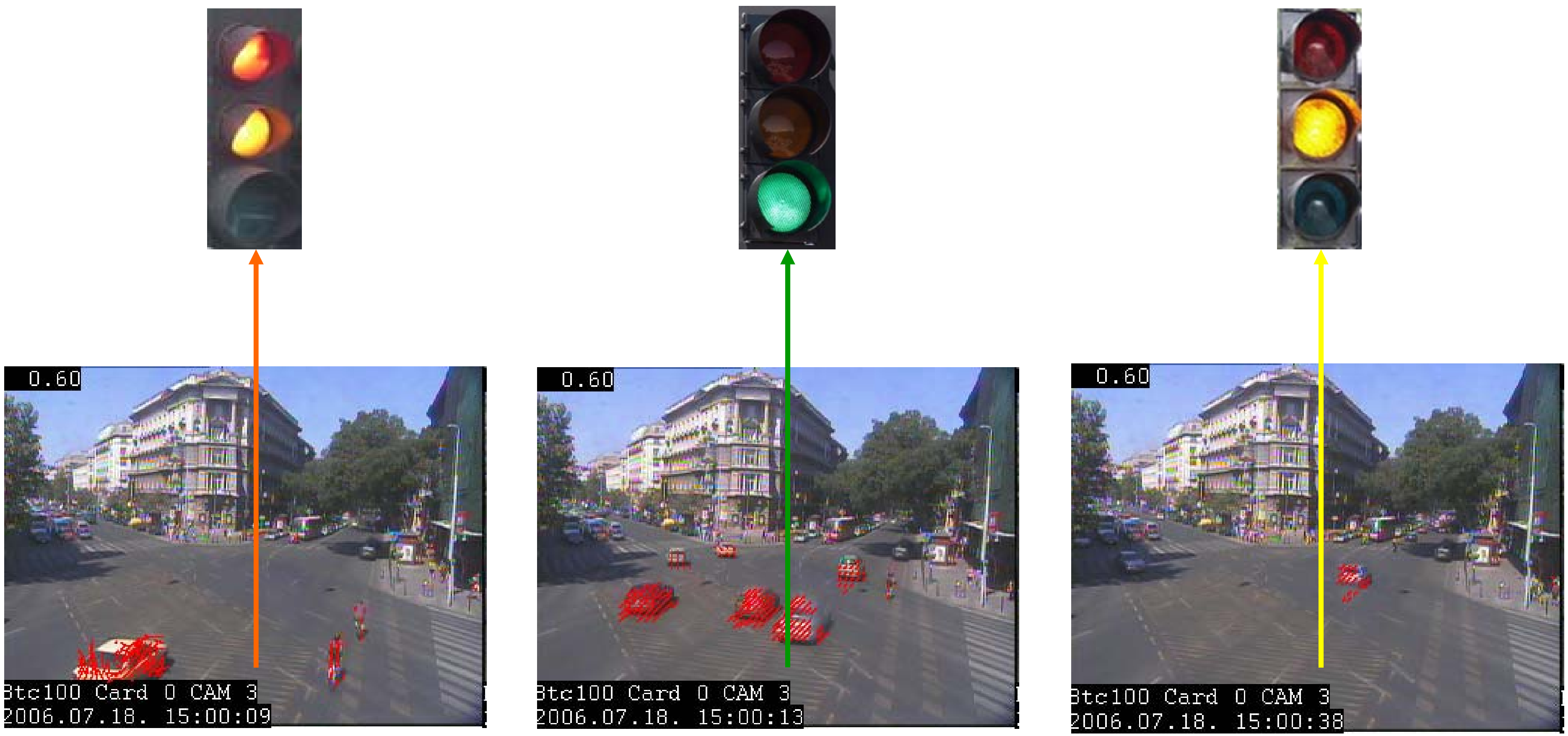
- What is a HMM?
  - A **system** which has finite number of **states** and certain rules (**transitions** with **Markov** property).
  - Process: the states are **hidden**, but the system generates an **observable** process.
- In our case:
  - **System** = traffic lamp system in the crossroad
  - **States** = traffic rules controlled by traffic lamp configurations (green, yellow, red)
  - **Transitions** = changes in the traffic lamp configurations
  - **Observation** = localized motion directions (x,y,d)

Model parameters

- Hidden Markov Model: $\lambda=(\pi, A, B)$
- Initial state probabilities (π): the probability that a process starts with a state
- Transition probabilities (A): the probability of changing to a state from the previous state
- Emission probabilities (B): the probability that a state generated a given observation


0.60
Btc100 Card 0 CAM 3
2006.07.18. 14:59:21
0.60
Btc100 Card 0 CAM 3
2006.07.18. 14:59:47
0.60
Btc100 Card 0 CAM 3
2006.07.18. 15:00:12

How many states?



Observation data

- Select a region (ROI)
- Observe moving blobs in the ROI
- Fit a MOG on motion directions in each blob
  - Only a few iterations (**real-time**!)
- **Observations** at time *t* = Mean directions of MOGs

Selected ROI

Emission probability

- We have K_t motion directions in time t, thus our observation in time t is $O_t=o_{t,1},\ldots,o_{t,K_t}$
- We use Mixture of M Gaussians, i.e.

$$b_i(O_t) = \prod_{k=1}^{K_t} b_i(o_{t,k})$$

$$b_i(o_{t,k}) = \sum_{l=1}^{M} w_{i,l} b_{i,l}(o_{t,k})$$

$$b_{i,l}(o_{t,k}) = \mathcal{N}(o_{t,k}|\mu_{i,l},\Sigma_{i,l})$$

Training HMM
(learning problem)

- Given an observation sequence $O=O_1,\ldots,O_T$
- Problem: How to adjust the model parameters π, A, B to maximize $P(O|\lambda)$?
- Expectation Maximization: using the iterative Baum-Welch re-estimation formula.

Precision Problem

- In the Baum-Welch algorithm the emission probabilities (b_j) of the observations are calculated to re-estimate the model parameters.
- The observations are heavily loaded with ***noise***, resulting in large covariances in the MOGs, resulting in ***very small probability values***.
- And b_j was defined as a product:

$$b_i(O_t) = \prod_{k=1}^{K_t} b_i(o_{t,k})$$

- Precision problem: the probabilities are small values and the product will head exponentally to zero!

Solution

- Scaling by relative emission:
  - Original emission: what is the probability that the state generated the observation?
  - Relative emission: what is the probability that the state generated the observation compared to the other states?

Original

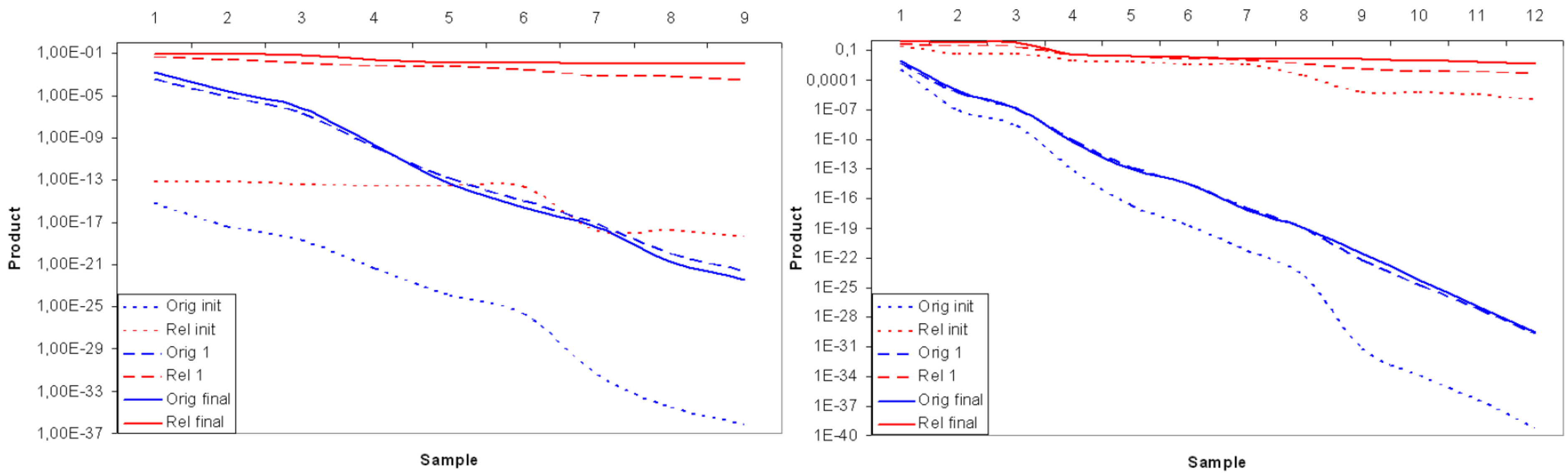
$$b_i(O_t) = \prod_{k=1}^{K_t} b_i(o_{t,k})$$

Relative

$$\tilde{b}_i(O_t) = \prod_{k=1}^{K_t} \frac{b_i(o_{t,k})}{\left[\sum_{j=1}^{N} b_j(o_{t,k})\right]}$$

- **The original Baum Welch re-estimation formula can be used with relative emissions!**

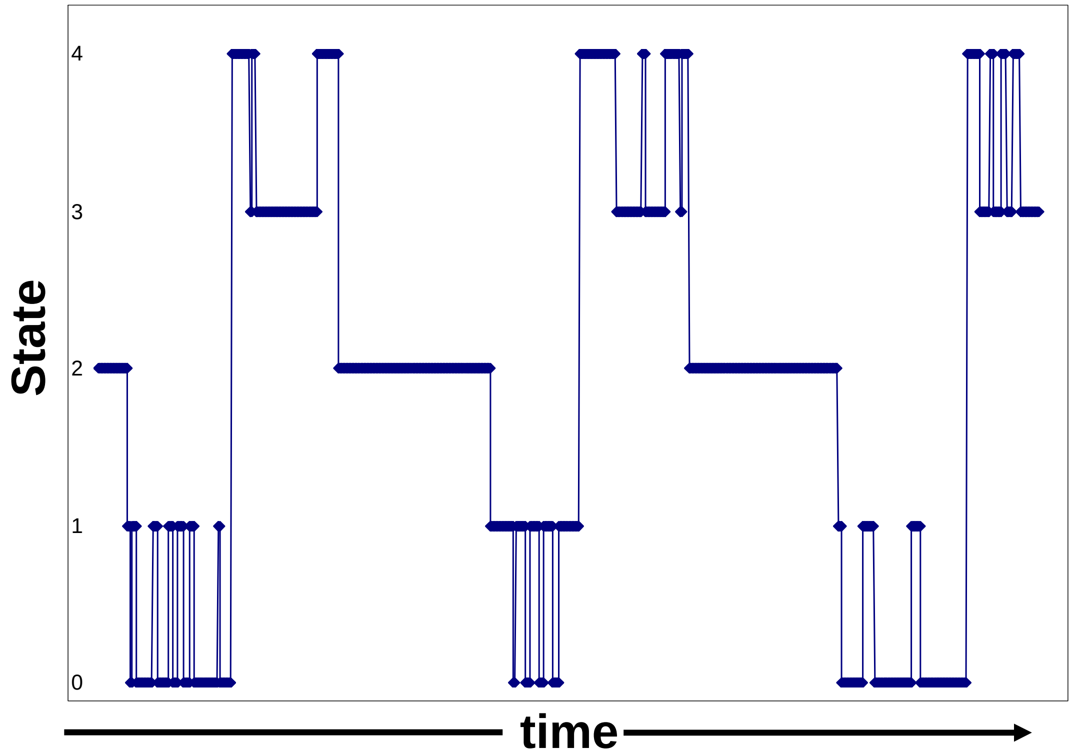
Effectiveness in training



- Horizontal: number of samples (in the product)
- Vertical: value of the product (logarithmic scale)
  - Blue: original emission probability
  - Red: relative emission

Detecting state sequence
(decoding problem)

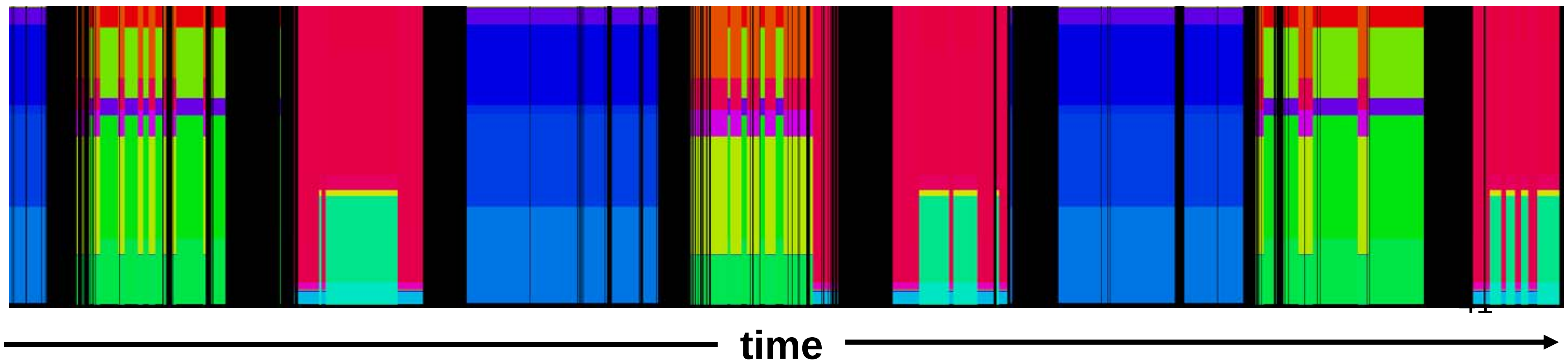
- Given the observation sequence $O=O_1,\ldots,O_T$
- What is the state sequence Q generated O?
- The Viterbi algorithm gives the answer



Detected state sequence

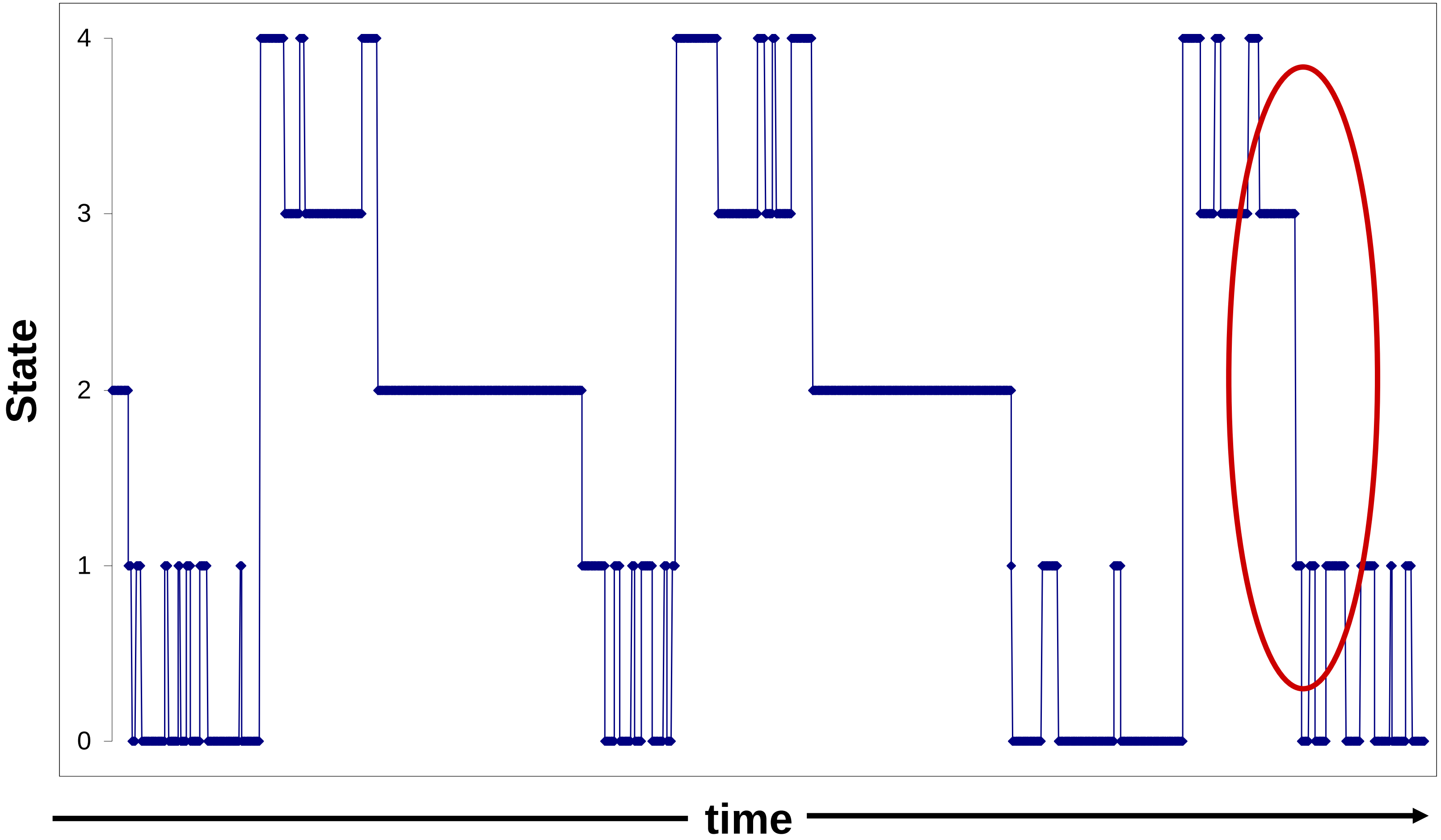
Visualization

- Take the generated state sequence
- Plot mean directions of the states on timeline using the HSV space (hue = direction angle)
- Height = weight of the component in MOG
- Black = no motion



Anomaly detection I

- Cut one phase from the video

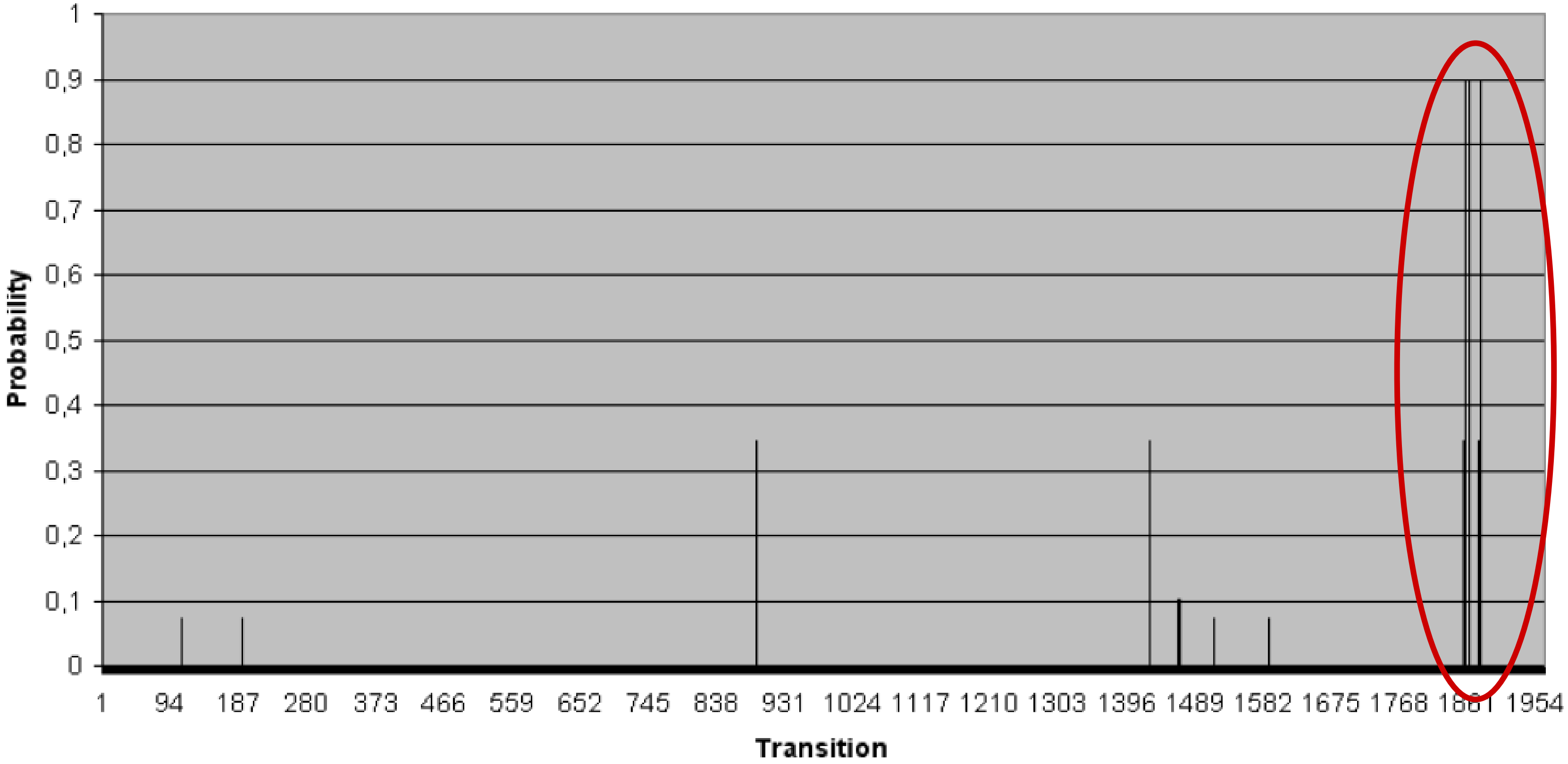


Anomaly detection II

- Car crossing the traffic

Anomaly detection II

- Generate state sequence for 3 non-empty frames
- Analyze the state transtions, plot on graph



Performance

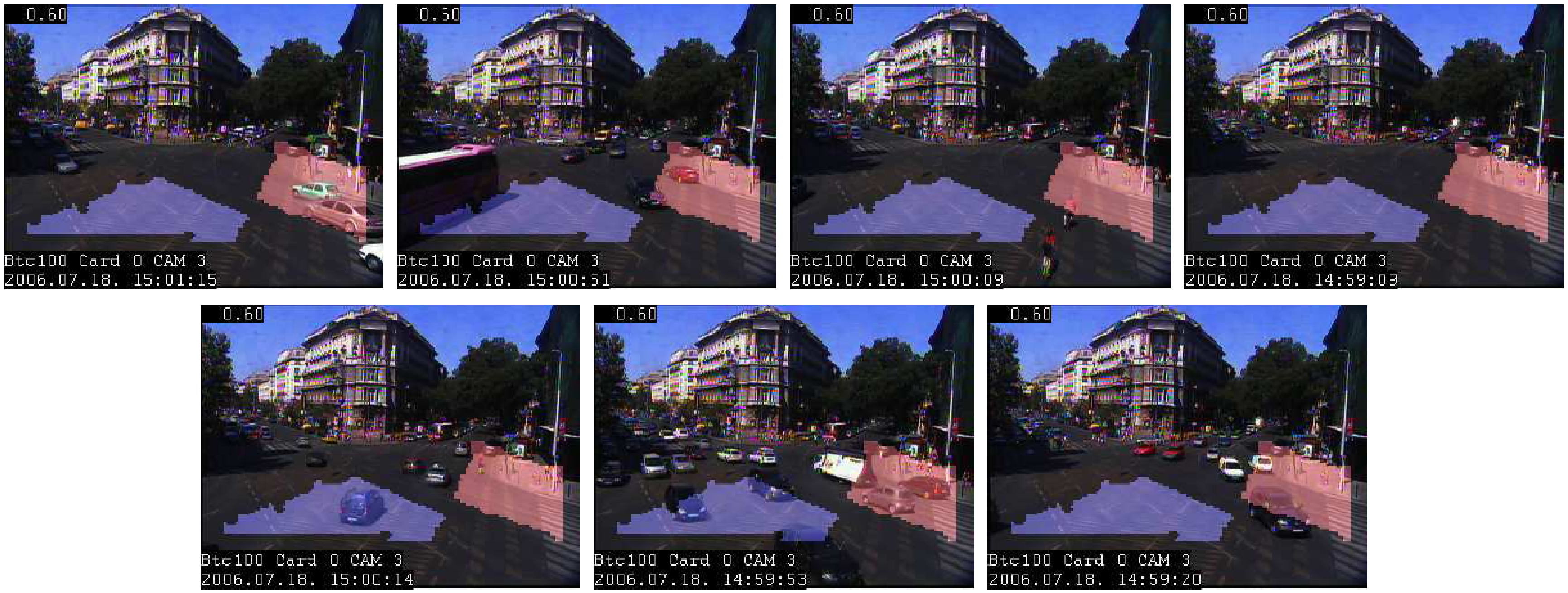
- 3 main phases:
  - **Preprocessing**: background-foreground separation + optical flow calculation and filtering, connected components
  - **Observation construction**: Fit MOG on components' motion directions inside ROI
  - **Anomaly detection II**

	Preprocessing	Observations	Detection	Total
Time (msec)	51.9	17.24	0.79	69.93

- Performance: 14 FPS on 160x120 video

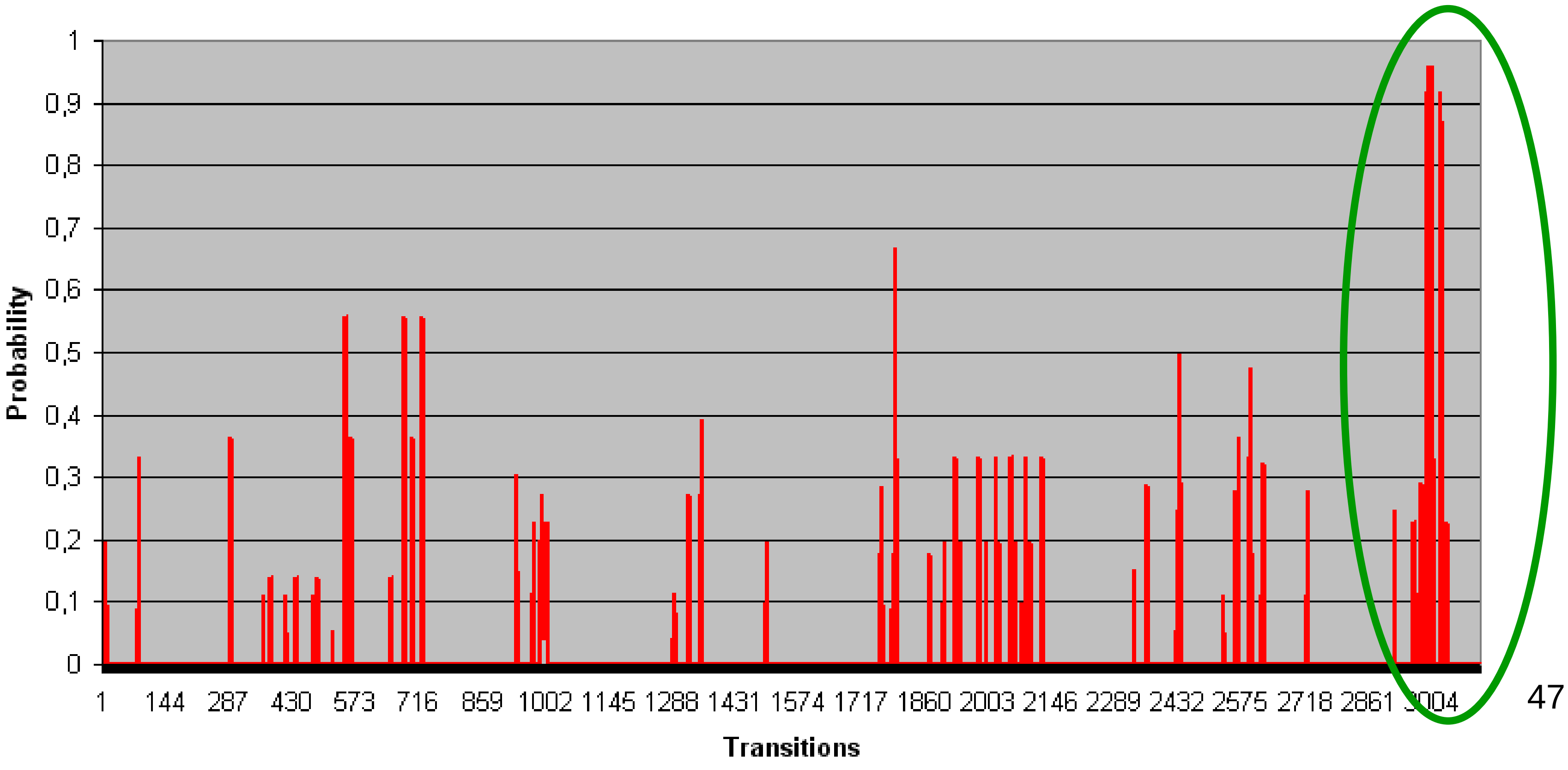
Hierarchical HMM

- Higher-level HMM, built on top of several ROI models
- Includes explicit modeling of „no motion”
- States after HMM training:



Anomaly detection III.

- Generate state sequence for 3 frames
- Analyze the state transtions, plot on graph



Anomaly detection III.

Detected frames

Conclusion

- Hidden Markov Model approach
  - Modeling motion directions in urban traffic
  - No tracking is necessary for anomaly detection
  - Solved a probability representation problem -> now it is possible to model the blobs with MOGs
- Hierarchical HMM approach
  - Explicit modelling of „no motion"
- Visualization of the traffic
- Anomaly detection
- Performance test

Publications

- Á. Utasi, L. Czúni, „HMM-based Unusual Motion Detection without Tracking”, 19th Int. Conf. on Pattern Recognition, Tampa, USA, 08-11 Dec 2008
- Á. Utasi, L. Czúni, *„Visual Analysis of Urban Road Traffic”*, 15th Int. Conf. on Systems, Signals and Image Processing, Bratislava, Slovakia, 25-28 June 2008
- Z. Szlávik, L. Kovács, L. Havasi, Cs. Benedek, I. Petrás, Á. Utasi, A. Licsár, L. Czúni, T. Szirányi, *„Behavior and Event Detection for Annotation and Surveillance”*, 6th Int. Workshop on Content-Based Multimedia Indexing, London, UK, 18-20 June 2008.
- Á. Utasi, L. Czúni, *„Anomaly Detection with Low-level Processes in Videos”*, 3rd Int. Conf. on Computer Vision Theory and Applications 2008, Madeira, Portugal, 22-25. Jan 2008
- Á. Utasi, L. Czúni, „*Unusual Event Detection in Low-Quality Urban Surveillance Videos with Modelling Motion Directions*”, Asia-Pacific Workshop on Visual Information Processing, Tainan, Taiwan, 15-17. Dec. 2007