BOSTON UNIVERSITY

http://iss.bu.edu/srv

Video Analytics for Multi-Camera Systems

Venkatesh Saligrama

Boston University

Faculty: P.M. Jodoin, J. Konrad

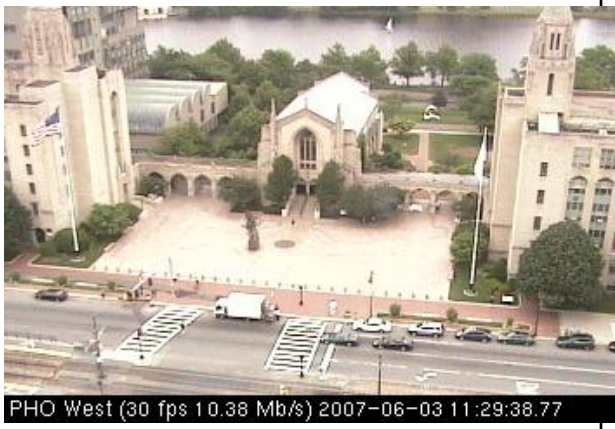
Students: E. Ermis, K. Rana, H. Ozkan @ BU

P. Clarot, Y. Benezeth @ U.Sherbrooke

Video Analytics for Camera Webs: The Need

"30 million cameras produce 4 billion hours of video footage each week in US alone"

Popular Mechanics, Jan. 2008

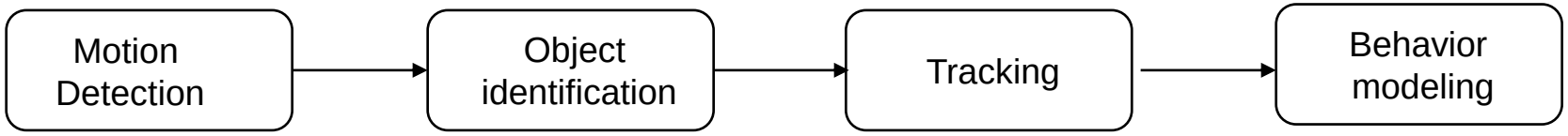


- Most video footage stored but rarely analyzed:
  - high cost + operator boredom & fatigue
- Source of information that should be tapped:
  - Alerts: real-time analysis, Post-event: video forensics
- BU Testbed: 12 Cameras, 2 PTZ, Networked, Web interface

What is Video Analytics?

- Know what you are looking for --- list of bad stuff!!! Search for it…
  - Intrusion detection, License plate ID, Facial ID, abandoned objects etc.
- Do not know what to look for: Have samples of what is good!!
  - Identify Unpredictable Behavior
    - Identifying suspicious movement of people or assets
    - Large crowd; Sudden traffic congestion, …
- Estimate Macroscopic Behavior
  - Counting pedestrian, traffic, long-term trends in behavior,
- Practical Aspects: Multi-Camera System
  - Autonomous, Removal/Insertion of assets, PTZs, Heterogenous Assets

Previous Work & Current Architecture



Two families of video analytics methods for surveillance:

- Object Based methods
  - Methods for which we exactly know in advance what is to be detected
  - Learning-based methods
    - Estimate ``normal" behavior of objects from training data

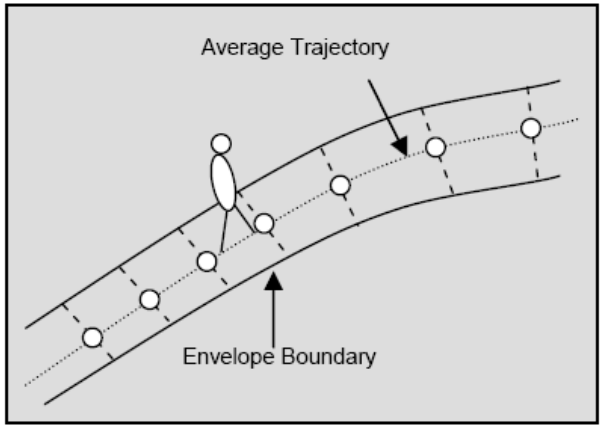
Know in advance what is to be detected

Face matching

Do not Know what you are looking for: Learn Normal Tracks

Normal people/cars are tracked during a training phase. Unusual people are those with an **unusual path**.

[Johnson and Hogg. 95]

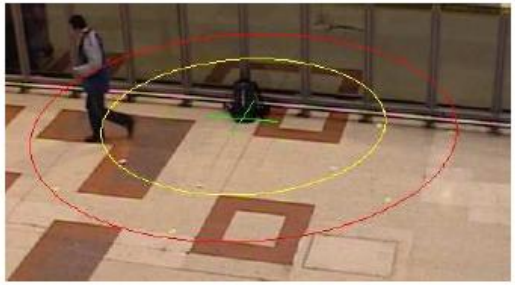


[Junejo et al. 04]

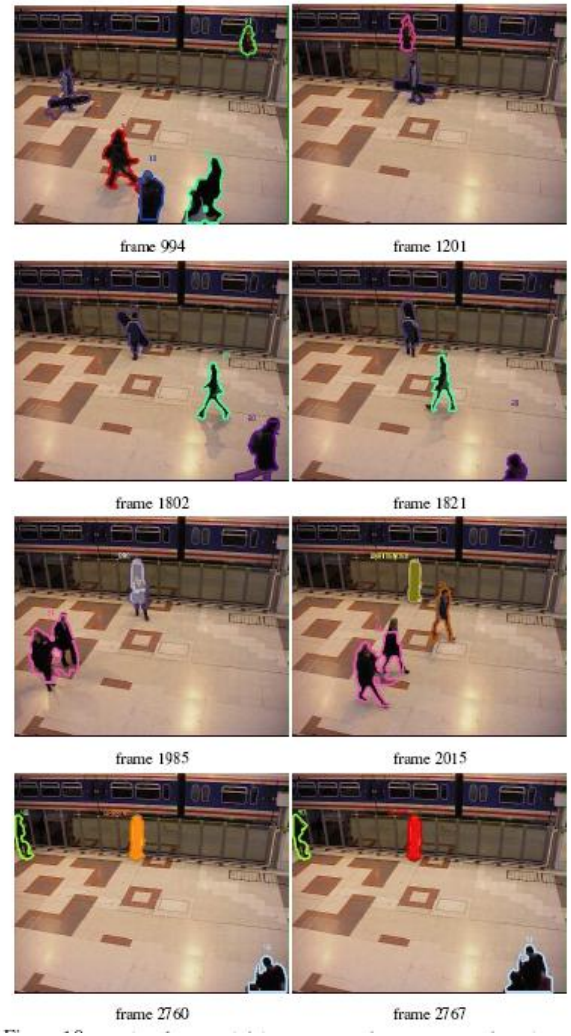
[Makris and Ellis 05]

Know in advance what is to be detected

Abandoned object detection

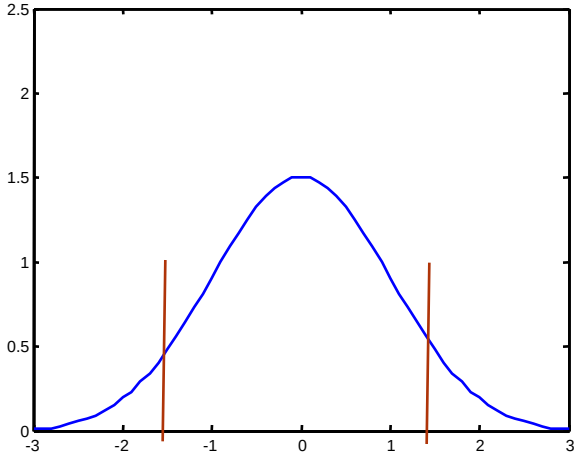


Smith et al. 06

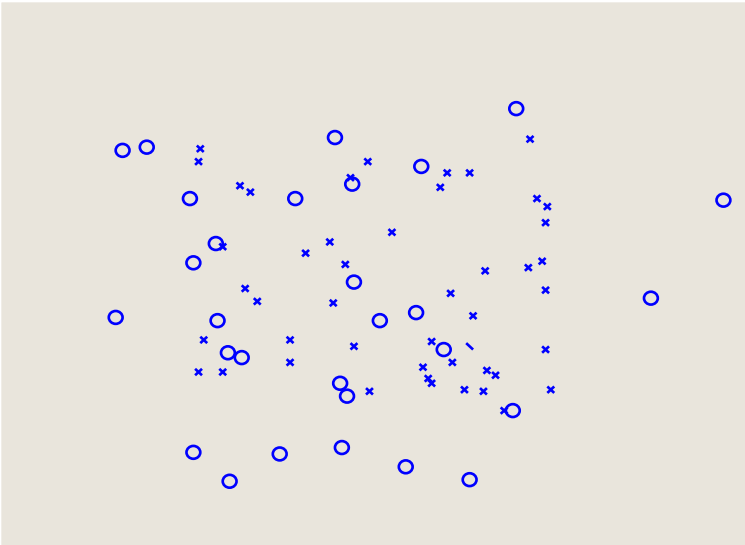


Learning Normal Behavior

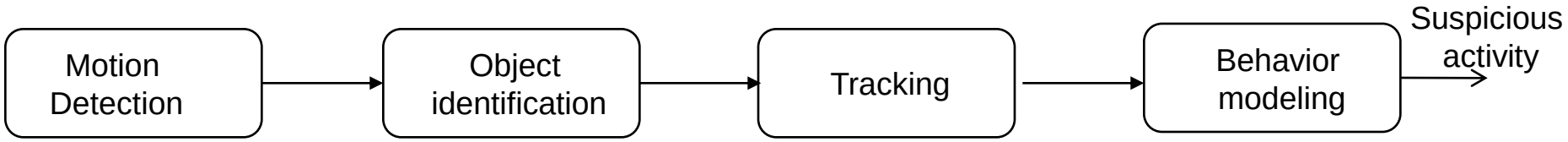
- Our Focus: Motion Anomaly + Don't know what to look for



- What features X to use?
  - Size, Color, Tracks?
- How are they distributed?
  - PDFs, Outliers, ….



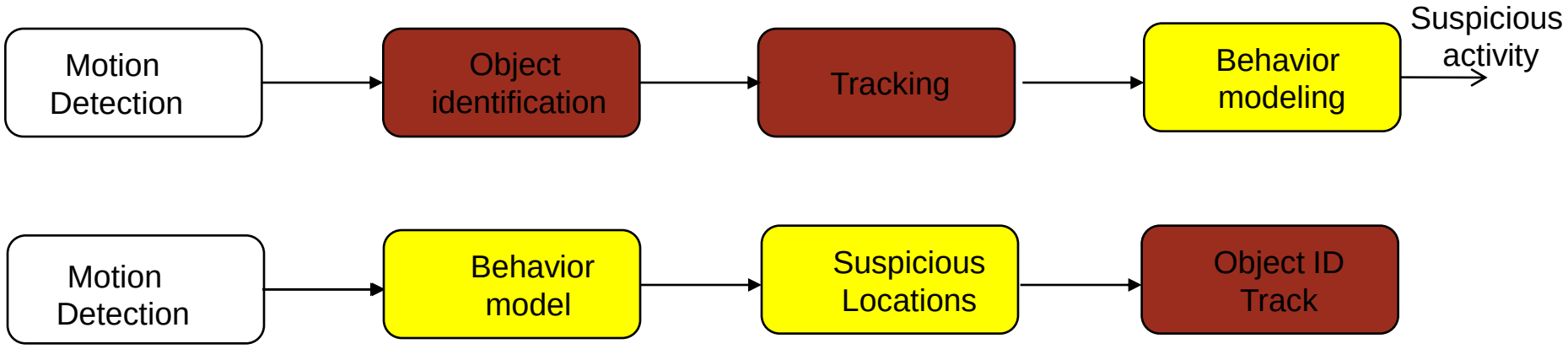
Object/Track as Features



- Object Based Approach
  - Locate, Tag & Track every object
  - Classify Outliers
  - Features are tracks

- Issues
  - Cascaded failure
  - Does not scale with cluttered urban scenario
  - Inconsistent with multi-camera views

Main Idea & Outline



- Tracking/identifying every object is very difficult
  - Locate/track/classify only suspicious objects?

- ``Location-based behavior Modeling" Approach
  - What is (un)expected behavior at a ``location"?
  - Algorithms based on activity patterns
  - Multi-Camera Fusion

Our method

Like most learning methods, our approach works in two phases :

- **Training phase**:
  - learn background activity
  - What is expected at a location?
- **Detection phase**:
  - detect foreground activity
  - Is it unexpected?
- Do not use tracks but low-dimensional features!!

Normal Behavior: What to learn in training phase?

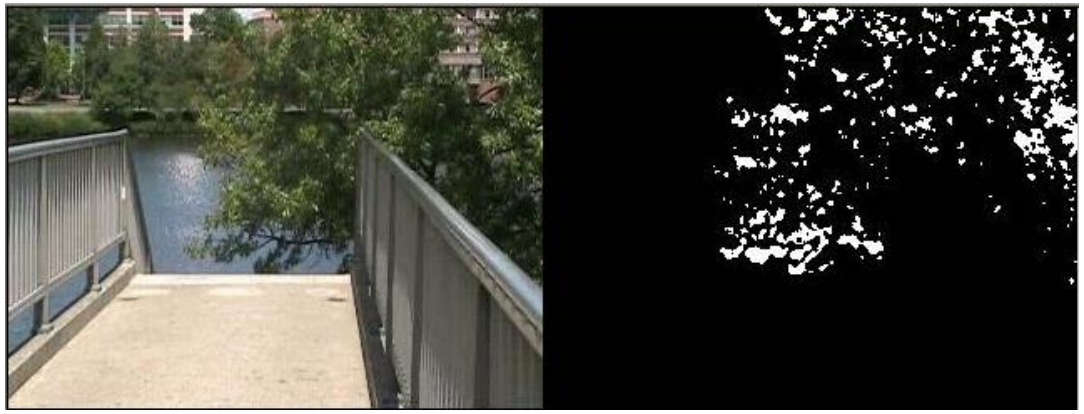
Many surveillance video sequences are made of repetitive background activities such as:

Highway traffic

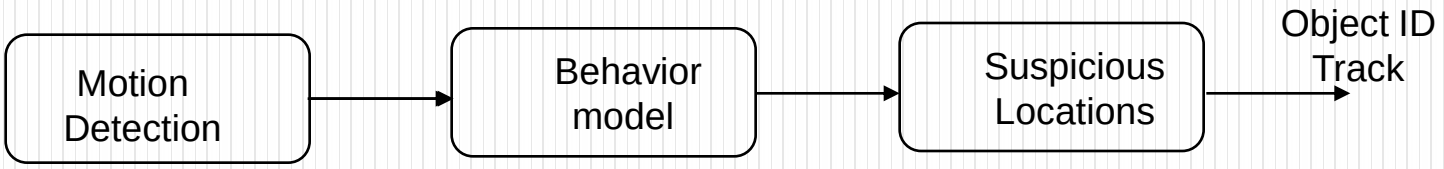
Input video | **Detected motion**

Leaves in the wind

Input video | **Detected motion**

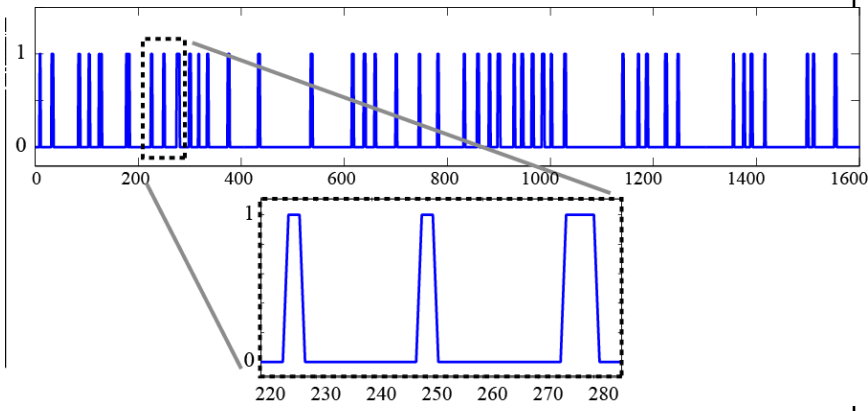


Key Steps


Motion Detection
Behavior model
Suspicious Locations
Object ID Track

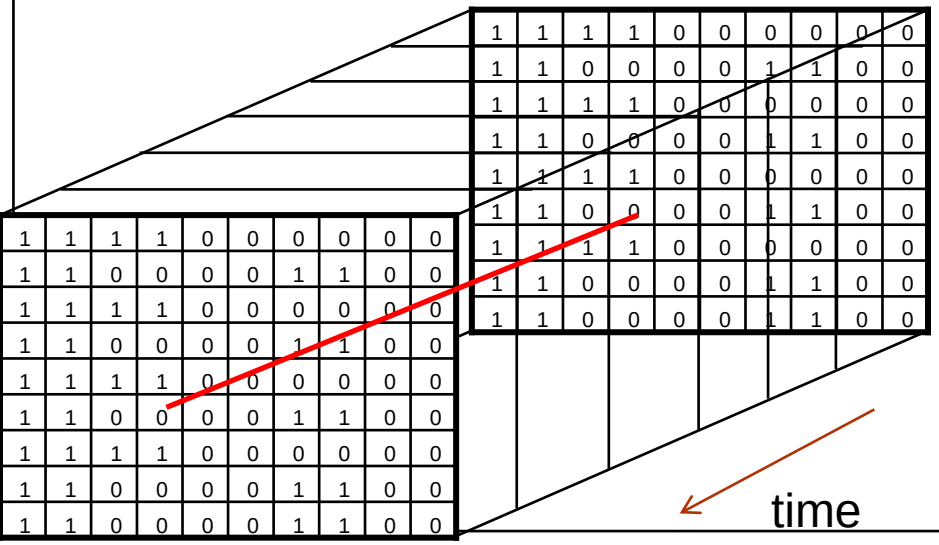
Frames to Detected Motion

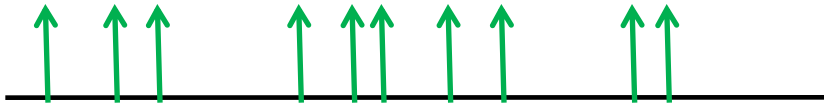
Motion to Event Representation



Temporal plot of one pixel

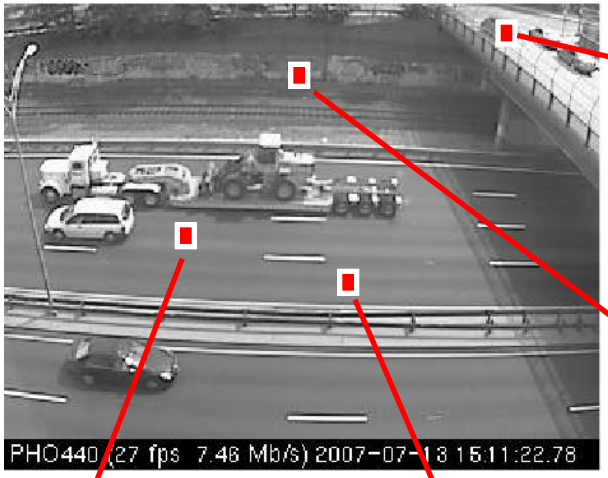
Binary Matrix of Motion Labels





Event Train
Onset of busy times

Events to Pixel Activity Model



$P(A(\vec{x}))$

A

$P(A(\vec{x}))$

A

$P(A(\vec{x}))$

A

$P(A(\vec{x}))$

A

2-D Normal Activity Model

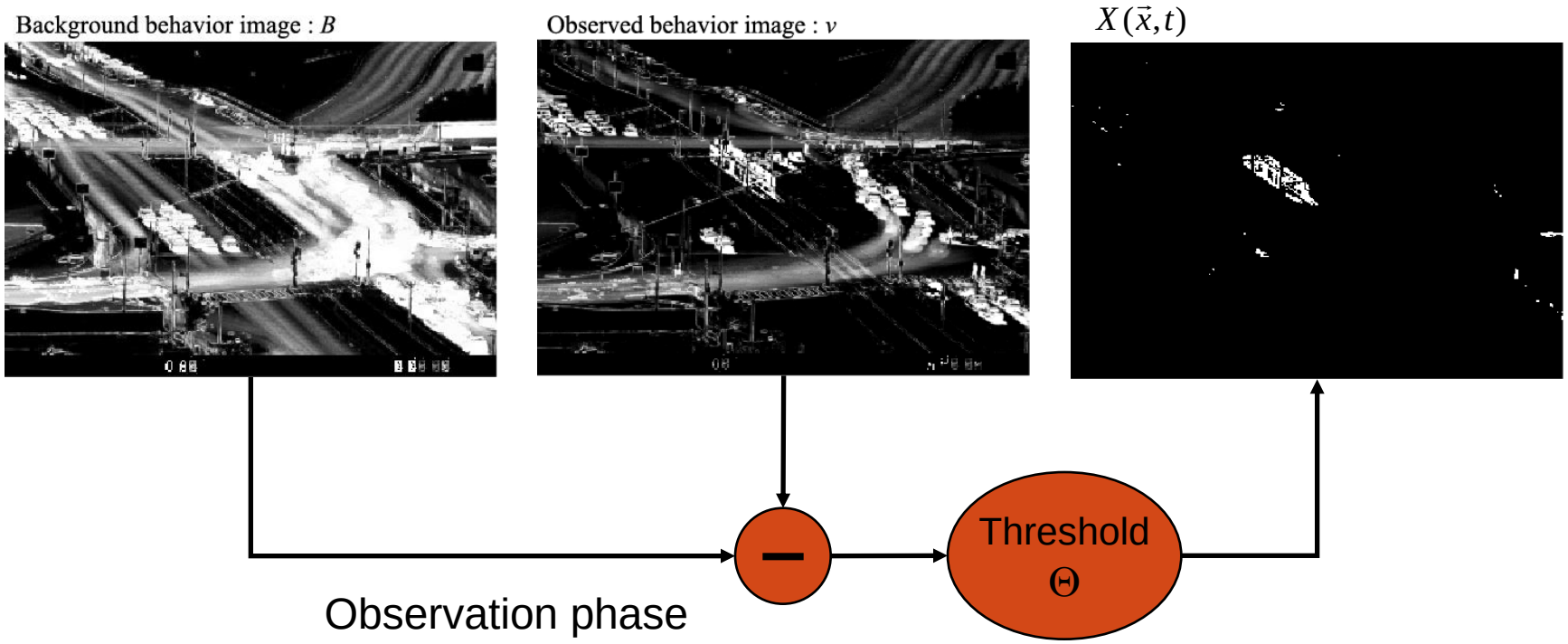
Training frame : $\hat{I}$

Background behavior image : B

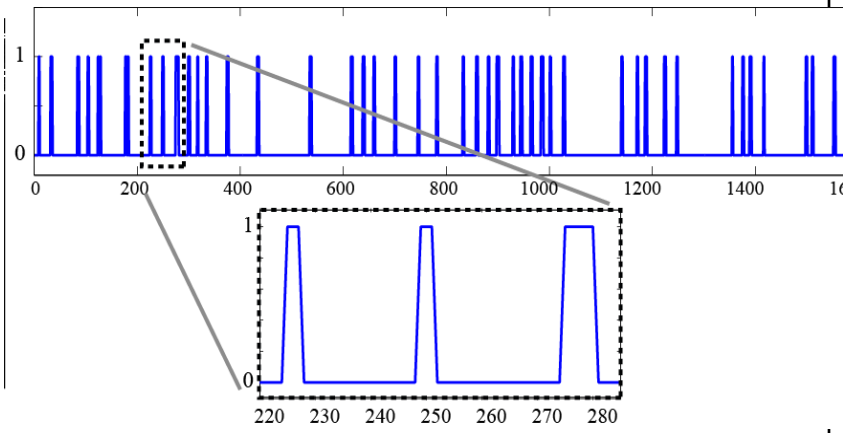
B is called the “Behavior background image”.

Behavior Subtraction: Anomaly Detection

Our method can be seen as a background subtraction approach.

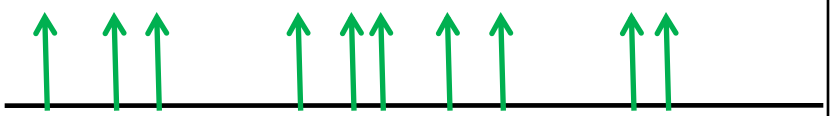


Key Technical Tools



- Define Behavior & Events at pixel p,
  - $\Lambda_p(j) = (B_j, I_j)$, $j=1, 2, \ldots$

- Behavior → Events → Renewal Processes
  - IID Sequence → Learning Theory
  - Size, color, shape conditioned on activity



Event Train

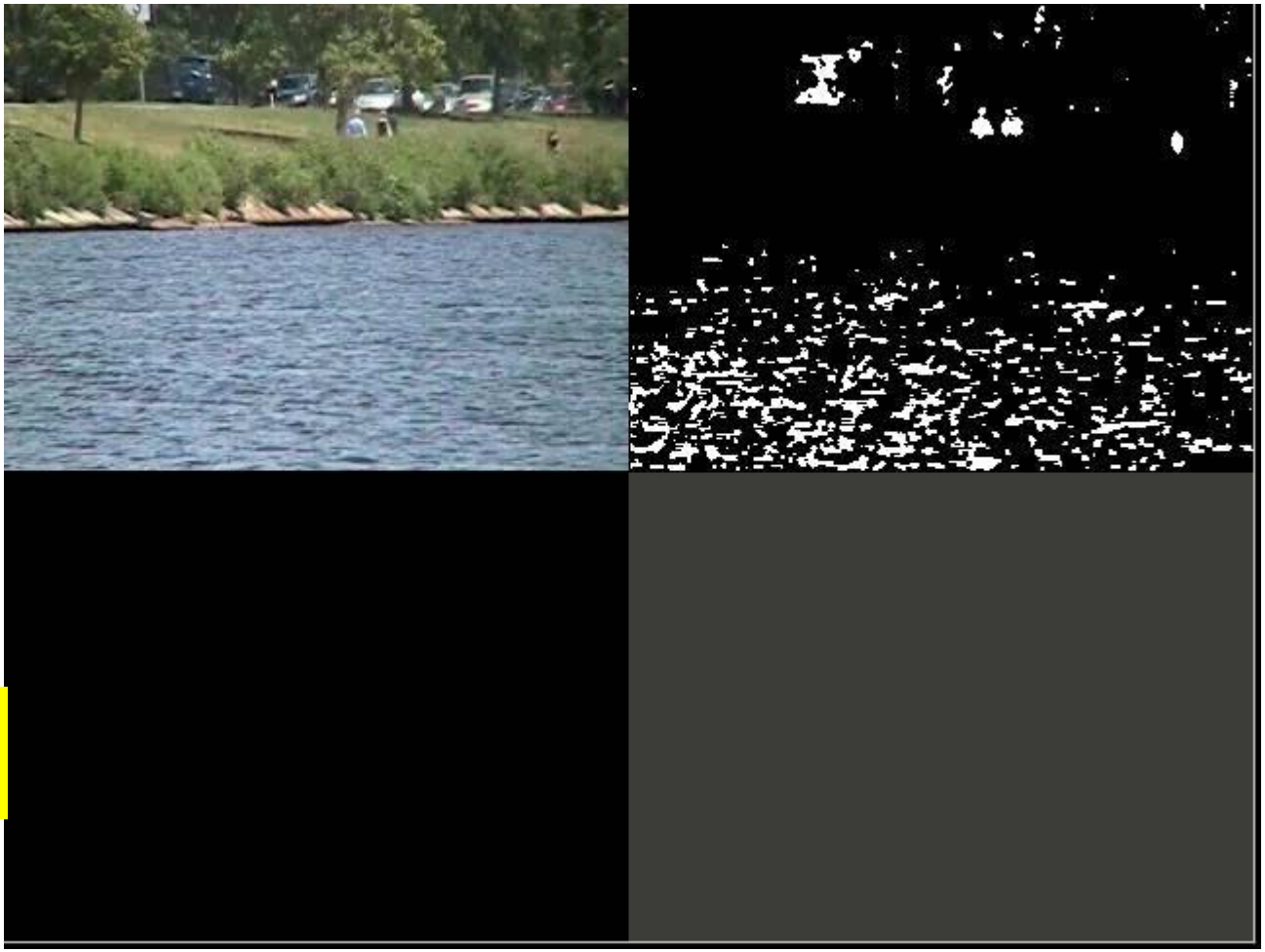
- Geometric Invariance
  - Busy-idle sequence invariant to zoom & angular displacement

- Uniqueness: Two locations with different behaviors are independent

Robust Intrusion Detection

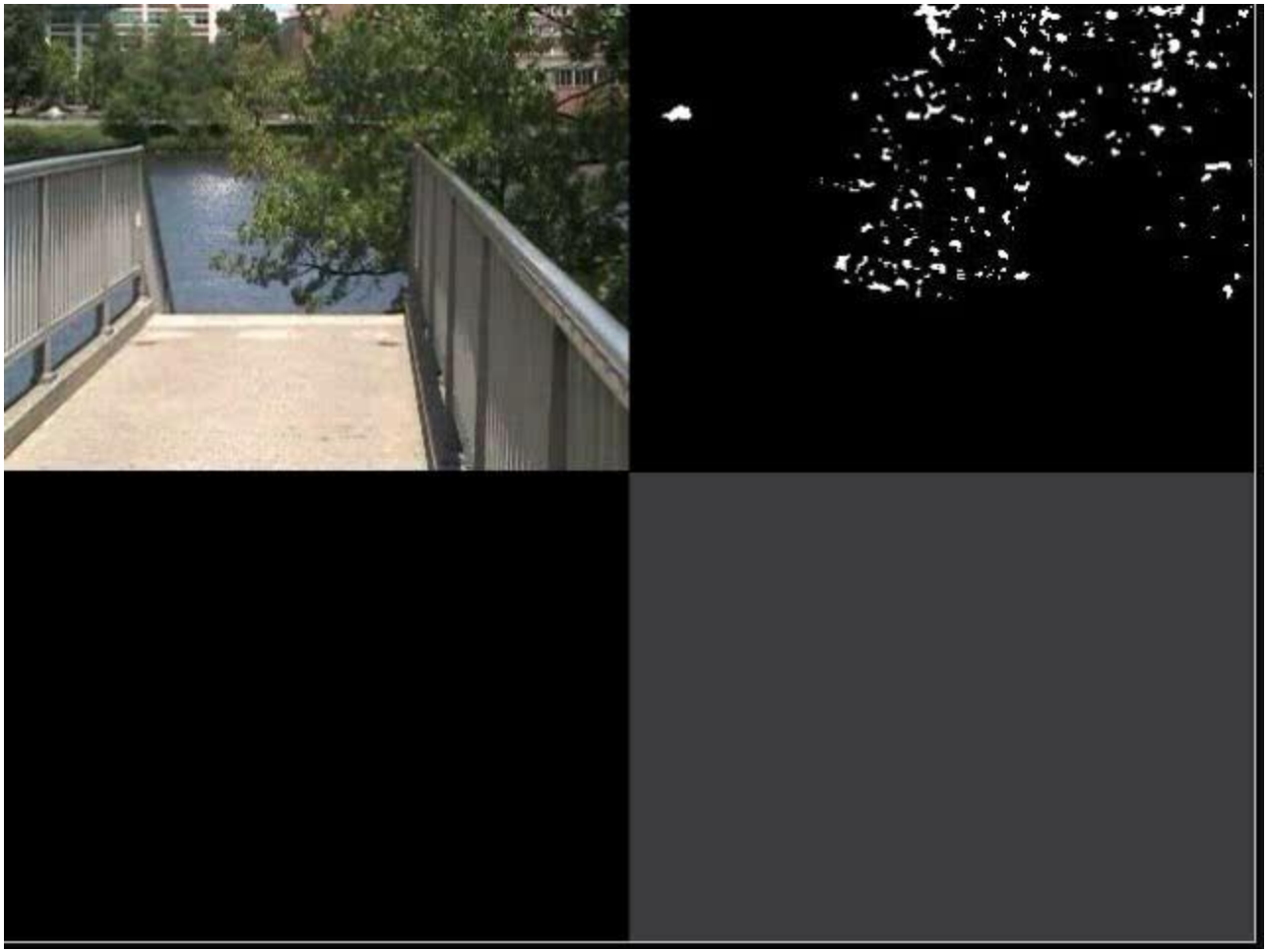
BOSTON UNIVERSITY

Scenario: Sailboat on Charles River



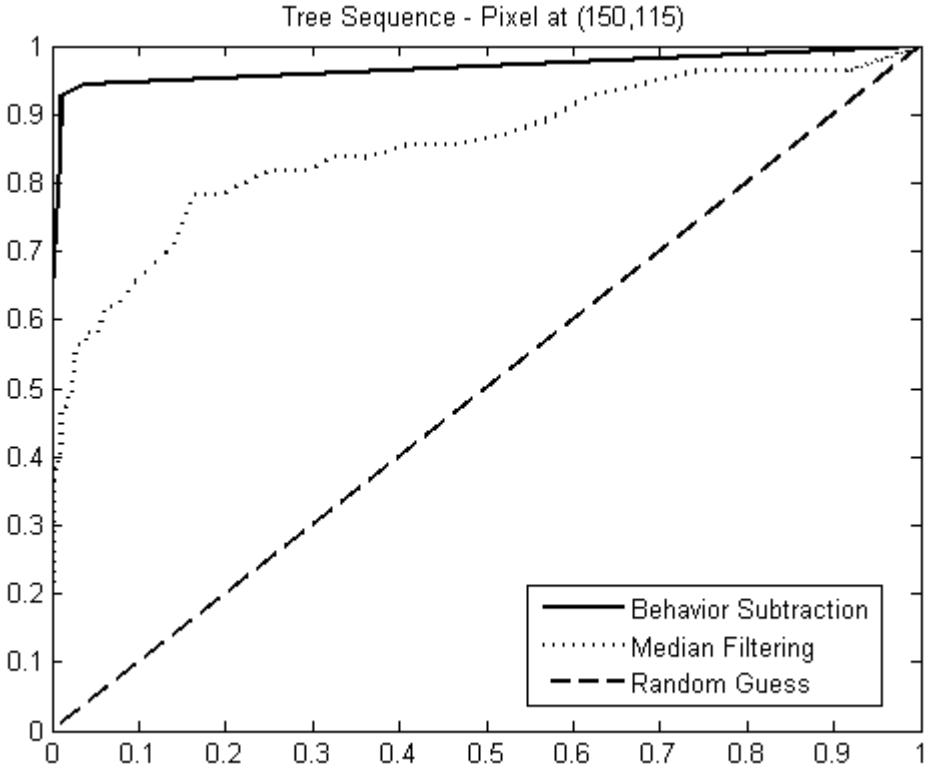
ABNORMAL VIDEO

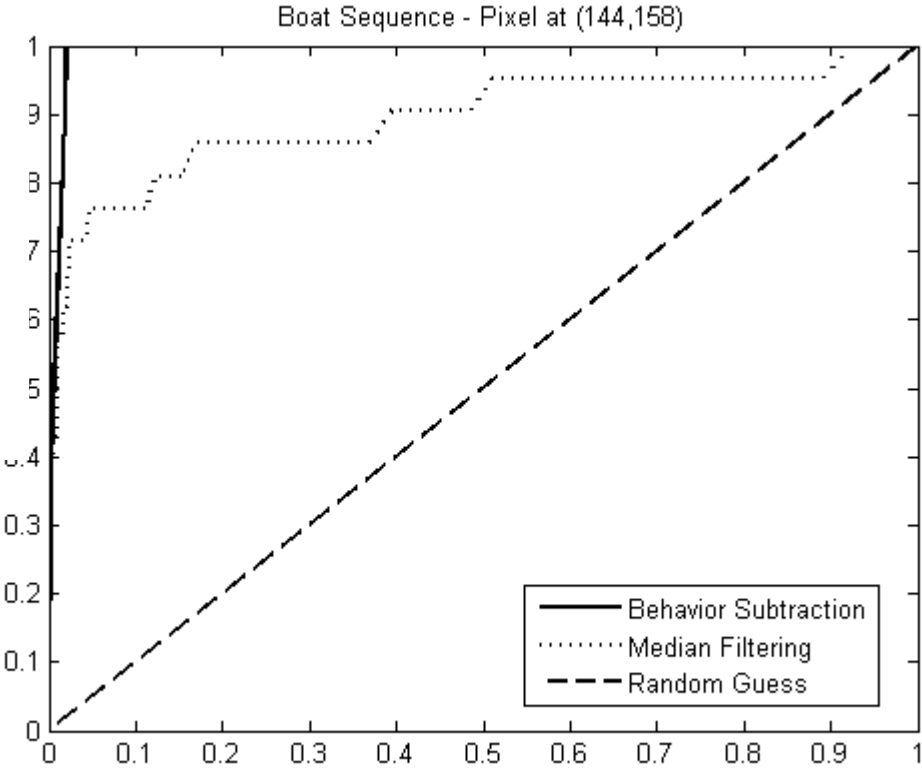
Scenario: Shaking Trees & River



ABNORMAL VIDEO

ROC Curves: Intrusion Detection in Cluttered Scenes

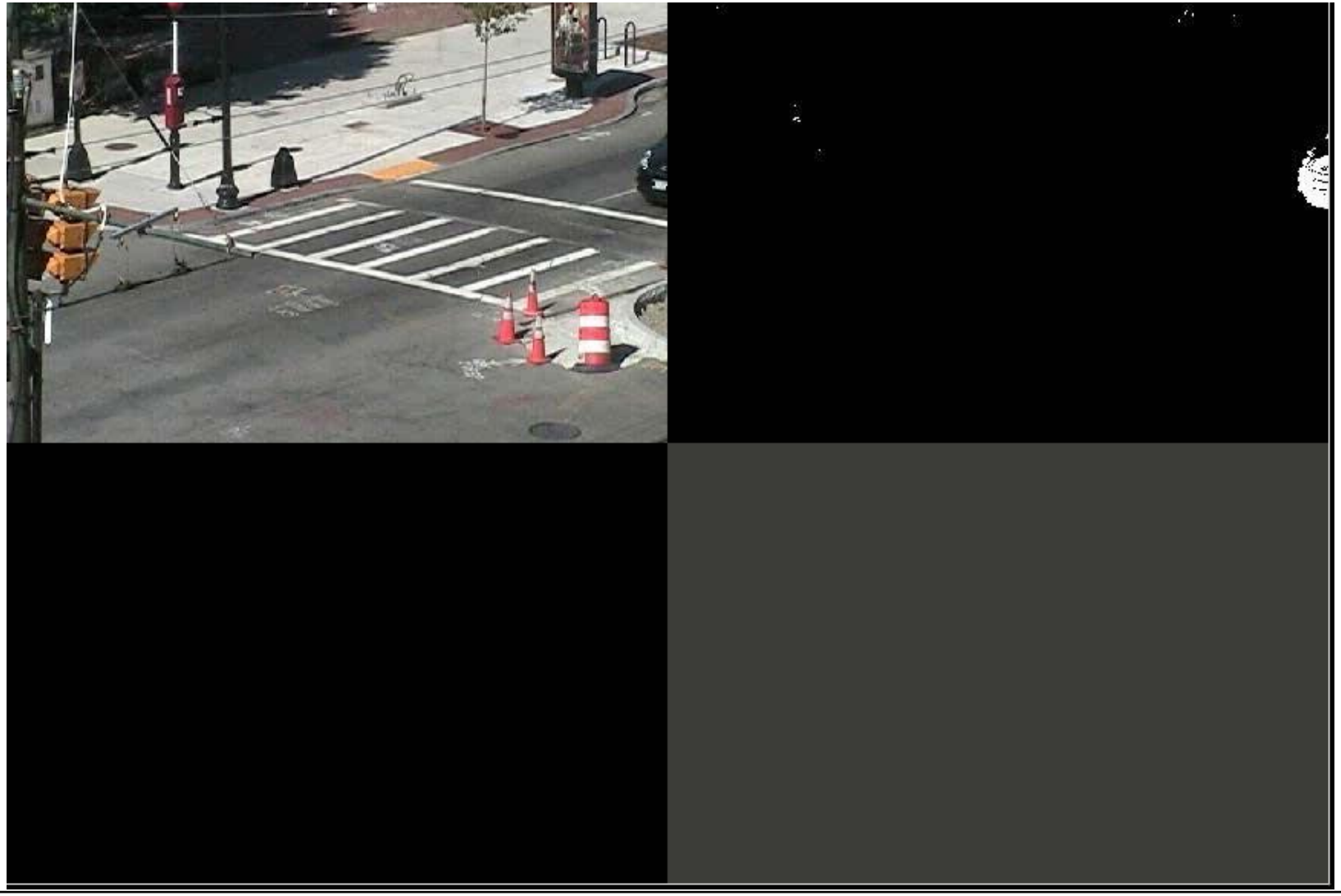




Suspicion due to inactivity

BOSTON UNIVERSITY

Scenario: Abandoned Objects



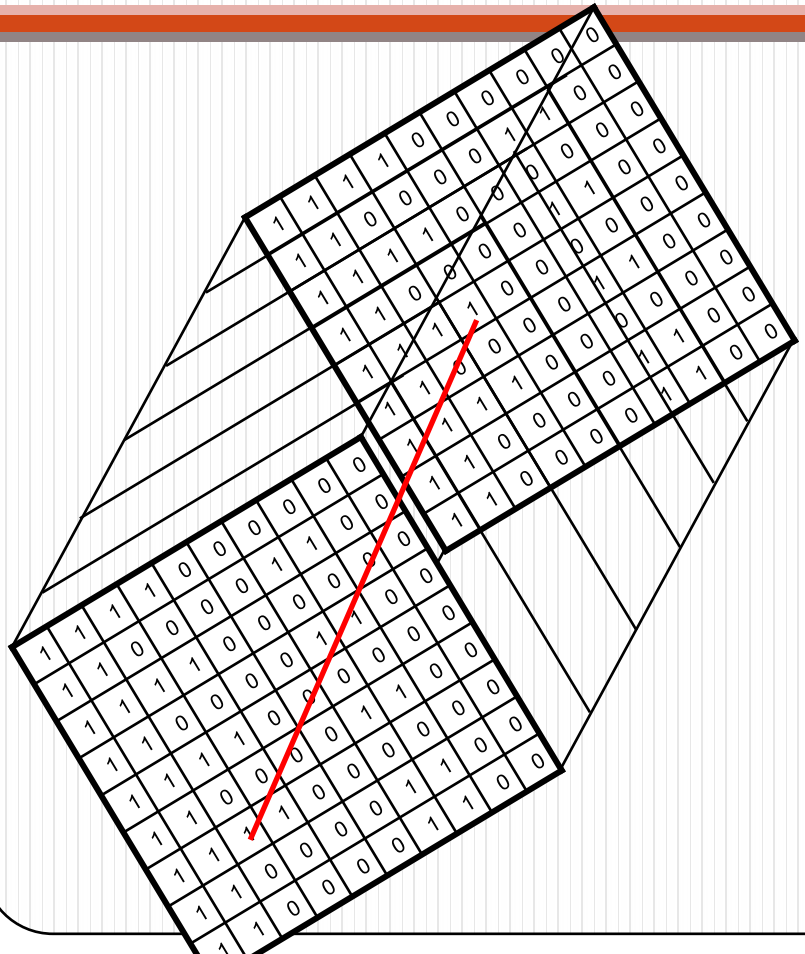
ABNORMAL VIDEO

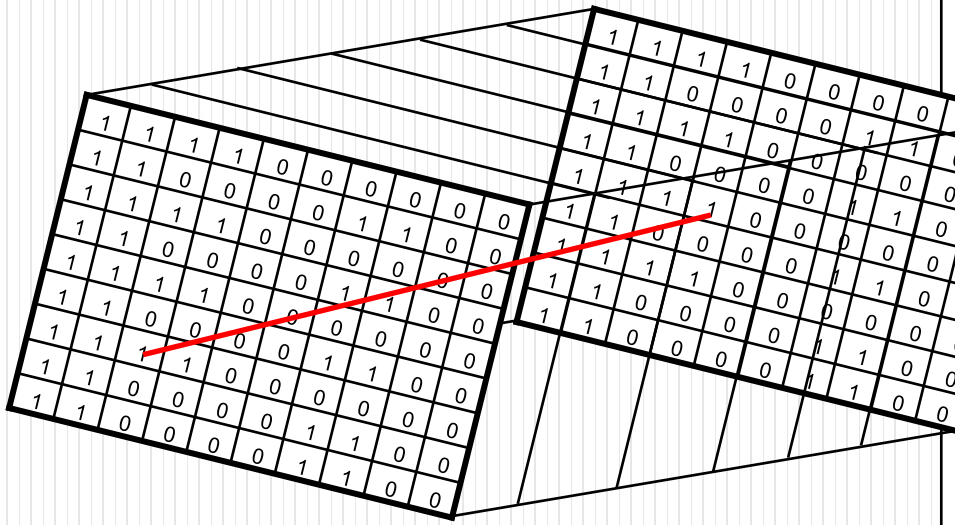
BOSTON UNIVERSITY

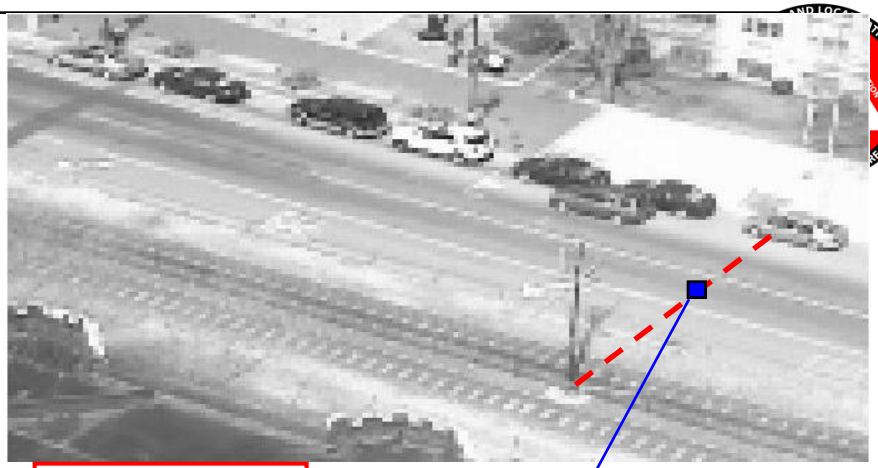
Scenario: Suspicious Activity on overpass

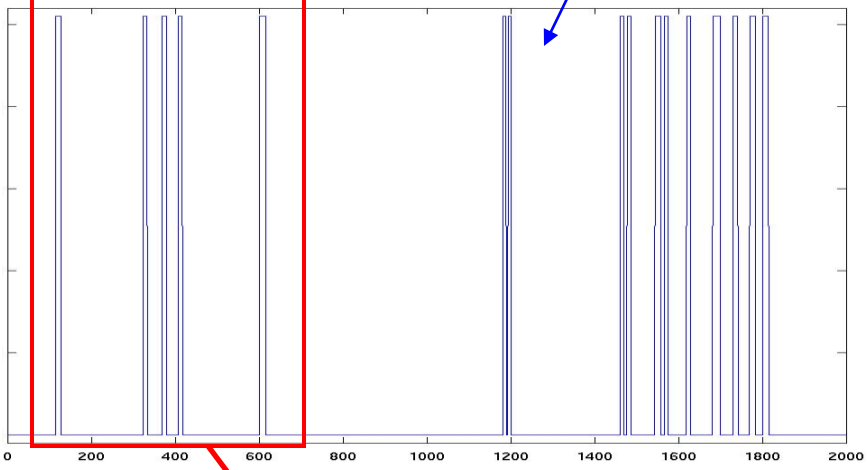
ABNORMAL VIDEO

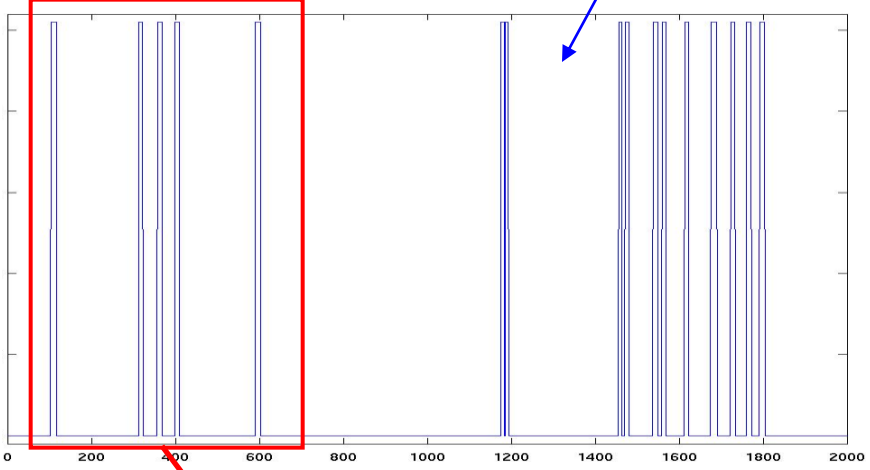
Multi-Camera Systems

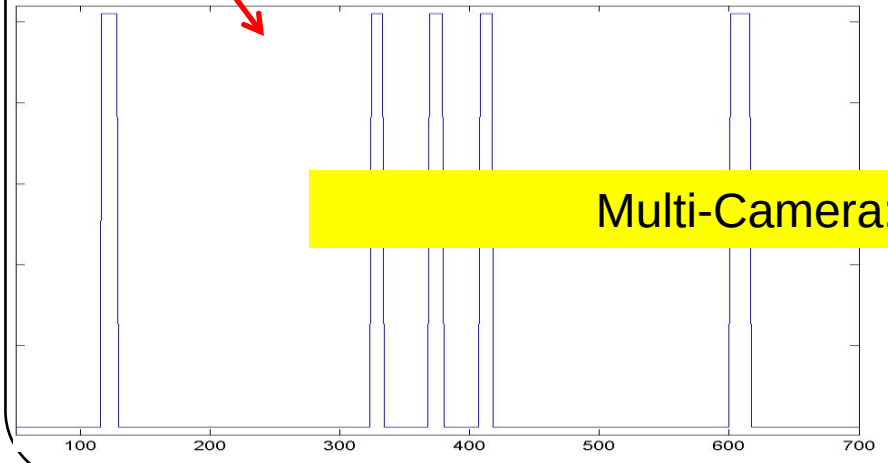


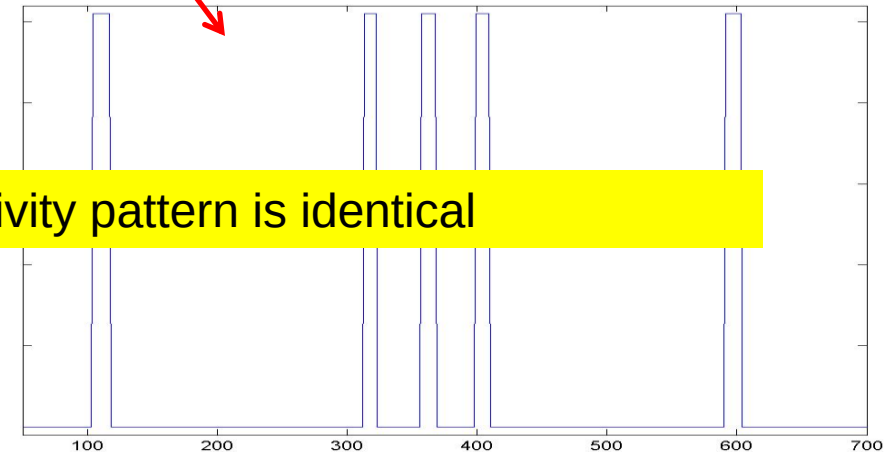




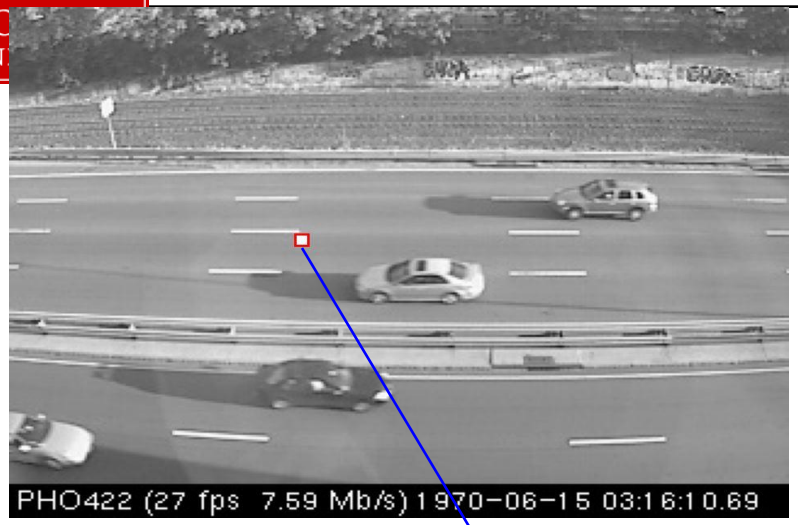


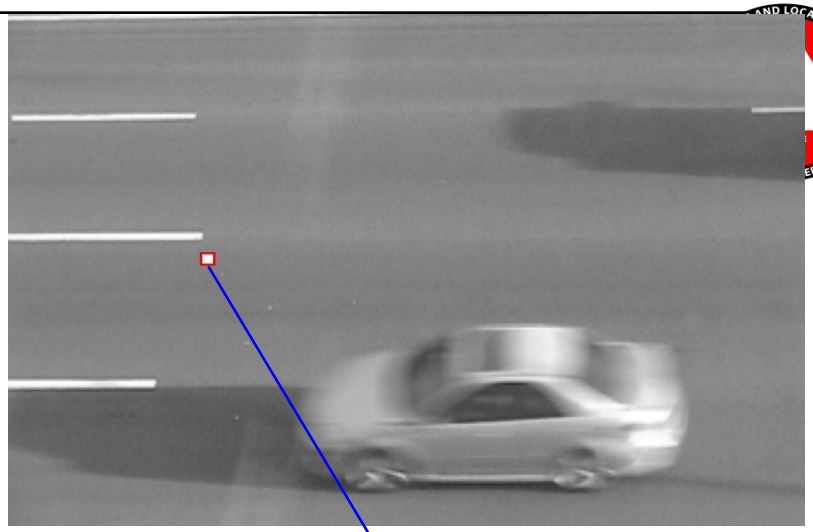


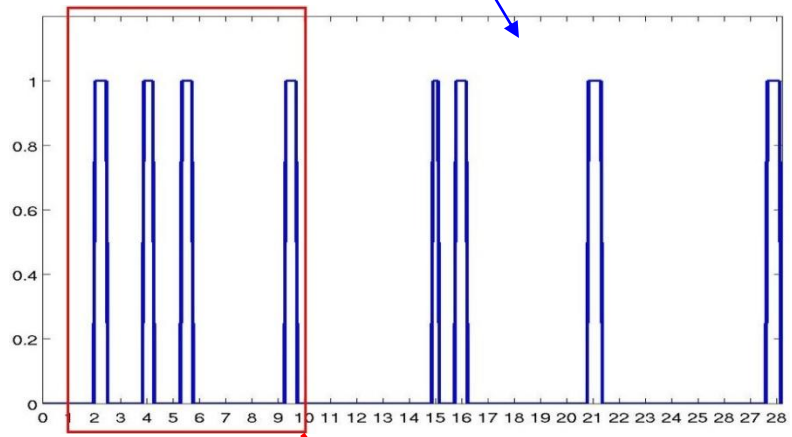


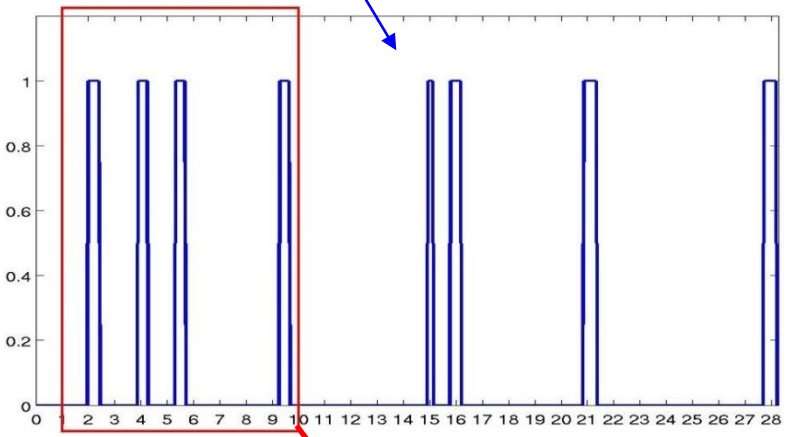


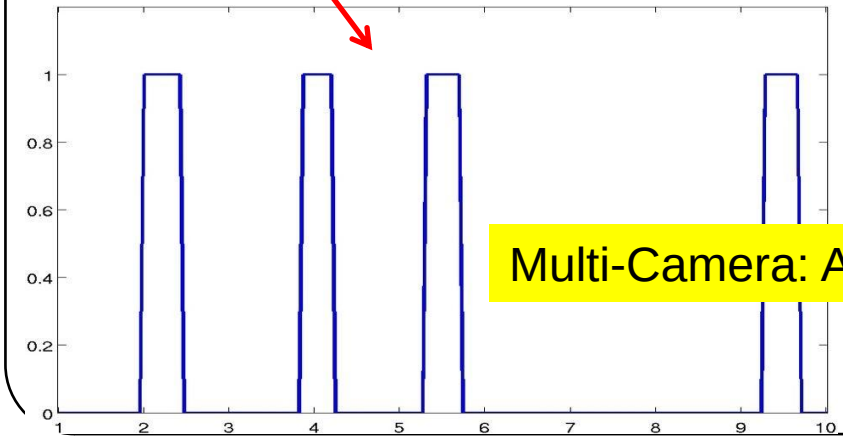
Multi-Camera: Activity pattern is identical

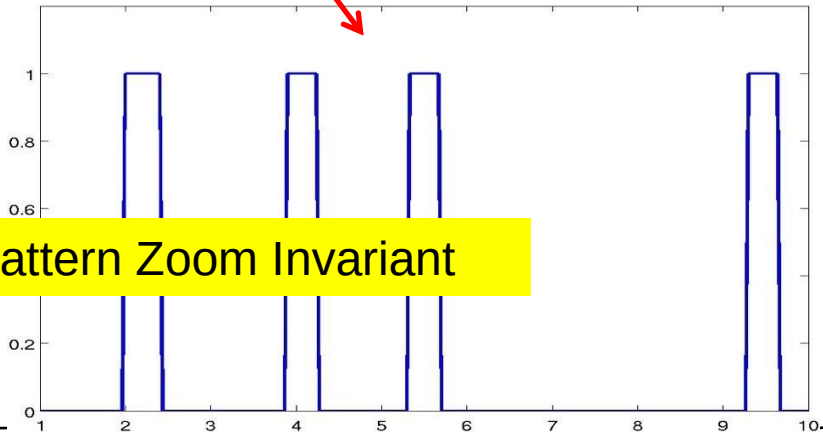






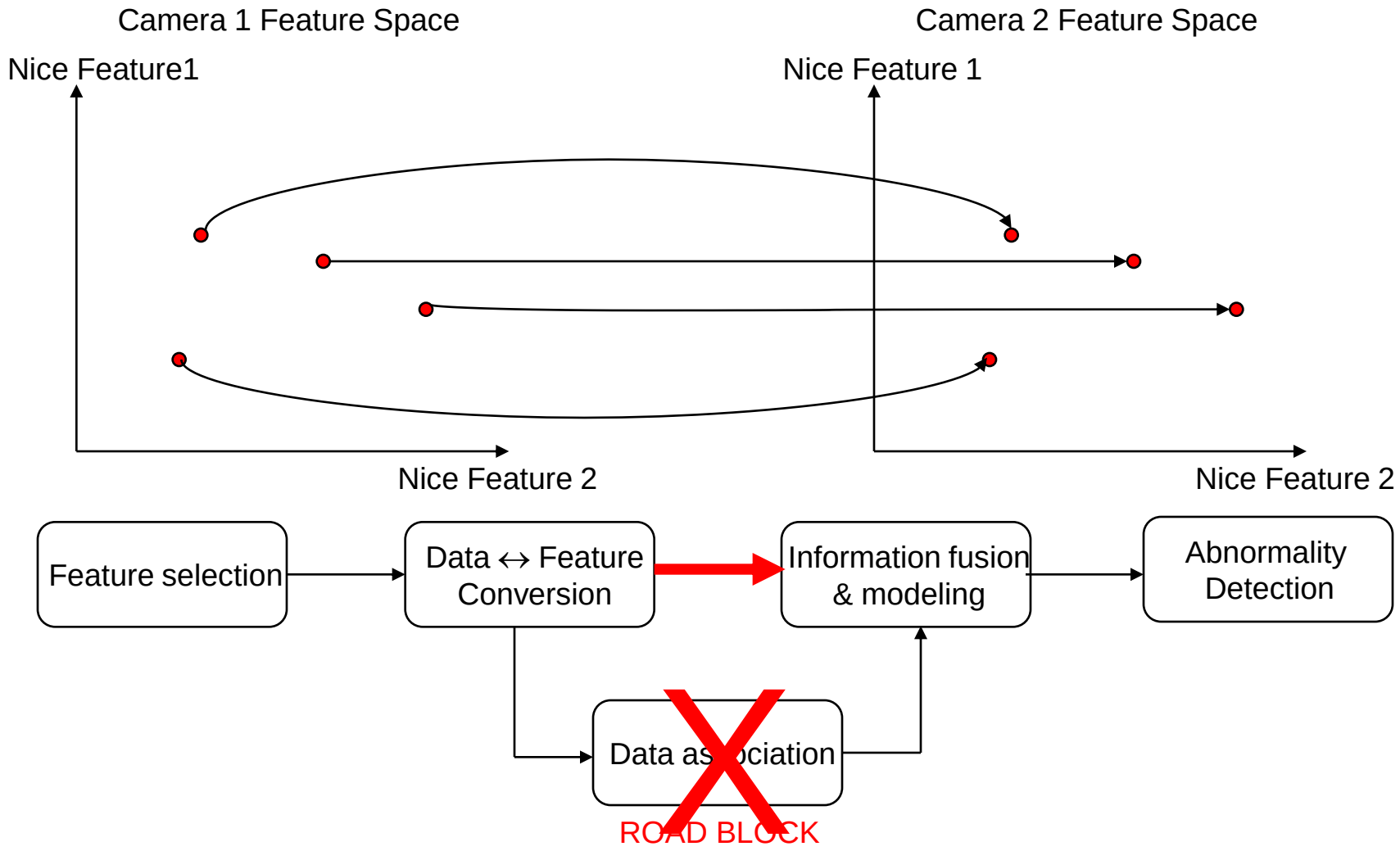


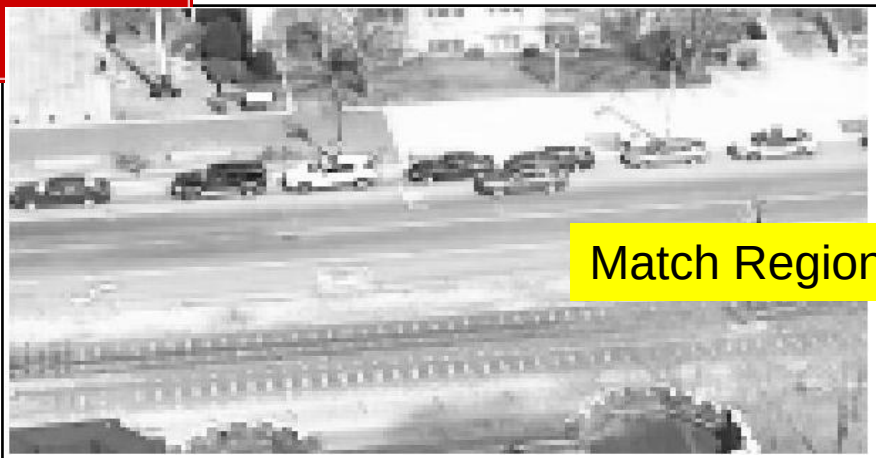




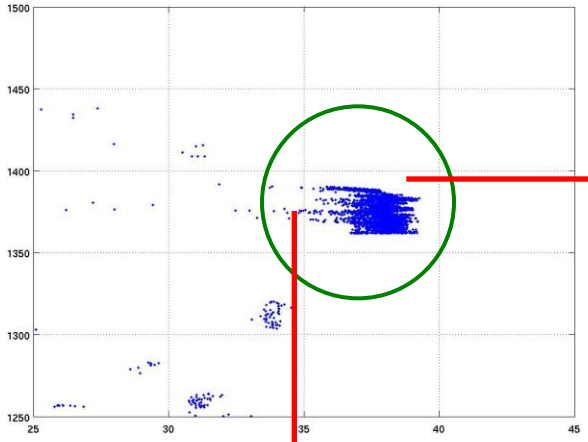
Multi-Camera: Activity pattern Zoom Invariant

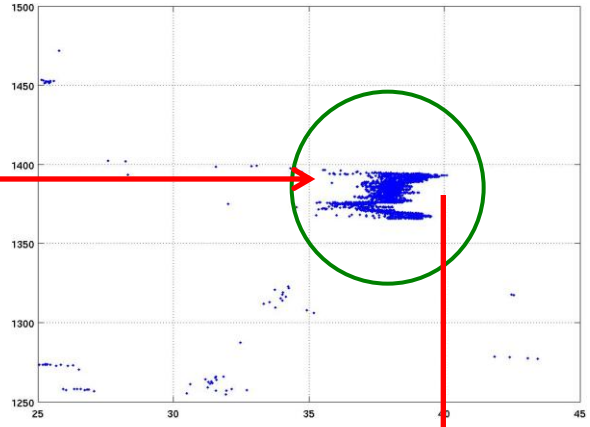
Video-Analytics for Camera Nets

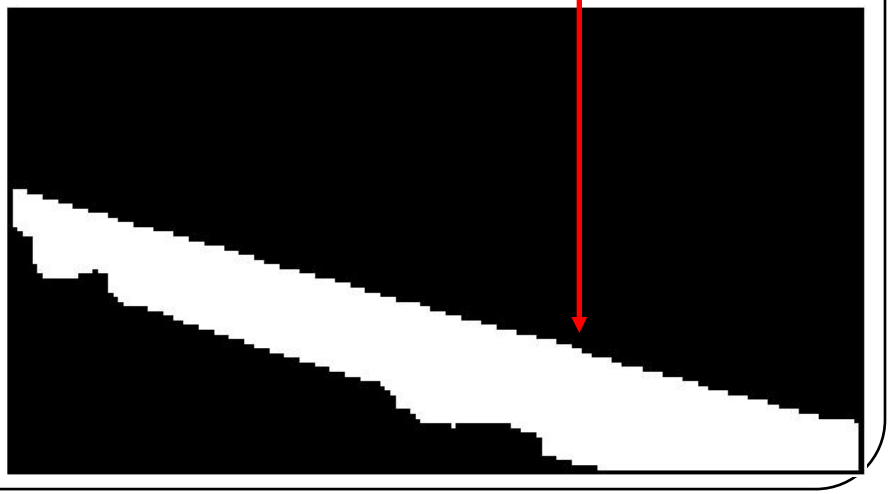




Match Regions with Similar Behavior







Networked Cameras: Diff. Views

Training on Camera 1

Camera 2

Camera 2: Motion

Abnormality: Camera 2

```
 'movie1.avi');
 = 30;
ile(aviFileName,'fps',framesPerSec,'COMPRESSION','Cinepak');

e(inputvideo,3)
queeze(inputvideo(:,:,i));
ame);
```

- Learn Behavior on Camera 1
  - Segment behaviors
- Detect in Camera 2 using Camera 1 behavior

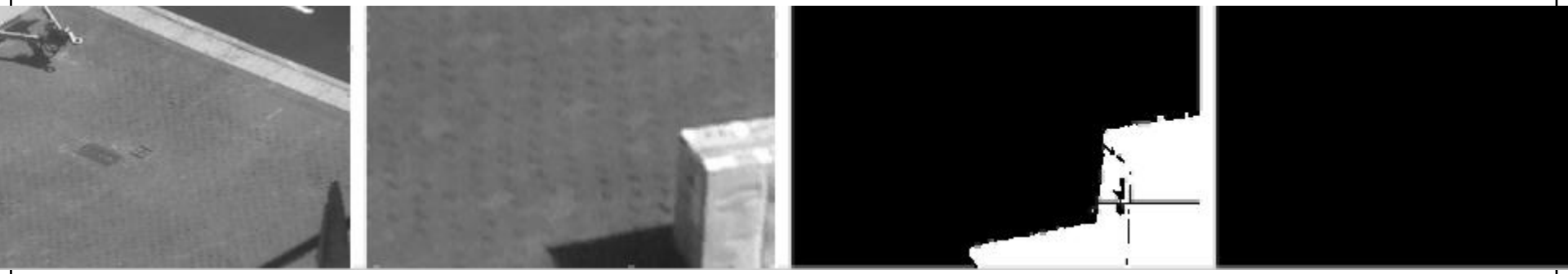
Networked Cameras: Different zooms

Training Camera 1

Camera 2 Zoomed

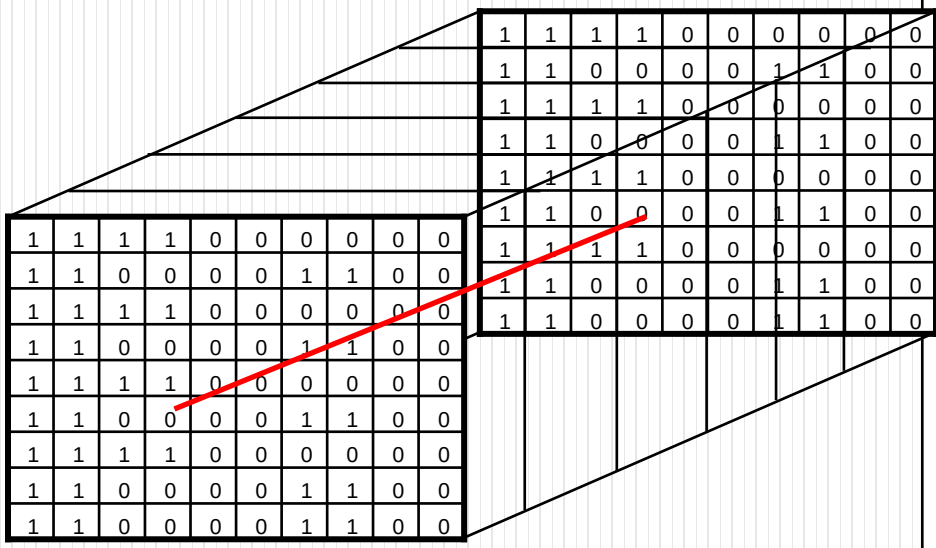
Detected motion

Abnormality

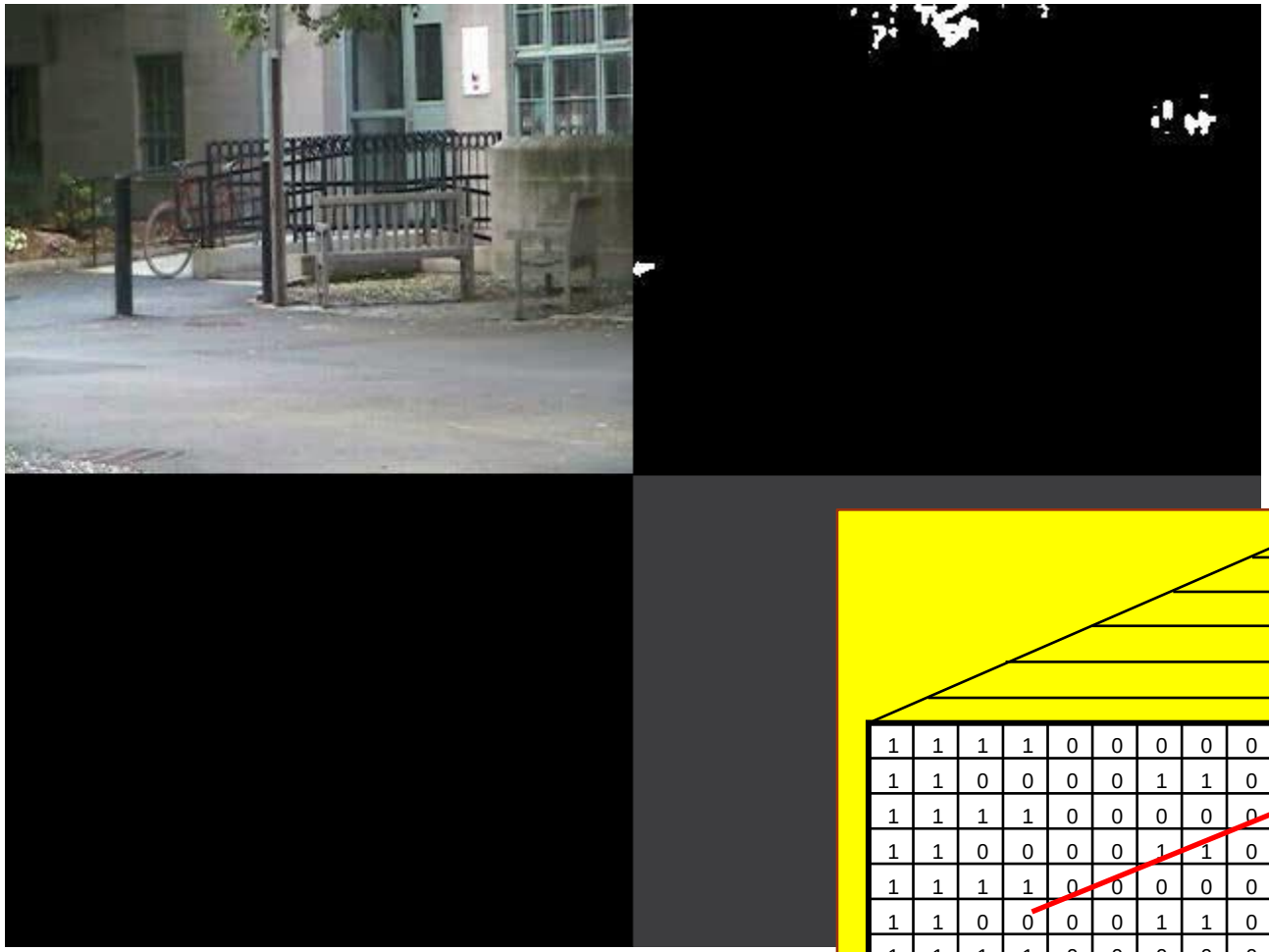


Behavior Models
Multi-Dimensional Features

- B-I Models, Features Size: Crowd Detection
- Co-Occurrence of Events
- Multi-Camera Models

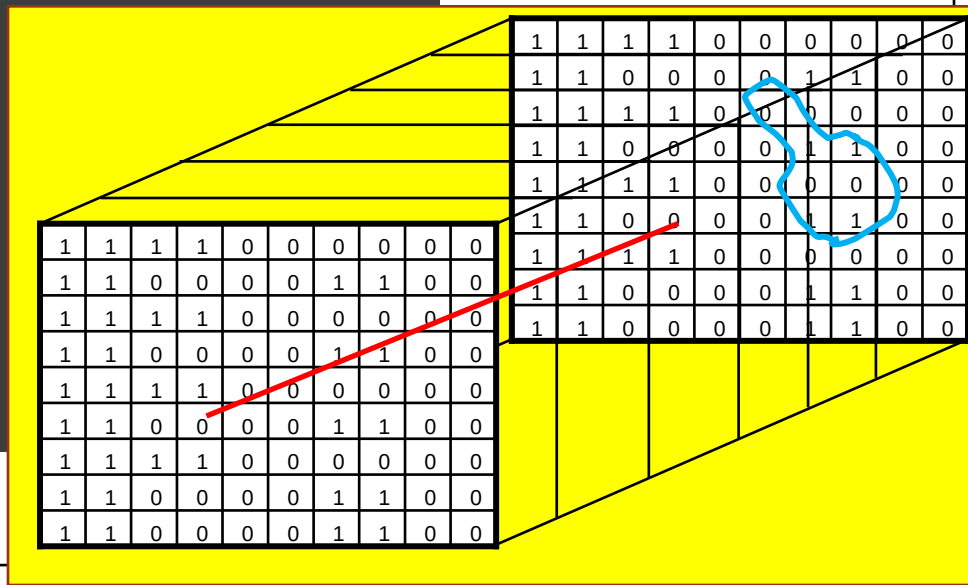


Scenario: Crowd Detection



Behavior@ p,

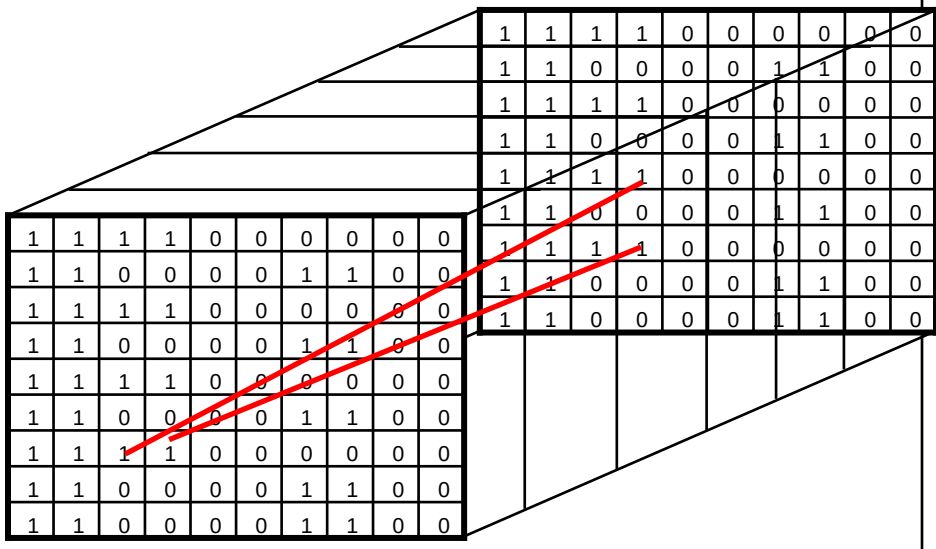
$\Lambda_p(j) = (B_j, I_j, S_j)$

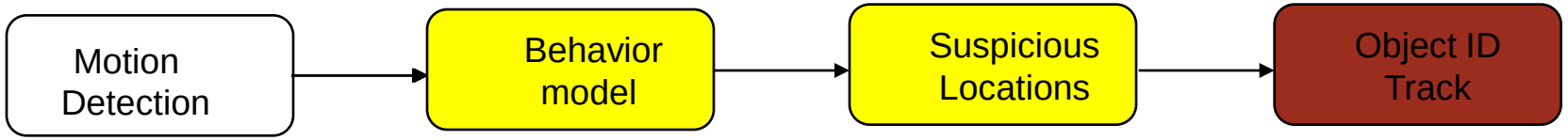


ABNORMAL VIDEO

Co-occurrences
Detection Anomalous Complex Event Sequences

- Events in one region "often" co-occurs with another region
  - → Co-occurrence Maps for pixels
    - Detect illegal U-turns
    - Detect baggage drops,
    - baggage exchange,
    - Abandoned baggage pickup





BOSTON UNIVERSITY

Scenario: Baggage Drop, track suspicious activity

BOSTON UNIVERSITY

Illegal U-Turn

Conclusions

- Take Home Message:
  - Why identify/track every object to find suspicious behavior
  - Pixel-based approach
  - Activity patterns contains significant information

- Algorithms amounts to Counting Bits in a Binary Matrix
  - Real-time tests currently underway at Phillips, TI
  - Several Scenarios → robust, meaningful outcomes

- Extensions to Multi-Camera, Co-occurrences (complex event seq)
  - Geometric, Zoom Invariance --- no need for camera registration!!!
    - Adapts to PTZ, bring/remove assets, heterogenous assets