# Where does Video Analytics go next for TSA

*Octavia Camps and Mario Sznaier*

*Students:*

*C. Dicle, T. Hebble, O. Lehmann, J. Romano, and F. Xiong*

Electrical and Computer Engineering

ALERT

AWARENESS AND LOCALIZATION OF EXPLOSIVES-RELATED THREATS

Northeastern

SIEMENS

# Video Analytics for TSA

- ***In-the-Exit Contraflow Detection***
  - Deployed at CLE Airport
    - Two different exits, different geometry
    - ~0 miss-detections, 1 to 2 false alarms/week
- ***Tag-and-Track across camera network***
  - In progress, to be deployed at CLE Airport soon
    - Two testbeds:
      - Parking to Terminal and
      - Checkpoint Exit to Three Concourses
- **Activity Recognition, Object Left Behind:**
  - Theory developed, tested on simple scenarios

**Dynamics as a key enabler to handle data deluge and obtain actionable information in a timely fashion.**

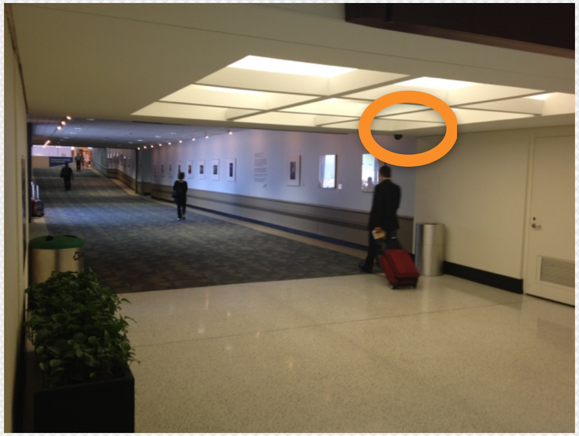
# Goals

Image from http://iware.pk/CCTVSystems.aspx

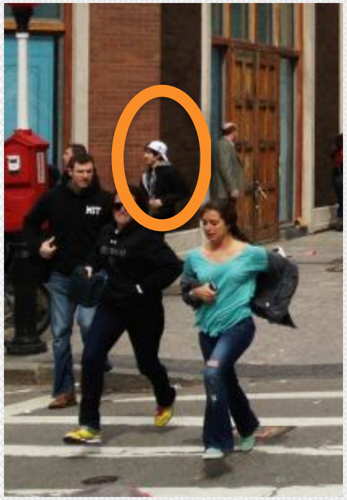
- **Detect Potential Threats & Track Suspects:**
  - Security breaches at portals
  - Track across cameras
  - Disruptive, suspicious behaviors
  - Objects left behind
- **Detect Other Emergencies:**
  - Person falling or hurt
  - Lost child
  - Stolen property
- **Customers:**
  - TSA
  - Law enforcement agencies
  - Sport venues, theme parks, etc.

# Requirements and Challenges

- Infrastructure:
  - Cameras already deployed
    - Security surveillance cameras
    - Citizen’s cell phones
  - Video is recorded, not processed
- Performance:
  - 0 misdetections, ~0 false alarms
  - Timely detection: real time processing
- Data Deluge Challenge:
  - Need to process vast amounts of highly complex data
  - Cope with environment, viewpoint, appearance changes
  - Ignore nuisance/clutter data
  - Find actionable/relevant information
  - ➡ Dynamics as key enabler

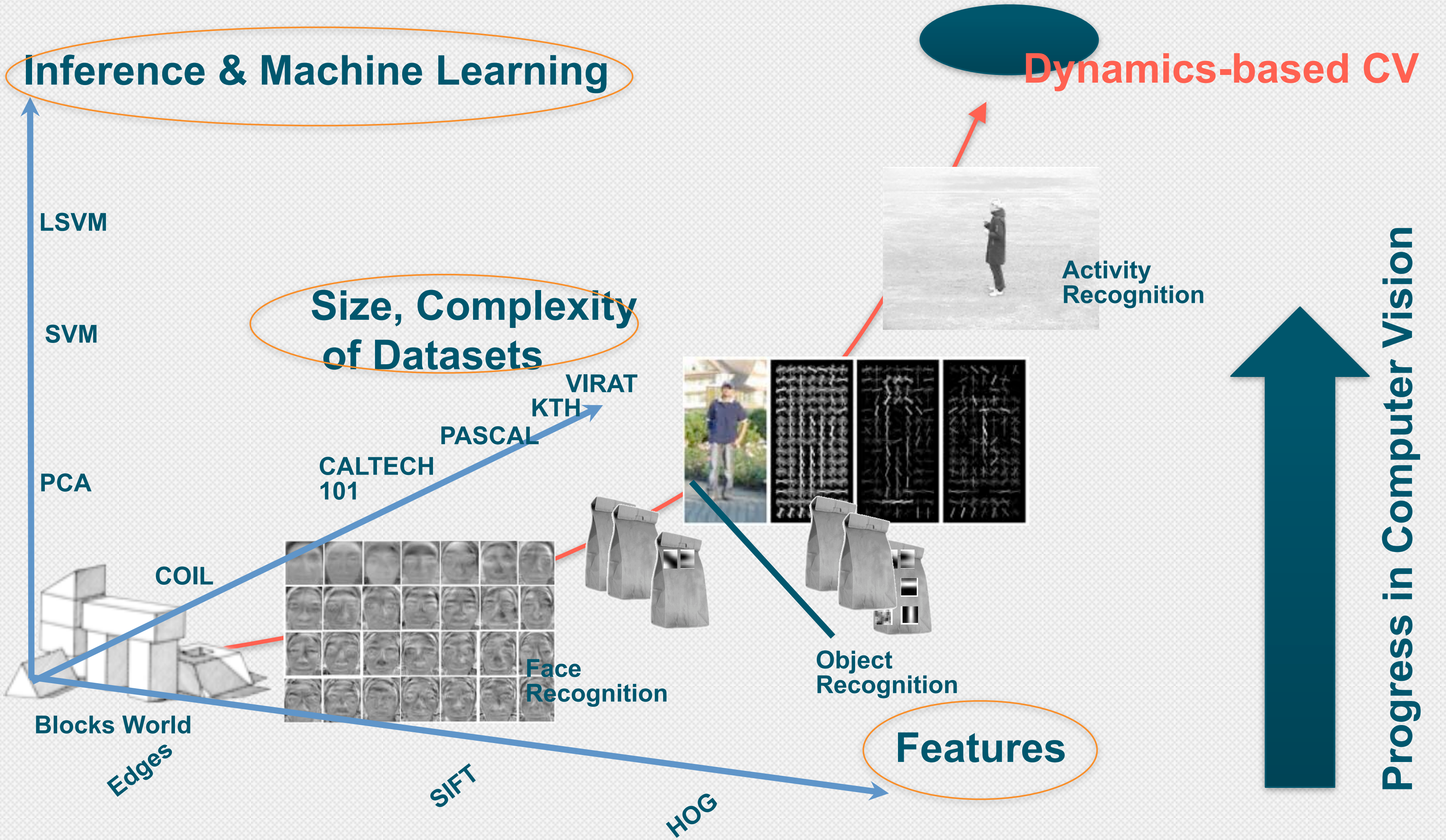


**Surveillance camera in an airport.**



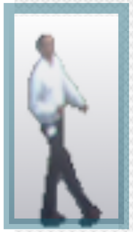
**Cell video showed the Boston Marathon Suspect(s)**

# Addressing the Challenges



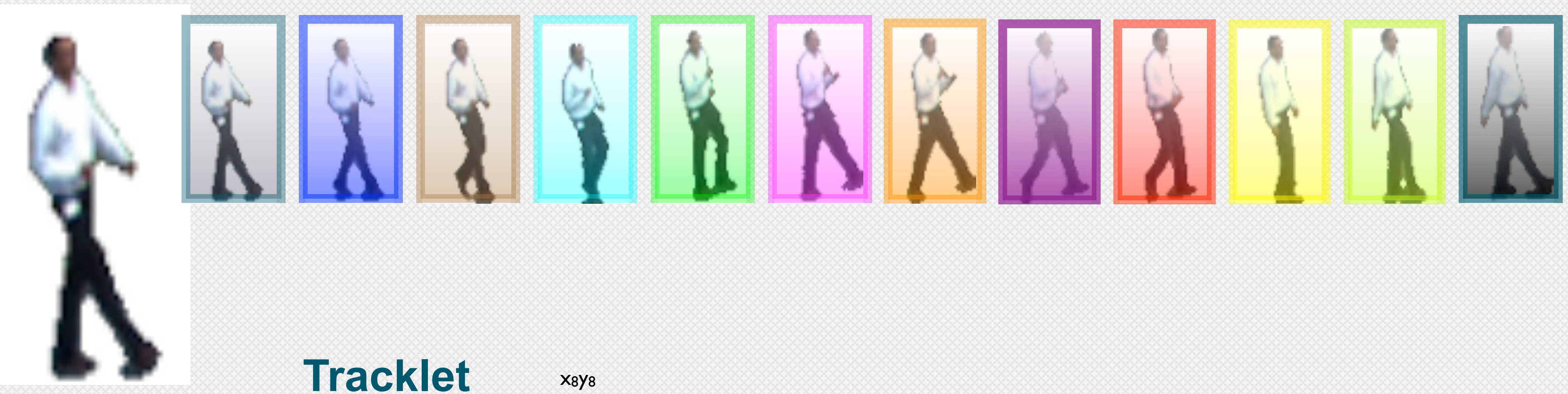
# Dynamics-based Feature

- Sequences as features capture the underlying dynamics



# Dynamics-based Feature

- Sequences as features capture the underlying dynamics



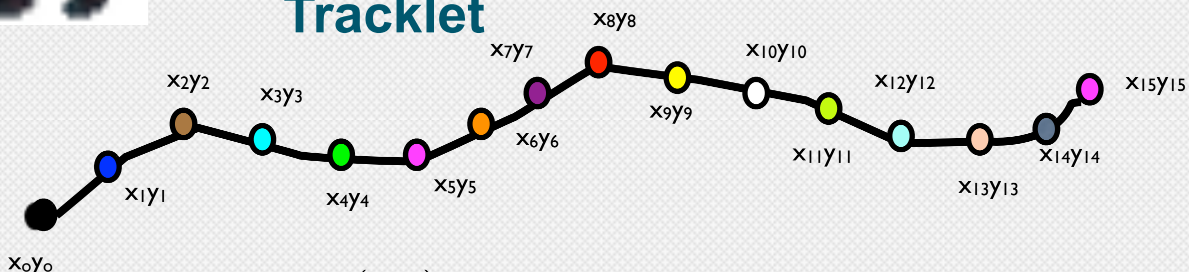

“Memory”

$$\hat{\mathbf{y}}_k = \sum_{i=1}^{n(\sigma_k)} a_i \hat{\mathbf{y}}_{k-i}$$

# Dynamics-based Feature

- Sequences as features capture the underlying dynamics

**Tracklet**

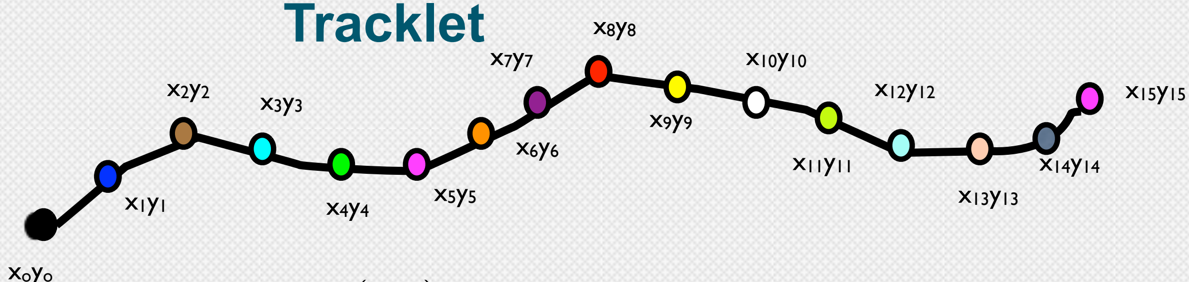



$$\hat{\mathbf{y}}_k = \sum_{i=1}^{n(\sigma_k)} a_i \hat{\mathbf{y}}_{k-i}$$

# Dynamics-based Feature

- Sequences as features capture the underlying dynamics

**Tracklet**




$$\hat{\mathbf{y}}_k = \sum_{i=1}^{n(\sigma_k)} a_i \hat{\mathbf{y}}_{k-i}$$

| | | | |
|---|---|---|---|
| $x_0$ $y_0$ | $x_1$ $y_1$ | $x_2$ $y_2$ | $x_3$ $y_3$ |
| $x_1$ $y_1$ | $x_2$ $y_2$ | $x_3$ $y_3$ | $x_4$ $y_4$ |
| $x_2$ $y_2$ | $x_3$ $y_3$ | $x_4$ $y_4$ | $x_5$ $y_5$ |
| $x_3$ $y_3$ | $x_4$ $y_4$ | $x_5$ $y_5$ | $x_6$ $y_6$ |
| $x_4$ $y_4$ | $x_5$ $y_5$ | $x_6$ $y_6$ | $x_7$ $y_7$ |
| $x_5$ $y_5$ | $x_6$ $y_6$ | $x_7$ $y_7$ | $x_8$ $y_8$ |
| $x_6$ $y_6$ | $x_7$ $y_7$ | $x_8$ $y_8$ | $x_9$ $y_9$ |
| $x_7$ $y_7$ | $x_8$ $y_8$ | $x_9$ $y_9$ | $x_{10}$ $y_{10}$ |
| $x_8$ $y_8$ | $x_9$ $y_9$ | $x_{10}$ $y_{10}$ | $x_{11}$ $y_{11}$ |
| $x_9$ $y_9$ | $x_{10}$ $y_{10}$ | $x_{11}$ $y_{11}$ | $x_{12}$ $y_{12}$ |
| $x_{10}$ $y_{10}$ | $x_{11}$ $y_{11}$ | $x_{12}$ $y_{12}$ | $x_{13}$ $y_{13}$ |

**Hankelet**

# Dynamics-based Feature

- Sequences as features capture the underlying dynamics


- **Hankelet Invariants:**
  - Rank n = complexity
  - AR is invariant to affine transformations
  - Columns Subspace is invariant to initial condition

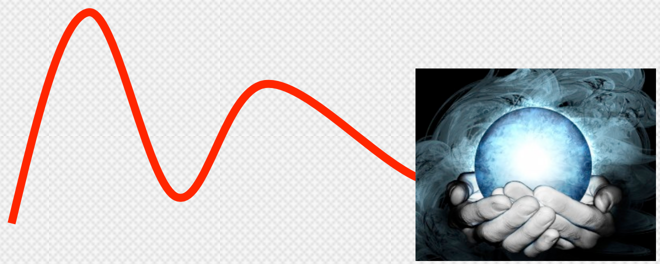
A compact, yet rich representation.

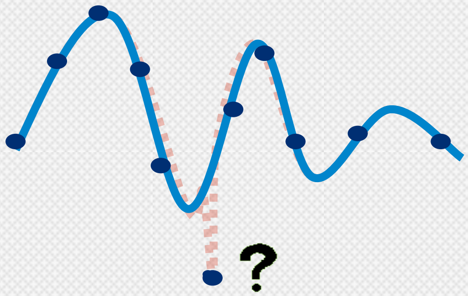
$$\begin{bmatrix} x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ x_2 & x_3 & x_4 & x_5 \\ y_2 & y_3 & y_4 & y_5 \\ x_3 & x_4 & x_5 & x_6 \\ y_3 & y_4 & y_5 & y_6 \\ x_4 & x_5 & x_6 & x_7 \\ y_4 & y_5 & y_6 & y_7 \\ x_5 & x_6 & x_7 & x_8 \\ y_5 & y_6 & y_7 & y_8 \\ x_6 & x_7 & x_8 & x_9 \\ y_6 & y_7 & y_8 & y_9 \\ x_7 & x_8 & x_9 & x_{10} \\ y_7 & y_8 & y_9 & y_{10} \\ x_8 & x_9 & x_{10} & x_{11} \\ y_8 & y_9 & y_{10} & y_{11} \\ x_9 & x_{10} & x_{11} & x_{12} \\ y_9 & y_{10} & y_{11} & y_{12} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ y_{10} & y_{11} & y_{12} & y_{13} \end{bmatrix}$$

**Hankelet**

# Dynamics-based Inference Methods

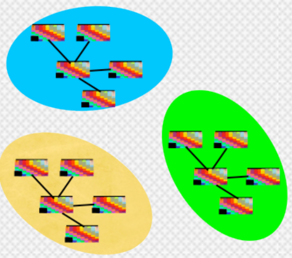
- Predict

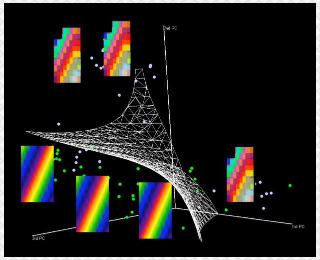





- Denoise
- Associate data


- Classify data
  - Unsupervised
  - Supervised

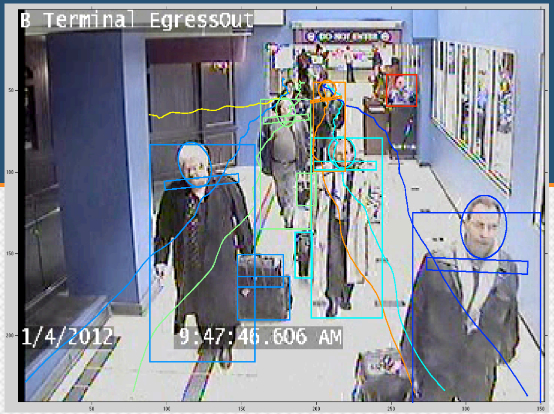




- Detect Causality

**Rank minimization of a structured Hankelet matrices**

# Real Surveillance Data




## Partnership with TSA and CLE Airport

- Access to live video:
  - 3 and 2 cameras at two exit lanes
  - 5 cameras from Parking structure to terminal
  - More to come

CLE

## Collaboration with Siemens Research

SIEMENS

- Support recording and accessing video
- Support transition technology to commercial surveillance systems

## Semi-automated annotation

- Ease ground-truth annotation of data
  - Location, attributes, Id-across cameras
- Facilitate both training and testing of algorithms

# Anchovy


11/29/2012 11:14:41:288
#23
50
100
150
200
250
300
350
400
450
0
100
200
300
400
500
600
700
Follow Single Object
Delete Future Annotations
Delete Past Annotations
Mark Object as Finished
Edit Bounding Box
Optimize Path
Show All Objects
Show Bounding Box
Show Tracks
Tracks Algebra
Save Tracks
Goto Frame...

# In-The-Exit Contraflow Detection

***In-the-exit contraflow:***

- Entrance through an airport exit:
  - Security-breach
  - Terminal(s) must be evacuated
  - Flights cancelled, millions of dollars cost

***In-the-exit contraflow:***

- Entrance through an airport exit:
  - Security-breach
  - Terminal(s) must be evacuated
  - Flights cancelled, millions of dollars cost

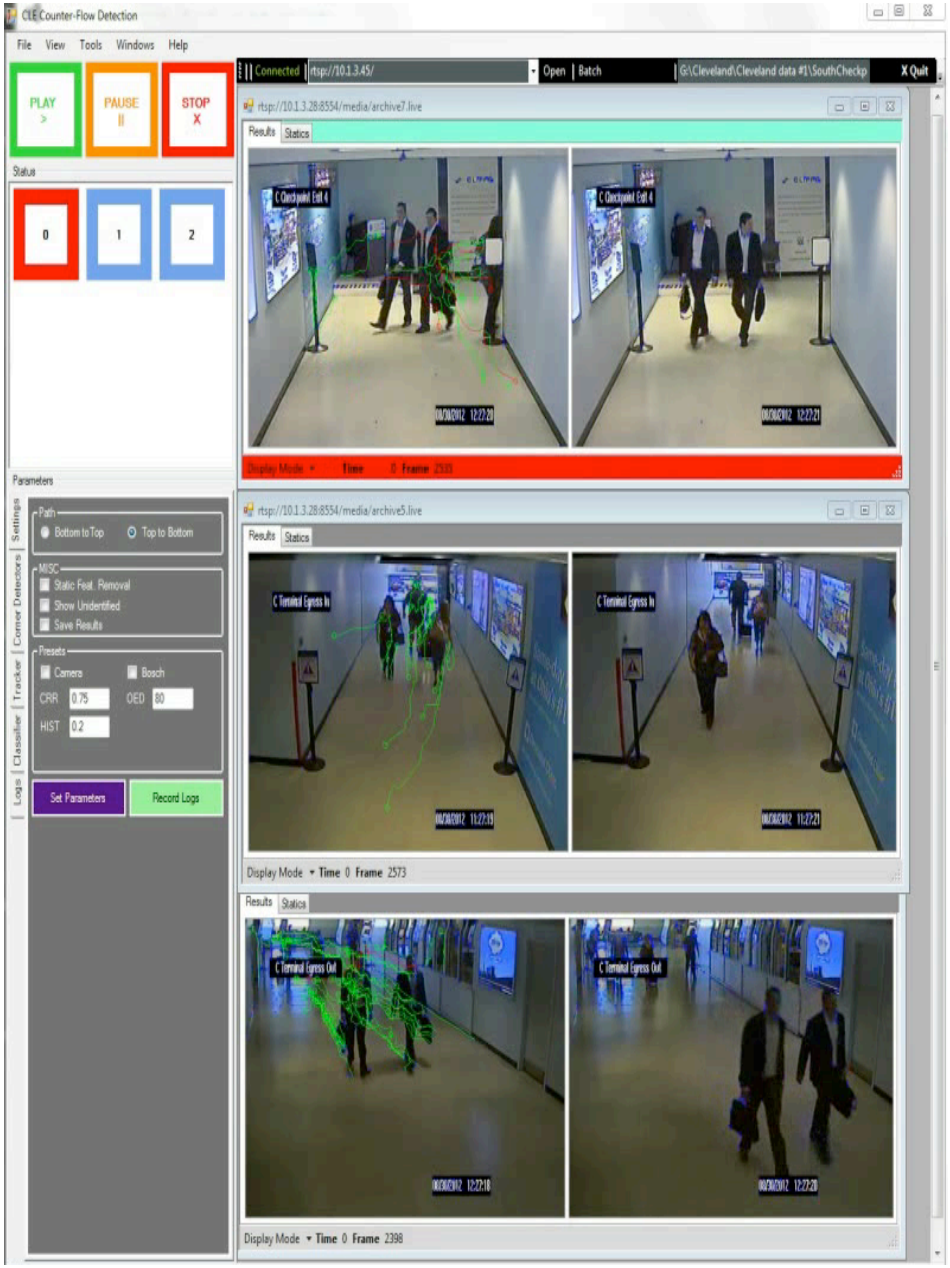
# In-The-Exit Contraflow Detection

- Bank of trackers + detection of contraflow motion
  - Real time GPU-based implementation
  - Currently deployed at CLE Airport, running 24/7, two exit lanes
  - Statistics on 10 weeks of video:
    - **628 detections; 0 Miss-detections, 12 False alarms**

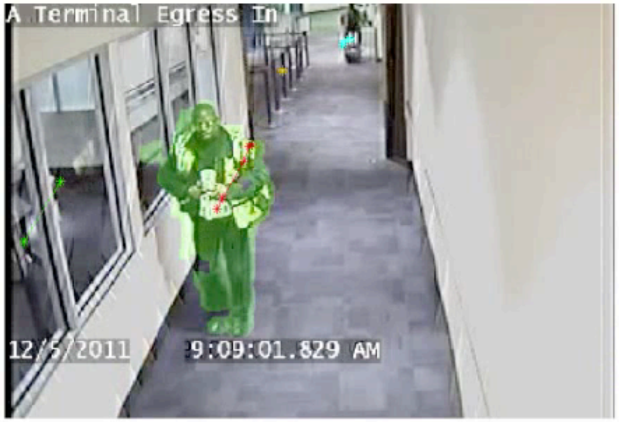

# In-The-Exit Contraflow Detection

- Bank of trackers + detection of contraflow motion
  - Real time GPU-based implementation
  - Currently deployed at CLE Airport, running 24/7, two exit lanes
  - Statistics on 10 weeks of video:
    - **628 detections; 0 Miss-detections, 12 False alarms**

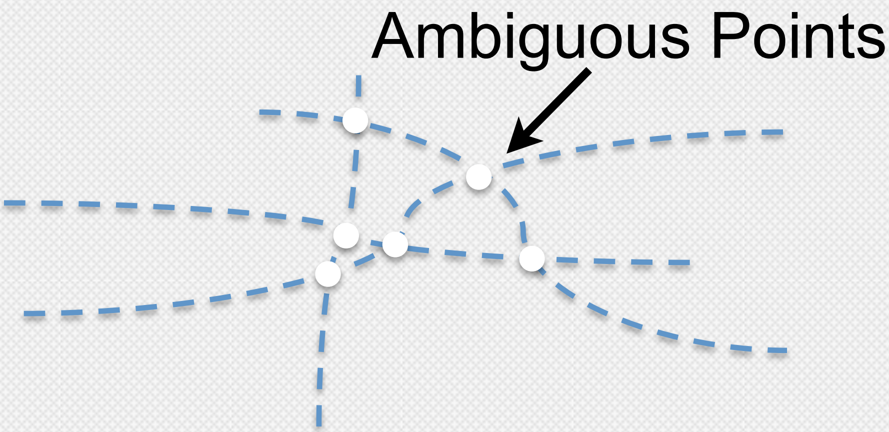

# RPI In-the-Exit (R. Radke)



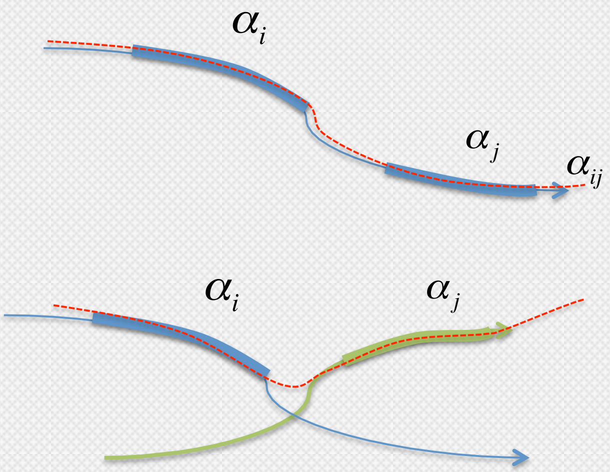

# BU In-the-Exit (V. Saligrama)



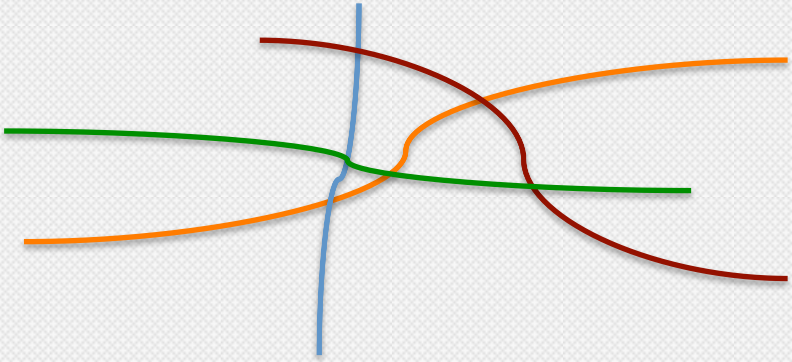

# Tracking in Crowded Scenes



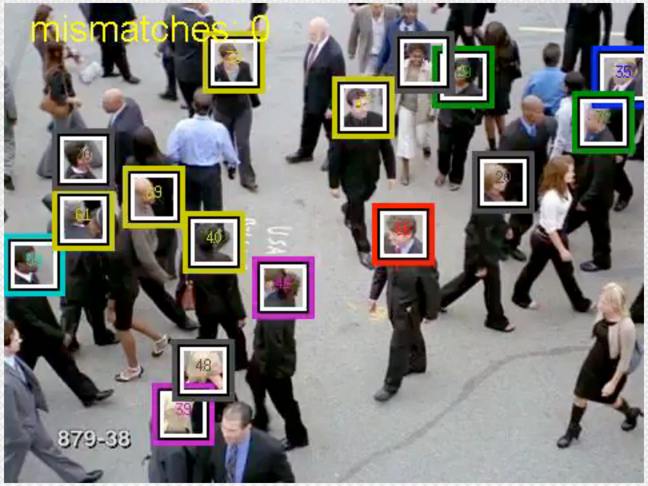

**Matching tracklets have simpler dynamics**




**Dynamics carry Id information of the target.**

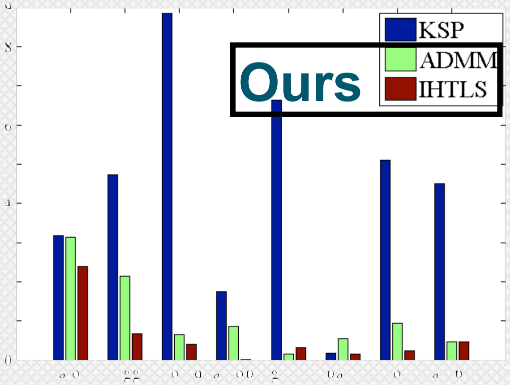


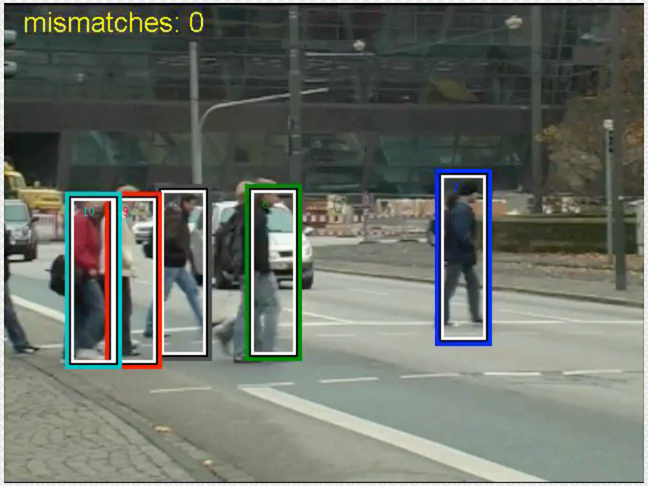
# Tracking in Crowded Scenes

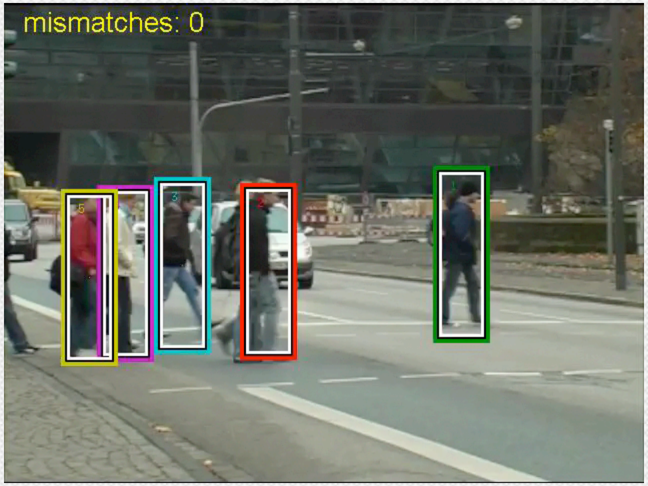




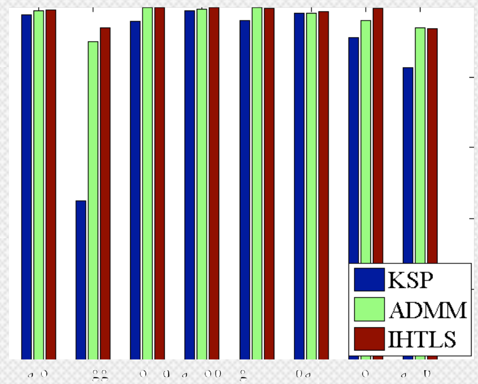
Switches




MOTA



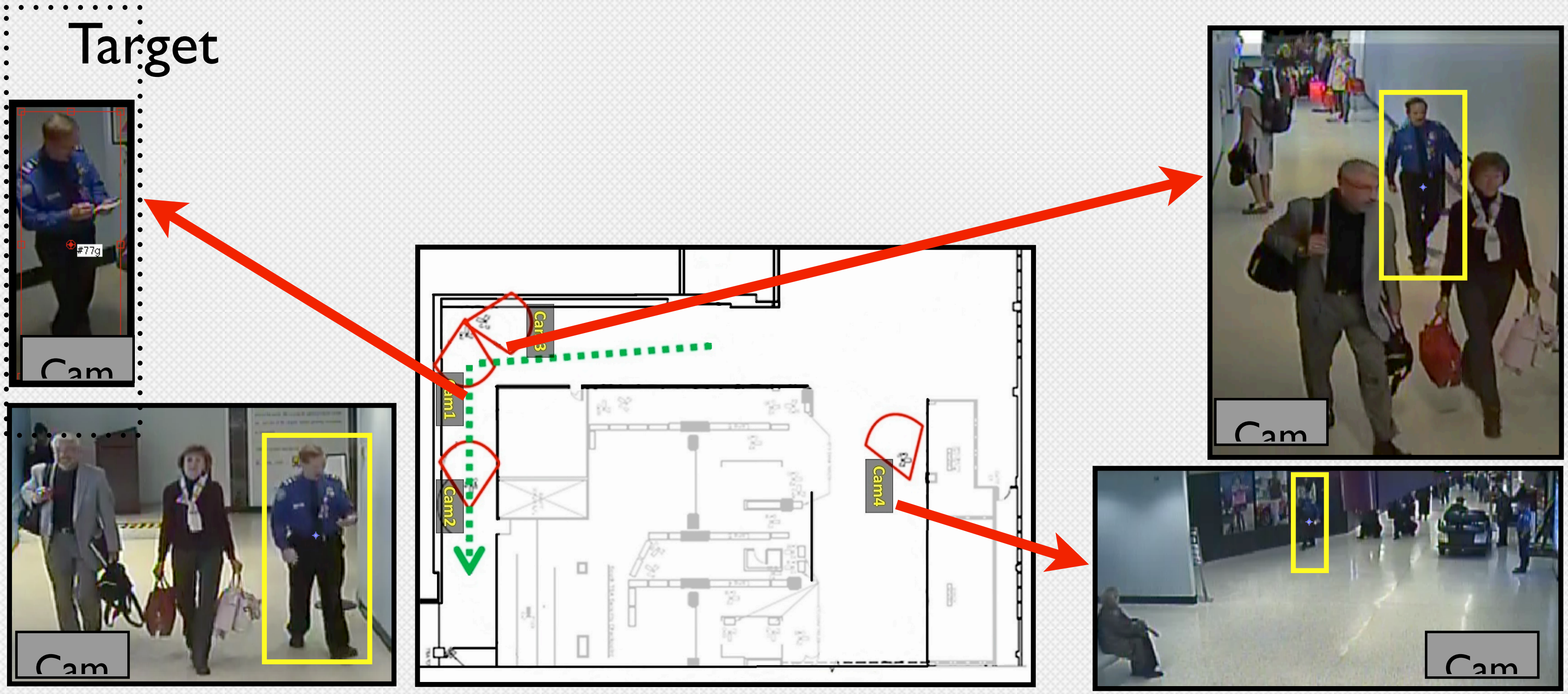

$$\text{MOTA} = 1 - \frac{\text{outliers} + \text{misses} + \text{switches}}{\text{all object occurrences}}$$

J. Berclaz, F. Fleuret, E. Turetken, and P. Fua.

Ours (ICCV 13)

# Tag-and-Track

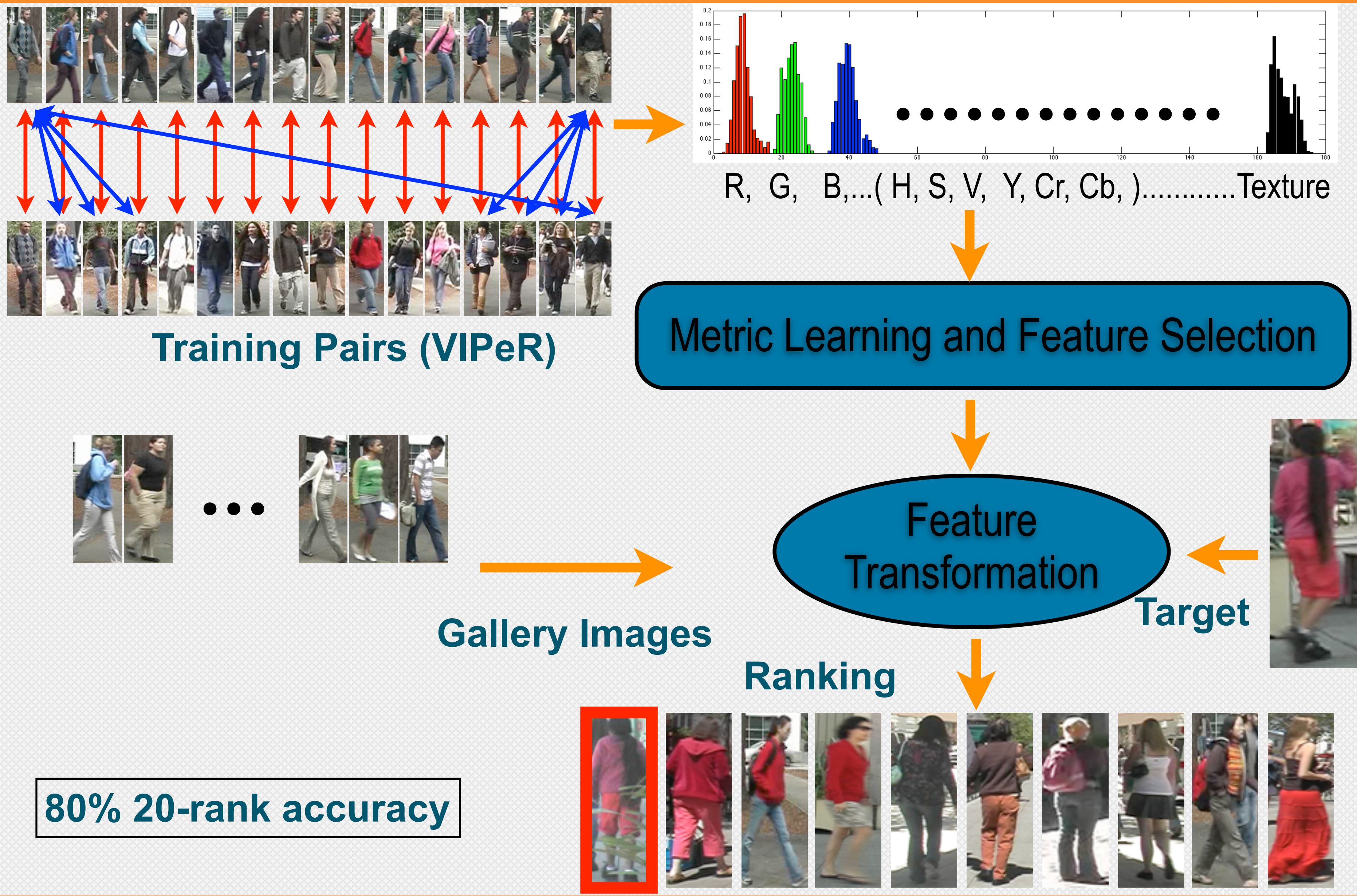
## Unsupervised, real-time tracking through a camera network




**Two testbeds at CLE Airport:**

**5 and 7 cameras, little or no overlapping field of view**

# Appearance-based ReId



# Using dynamic appearance

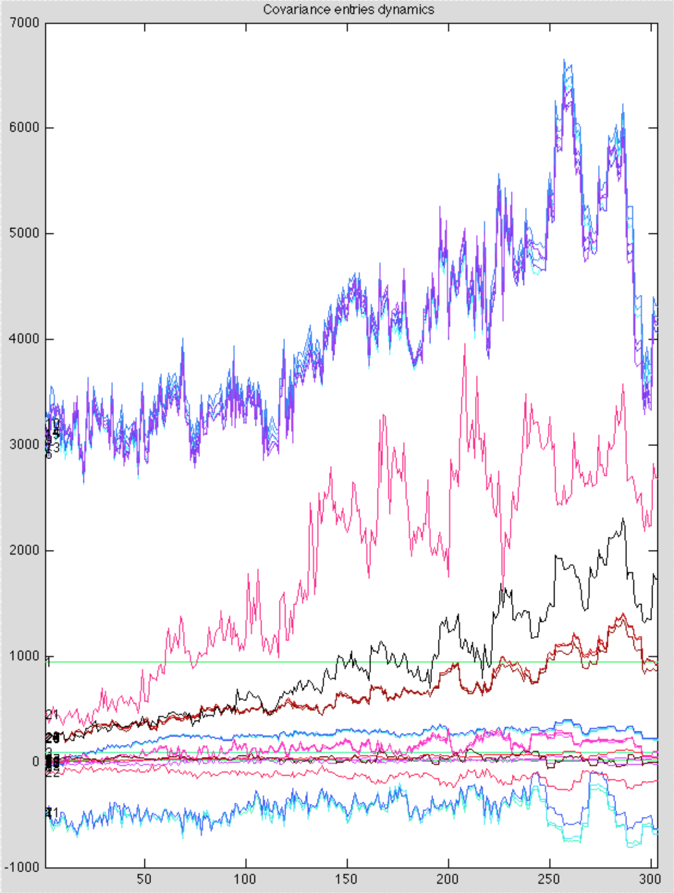
# Using dynamic appearance

| | x | \|y-y_{0}\| | R | G | B | I_{x} | \|I_{y}\| |
|---|---|---|---|---|---|---|---|
| x | 1 | 2 | 4 | 7 | 11 | 16 | 22 |
| \|y-y_{0}\| | | 3 | 5 | 8 | 12 | 17 | 23 |
| R | | | 6 | 9 | 13 | 18 | 24 |
| G | | | | 10 | 14 | 19 | 25 |
| B | | | | | 15 | 20 | 26 |
| I_{x} | | | | | | 21 | 27 |
| \|I_{y}\| | | | | | | | 28 |

- Use **Region Covariance** to model target appearance.

# Using dynamic appearance

| | x | \|y-y_{0}\| | R | G | B | I_{x} | \|I_{y}\| |
|---|---|---|---|---|---|---|---|
| x | 1 | 2 | 4 | 7 | 11 | 16 | 22 |
| \|y-y_{0}\| | | 3 | 5 | 8 | 12 | 17 | 23 |
| R | | | 6 | 9 | 13 | 18 | 24 |
| G | | | | 10 | 14 | 19 | 25 |
| B | | | | | 15 | 20 | 26 |
| I_{x} | | | | | | 21 | 27 |
| \|I_{y}\| | | | | | | | 28 |



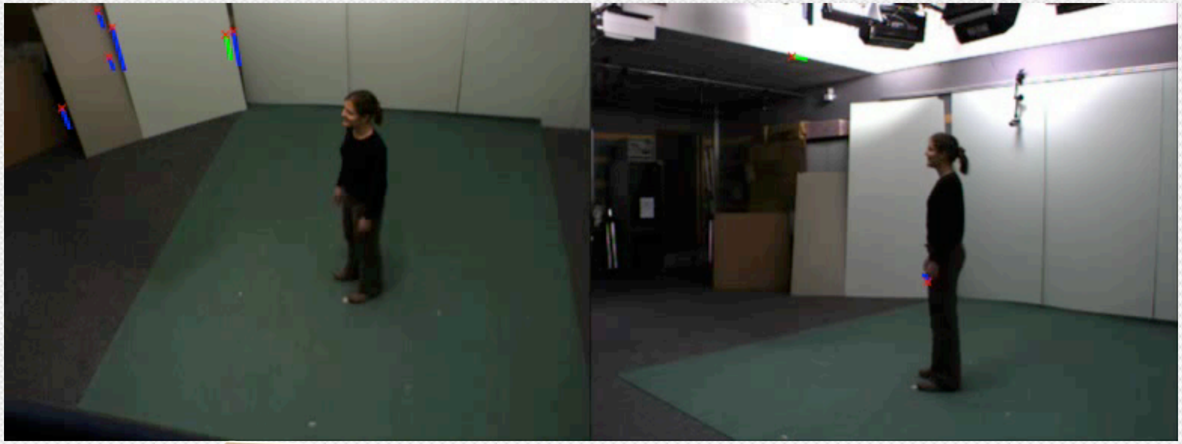

- Use **Region Covariance** to model target appearance.
- Model its dynamic evolution on the Lie group of positive definite matrices.

# Using dynamic appearance


- Use **Region Covariance** to model target appearance.
- Model its dynamic evolution on the Lie group of positive definite matrices.
- Compare appearances using their intrinsic distance on this manifold.

# Cross-View Activity Recognition

- IXMAS dataset (5 cameras, 12 actors, 11 activities):
  - Check watch,
  - cross arms,
  - scratch head,
  - sit down,
  - get up,
  - turn around,
  - walk,
  - wave,
  - punch,
  - kick,
  - pick up




=

**Affine and initial condition invariance**

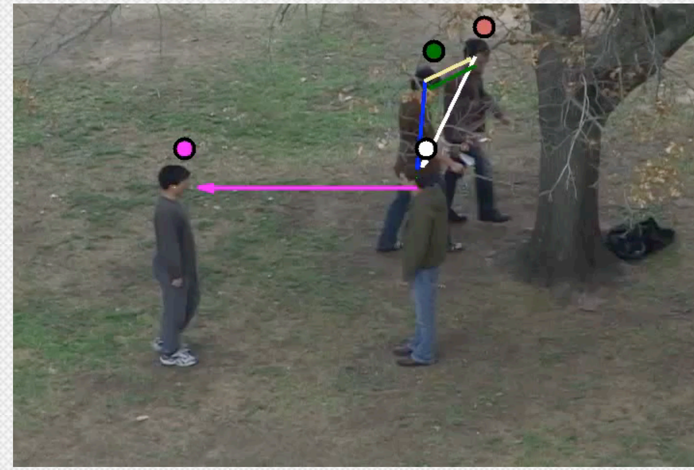

- 90.57% Average Accuracy
- Improvement of 20% over previous state of the art.

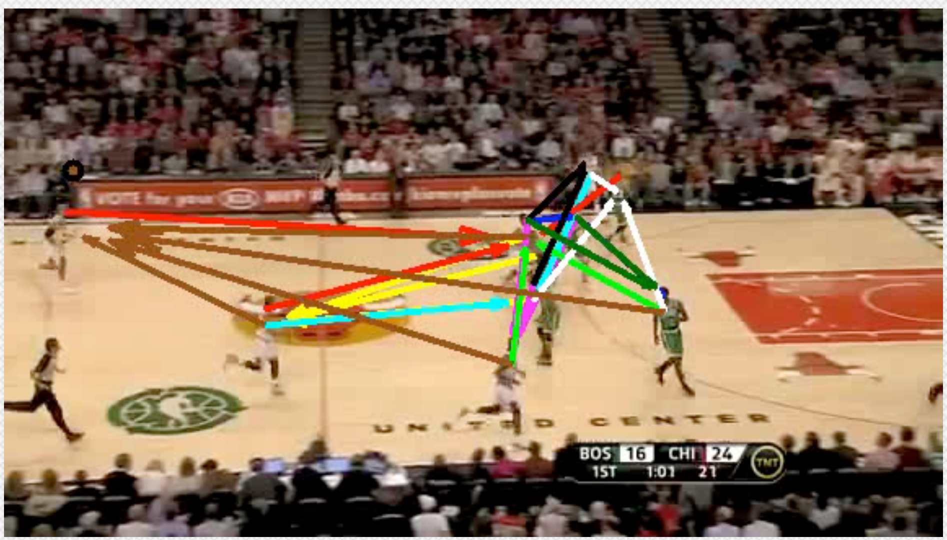
**(CVPR 12)**

NEXT: Crowded and clutter scenarios with multiple agents

# Coordinated Activities

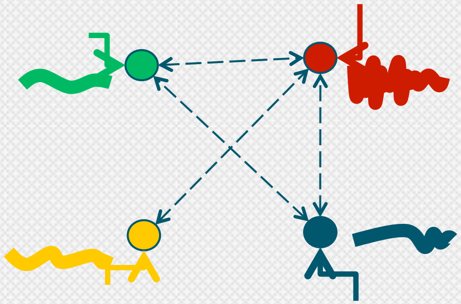
- **State of the Art**

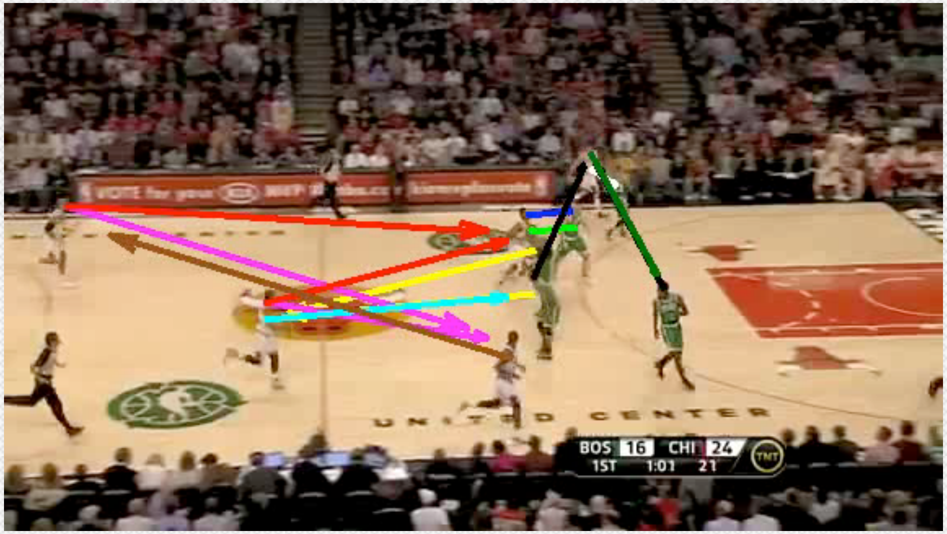




# Coordinated Activities

- **Using Sparse Dynamics:**

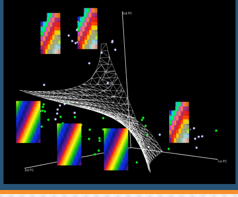





**(ICCV 13)**

Next: From seemingly normal individual activity to suspicious collective behavior.

# Who/What Where doing What



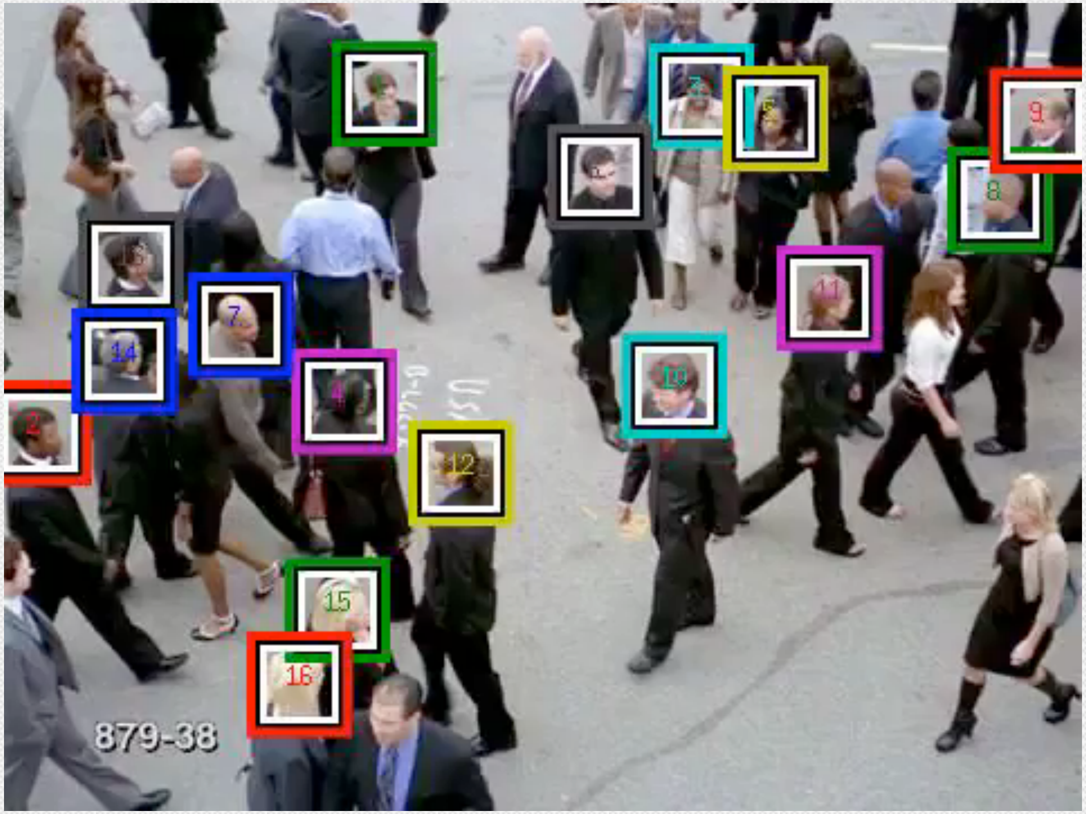

- Automatically joint segmentation and event detection
  - Uses both appearance and dynamics

NEXT: Threats (i.e. leaving bag behind) in crowded scenarios
- Real time implementations

# Video Analytics for TSA

- ***In-the-Exit Contraflow Detection***
    - Deployed at CLE Airport
        - Two different exits, different geometry
        - ~0 miss-detections, 1 to 2 false alarms/week
- ***Tag-and-Track across camera network***
    - In progress, to be deployed at CLE Airport soon
        - Two testbeds:
            - Parking to Terminal and
            - Checkpoint Exit to Three Concourses
- **Activity Recognition, Object Left Behind:**
    - Theory developed, tested on simple scenarios

**Dynamics as a key enabler to handle data deluge and obtain actionable information in a timely fashion.**




# THANKS!!