Learning Object Motion Patterns for Anomaly Detection and Improved Object Detection

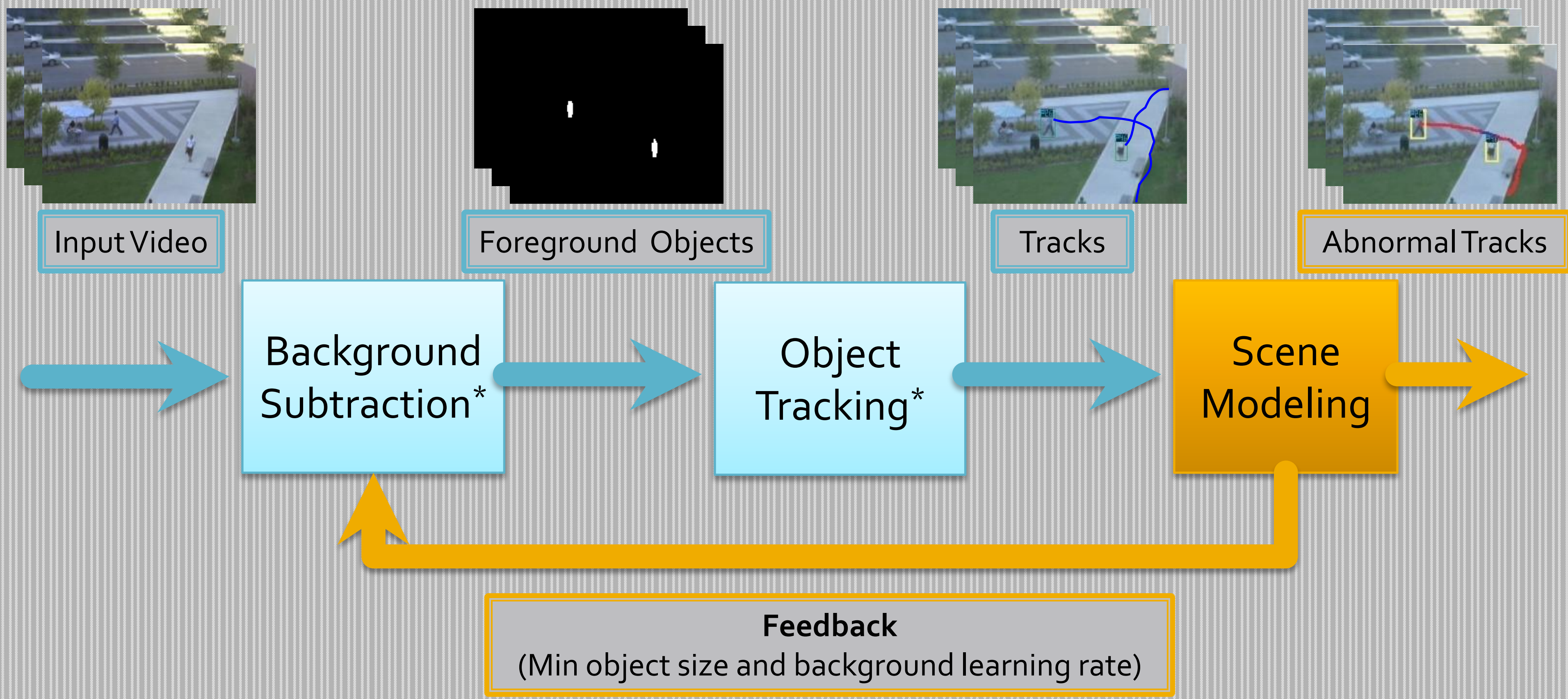
Arslan Basharat, Alexei Gritai, and Mubarak Shah

Presented at CVPR, June 2008 in Alaska

Motivation

- Various object detection & tracking approaches are available
- Need of higher level analysis
- Create scene model to
  - Detect abnormal behavior
  - Improve performance of the surveillance pipeline

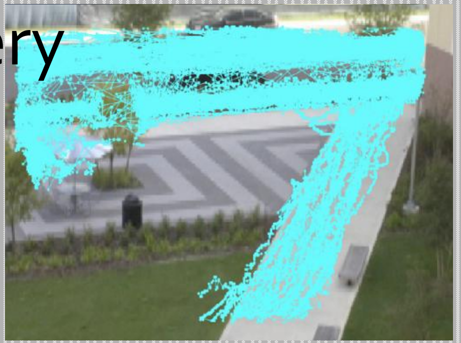
Proposed Approach



* UCF KNIGHT video surveillance system

Proposed Approach

- Learn a PDF of object *motion* *and* *size* at every pixel
  - Previously used for pixel intensities*
  - No need to explicitly cluster tracks into paths



- Detect abnormal events based on *local* and *global* behavior of tracks

- Scene model feedback
  - Minimum size of detected foreground blob
  - Background model learning rate

*C. Stauffer & W. Grimson, Adaptive background mixture models for real-time tracking, CVPR 1999

Model Learning

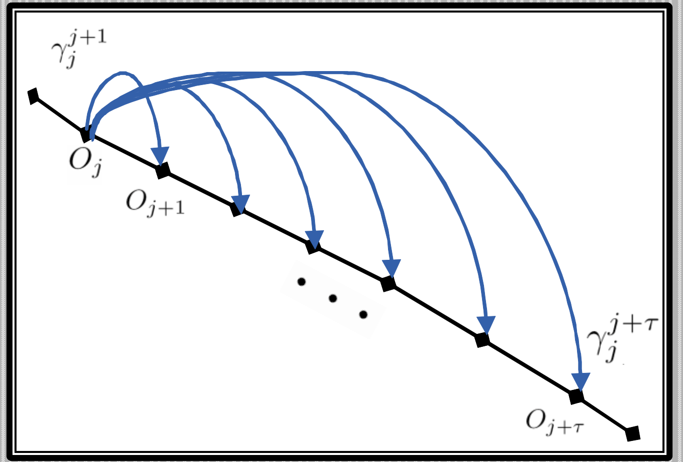
- Set of training tracks
- Track: set of observations

$$O_j = (x, y, t, w, h)$$

- Transition vector

$$\gamma = (x', y', \delta t, w, h)$$

  - Destination location
  - Transition time
  - Object width and height



Model Learning

- Using γ, create a PDF of normal behavior (motion & size) at every pixel
- Gaussian Mixture Model (GMM) of transition vectors at every pixel *l*

$$P(\Gamma_l = \gamma|\theta_l) = \sum_{i=1}^{n} \alpha_l^i p(\gamma|\theta_l^i)$$

$$p(\gamma|\theta_l^i) = \frac{1}{(2\pi)^{d/2}|\Sigma_l^i|^{1/2}} exp(-\frac{1}{2}(\gamma - \mu_l^i)^T {\Sigma_l^i}^{-1}(\gamma - \mu_l^i))$$

Model Learning

- Learn GMM parameters using EM*
  - E-step

$$\omega_l^i = \frac{\alpha_l^i(t)p(\gamma|\theta_l^i(t))}{\sum_{j=1}^k \alpha_l^j(t)p(\gamma|\theta_l^j(t))}$$

  - M-step

$$\hat{\alpha}_l^i(t+1) = \frac{max\{0, (\sum_{m=1}^S \omega_l^i(m)) - \frac{d}{2}\}}{\sum_{j=1}^k max\{0, (\sum_{m=1}^S \omega_l^i(m)) - \frac{d}{2}\}}$$

$$\hat{\theta}_l^i(t+1) = \arg\max_{\theta_l^i} Q(\theta_l, \hat{\theta}_l(t))$$

*M. Figueiredo and A. K. Jain, Unsupervised learning of finite mixture models. IEEE TPAMI 2002.

Model Learning

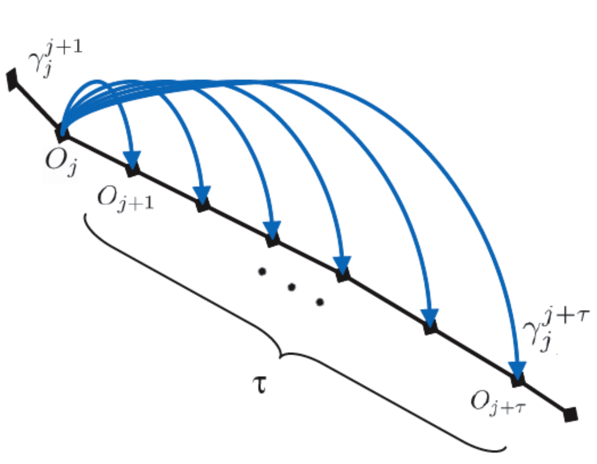
- Advantages of the proposed model
  - Unsupervised learning
  - Handles multiple paths at each location
  - Ability to perform online learning
  - Ability to marginalize to different parameters
  - Suitable for real-time surveillance system

Anomaly Detection

- Using $i = 1,\ldots,\tau$ transitions
- Use the least probable transition

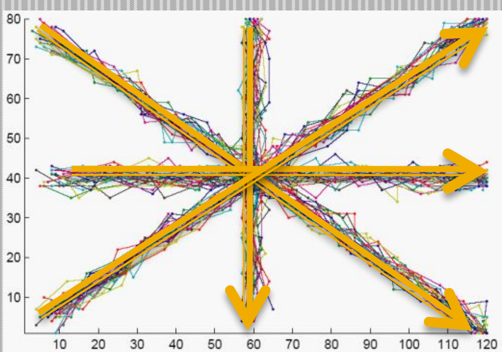
$$\beta_t = min_i P(\Gamma_{l(t-i)} = \gamma^t_{t-i})$$

$$\boxed{\beta_t < \lambda}$$

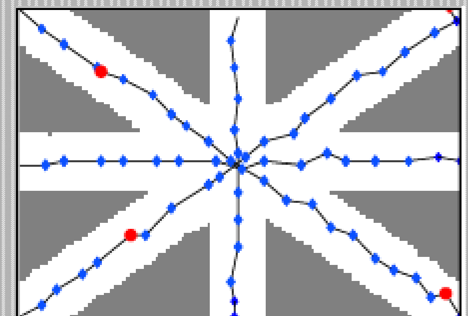


- Local vs. global anomalies
  - $\tau = 1$ vs. $\tau >> 1$

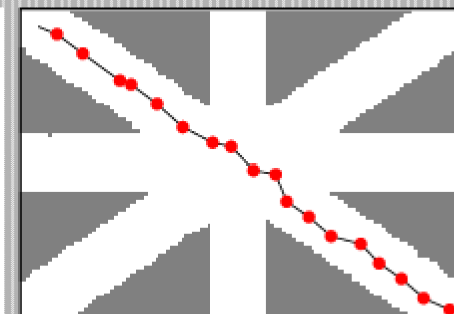
Synthetic Scenes



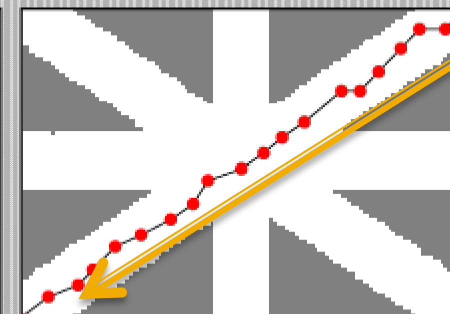
(a) Training Tracks (4 Paths)



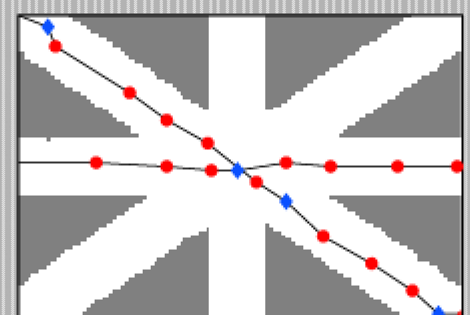
(b)Normal Tracks



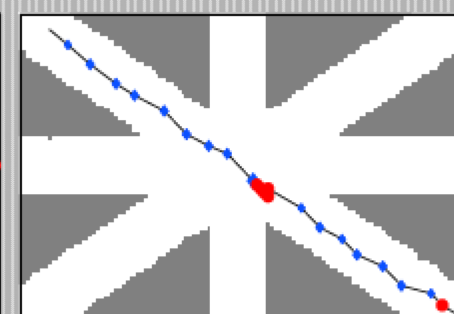
(c) Unusual Object Size



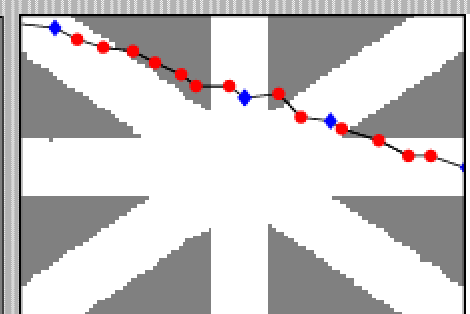
(d) One- way Violation



(e)Fast Motion



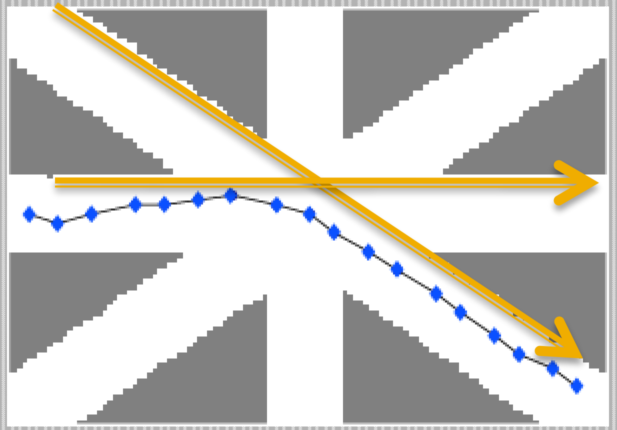
(f)Stopping Incident



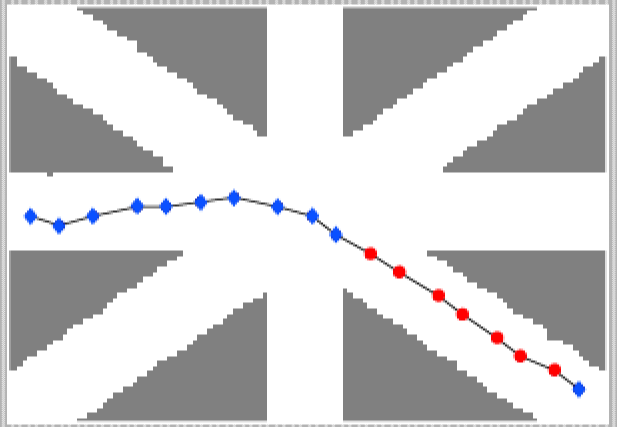
(g)Unusual Path

Global Anomalies

- Local analysis shows limited performance
- Common in previous approaches
- Global analysis (τ >> 1)
- Track structure is analyzed



(a) After local analysis



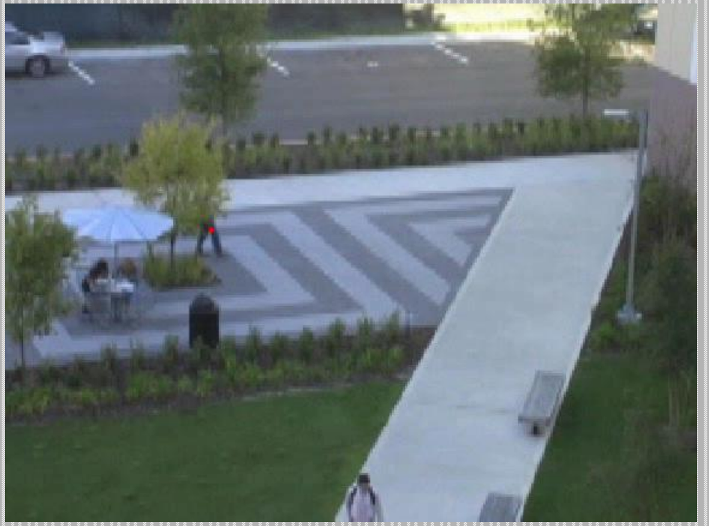
(b) After global analysis

Anomaly Detection

Pedestrian on the road

Skateboarder on the sidewalk

Bicyclist on the sidewalk

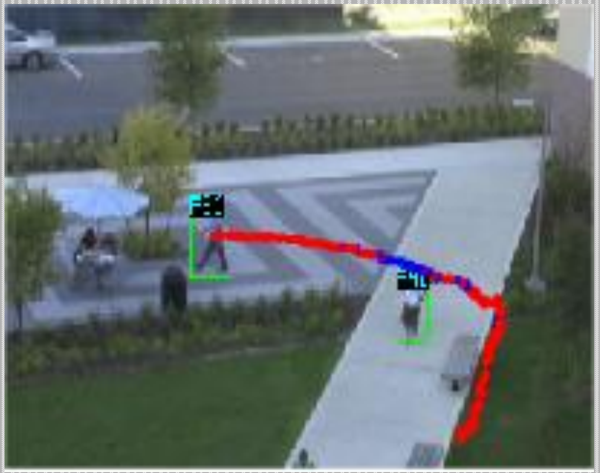


Unusual path

Anomaly Detection

Normal Track

Sitting on sidewalk



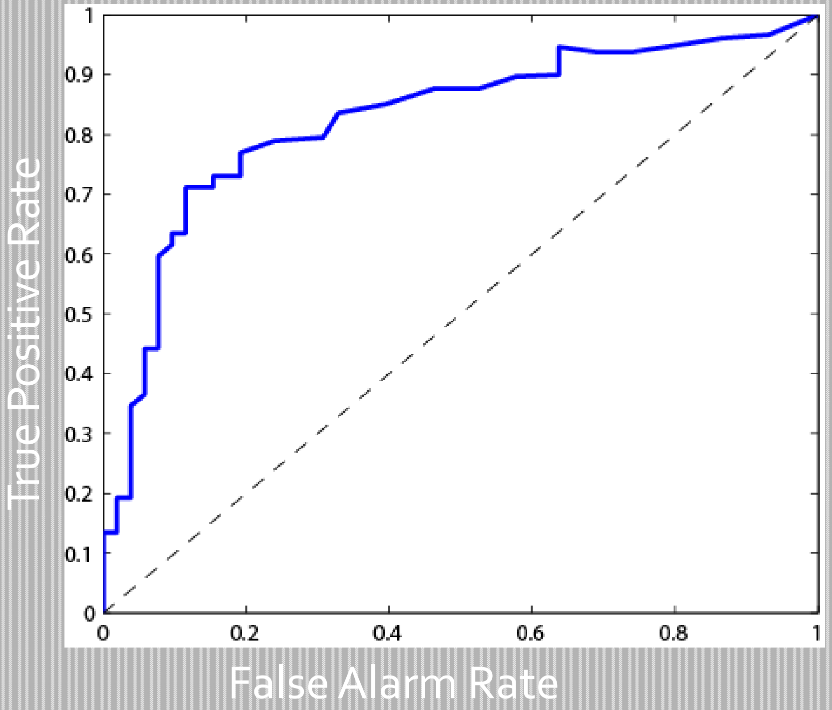
Unusual Path

Quantitative Analysis

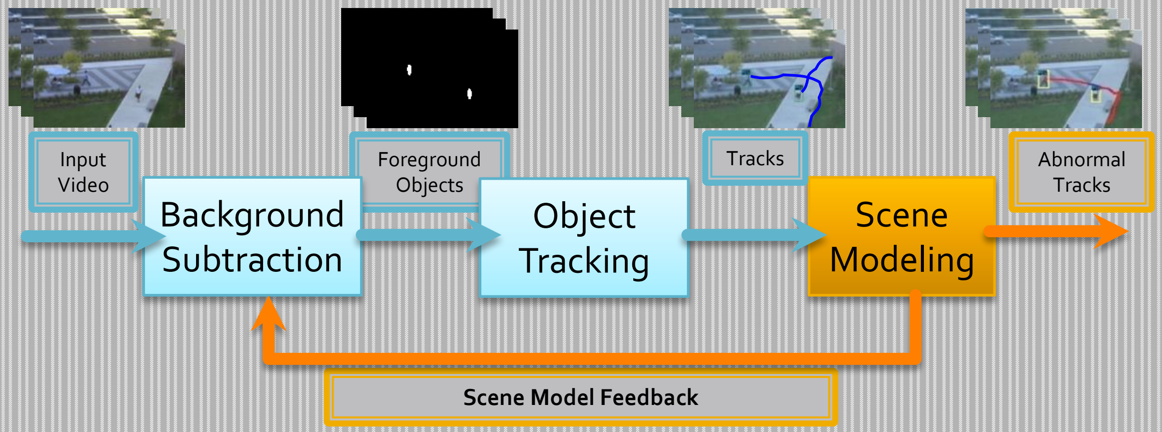
- Dataset used
  - Training: 90 mins sequence
  - Testing: 30 mins sequence

	Training	Testing
Normal Tracks	1342	217
Abnormal Tracks	0	31

- ROC Curve



Scene Model Feedback



- Learnt scene model
- Provide feedback for better performance
- Parameters of object detection
  - Minimum object size to detect
  - Background model learning rate

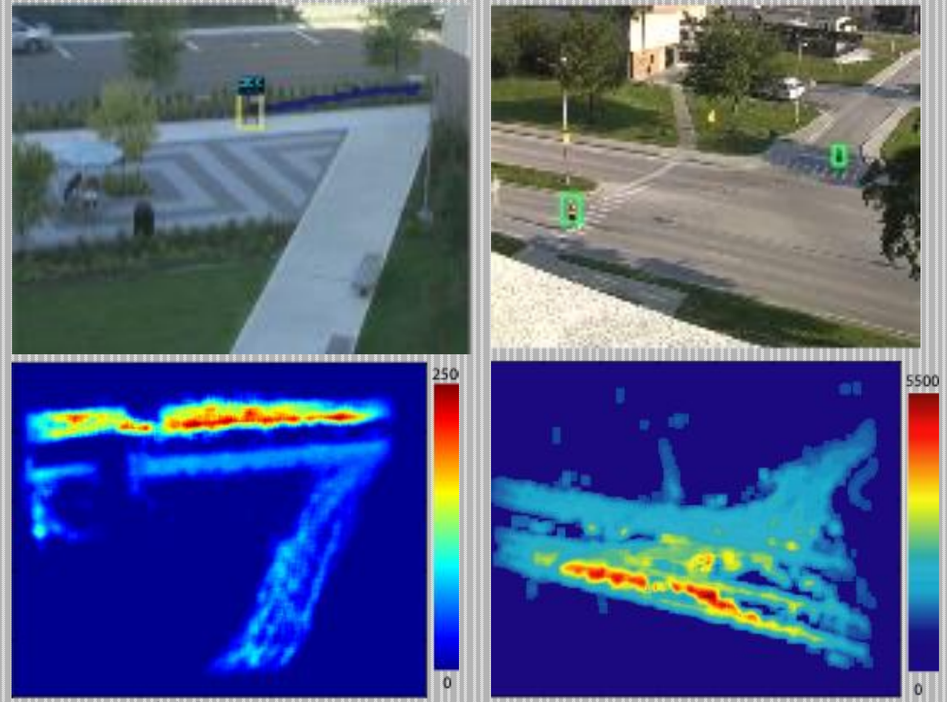
Feedback: Min Object Size

- Use the learnt PDF
- Marginal PDF of object size

$$P(w,h) = \sum_{x=1}^{C}\sum_{y=1}^{R}\sum_{t=0}^{T} P(x,y,t,w,h)$$

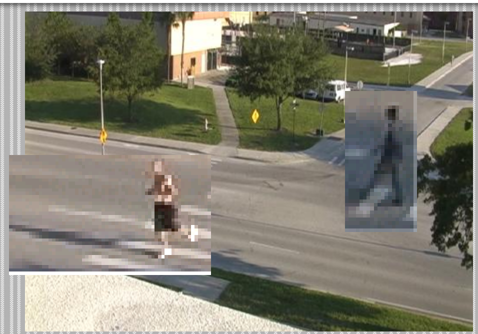
- Most probable size at every pixel
- Fixed vs. variable '*s*' in $[s_{min}, s_{max}]$

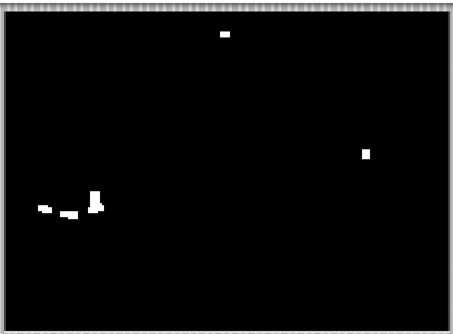
$$s = s_{min}P(w,h) + s_{max}(1 - P(w,h))$$



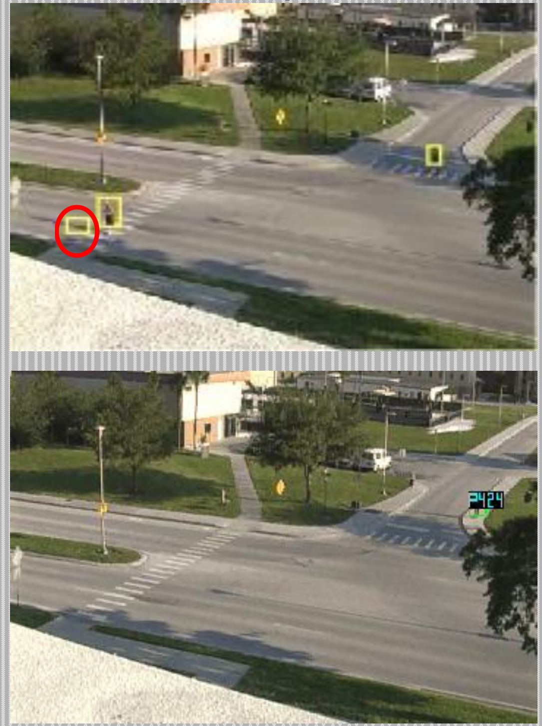
(a) Scene 1 (b) Scene 2

Feedback: Size Model

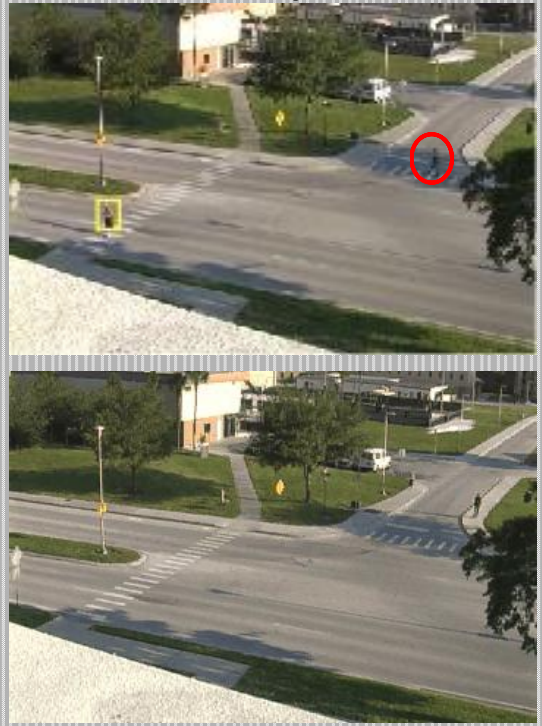




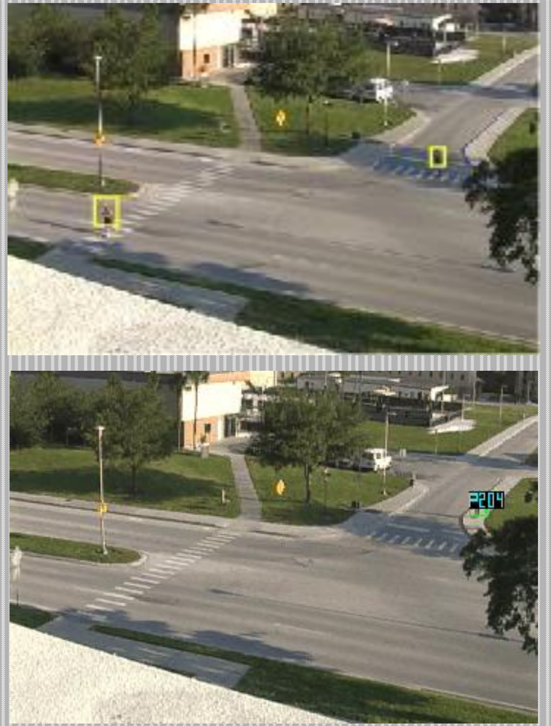
s = 50



s = 150



Variable 's'



Feedback: Min Object Size

s = 50

s = 150

Variable ‘s’

Feedback: Background Learning Rate ρ

- Marginal of speed
- Fixed vs. variable ρ
- Low ρ at low speeds

$$P(x,y,t) = \sum_{w=0}^{C}\sum_{h=0}^{R} P(x,y,t,w,h)$$

$$\rho = \rho_{min} P_v(u) + \rho_{max}(1 - P_v(u))$$

Fixed learning rate

Variable learning rate using feedback

Conclusion

- Proposed a new scene modeling approach
- Unsupervised approach
- No clustering of tracks into main paths
- Useful for various type of anomalies and scene model feedback
- Future directions
  - Modeling interaction of multiple tracks
  - Scene analysis in multiple cameras

Thank you!