ADSA15:
Attribute-Based Searching and 360^{o} Surveillance Video

Cindy Fang
fangc@ll.mit.edu
November 16, 2016

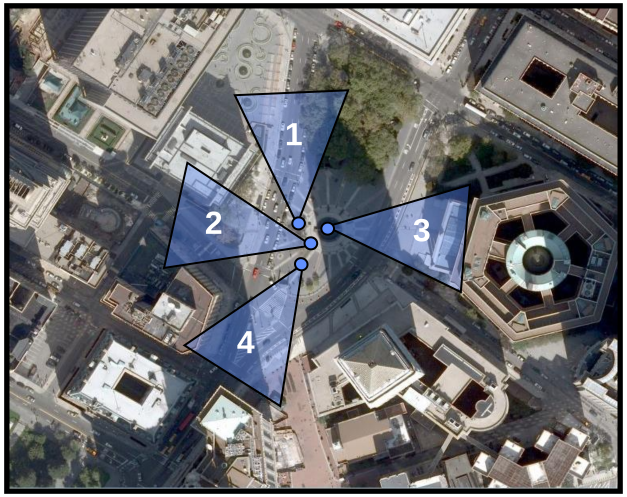
LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

So what? Who cares?

- **Topic Areas:**
  - **Video Surveillance Technologies and Video Analytics**
- **Challenges addressed**
  - **Video surveillance trade-off between coverage and resolution**
  - **Forensic video search for a person of interest**
- **What did we do?**
  - **Built a high resolution 360 degree camera system with COTS parts and demonstrated system at Logan Airport**
  - **Developed an attribute-based search tool**
- **Benefits:**
  - **High quality video with coverage everywhere ensures every incident is recorded**
  - **Improve forensic search time from hours to tens of minutes**

Critical Infrastructure Protection
Video Surveillance

Typical problem: Foley Square, NYC



Typical operator control room

- **Video surveillance challenges**
  - **Inadequate resolution**
  - **Disjointed situational awareness**
  - **Gaps in camera coverage**
  - **Operator fatigue / overload**

Immersive Imaging System (IIS)

Size	13.5" diameter
Resolution	1.2 cm @ 100m
Pixels	240 megapixels
Frame Rate	8 frames/second
Data Rate	1 TB / hour

- **360 degree coverage from sensor location**
- **Compact and lightweight**
- **On-board electronics and firmware**
- **Inexpensive components – cell phone imagers**

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

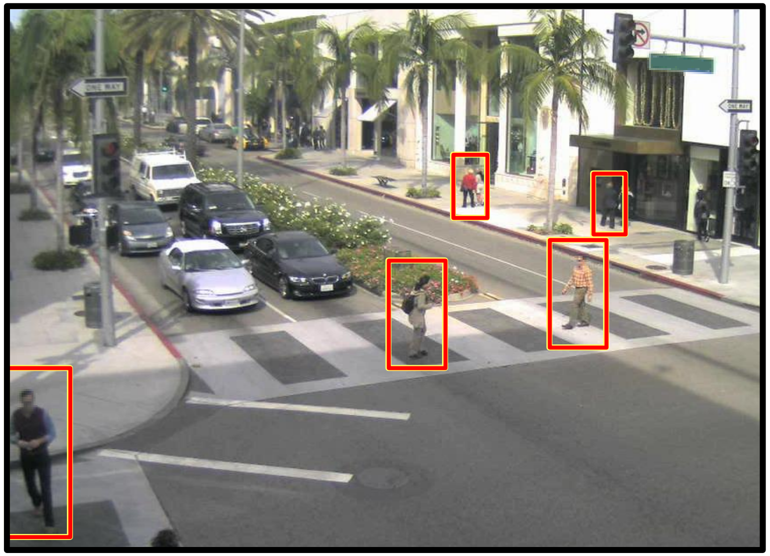
Logan Airport (Terminal A) Deployment

- **MASSPORT / TSA / Police Needs**
  - **Real-time scene monitoring and situational awareness**
    - **Comprehensive 360 degree coverage**
    - **Automated cuing of real-time events**
  - **Forensic incident review and suspect backtracking**
    - **Automated image search**

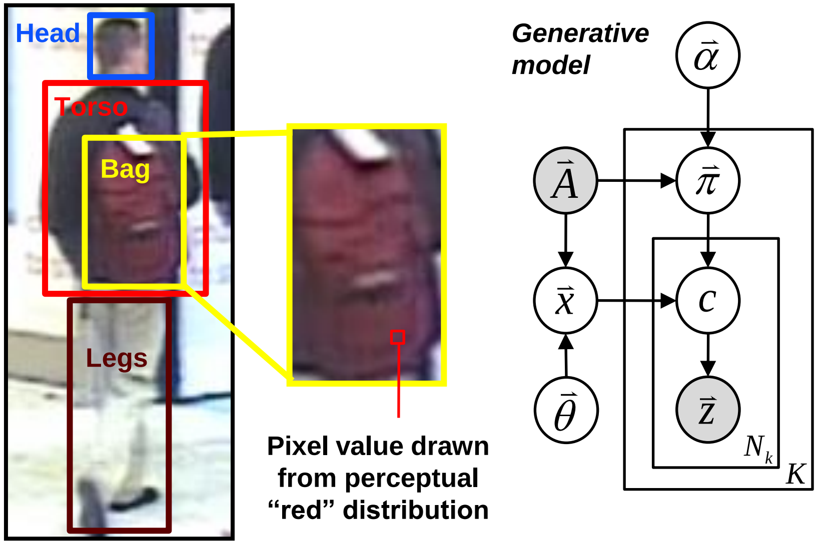
LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Attribute-Based Searching

Robust moving person detection

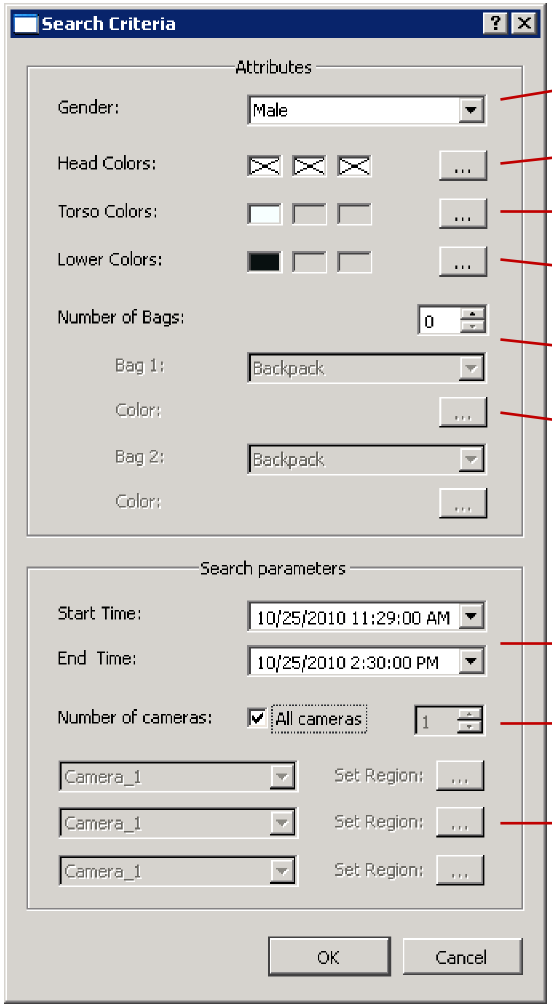


Probabilistic image interpretation



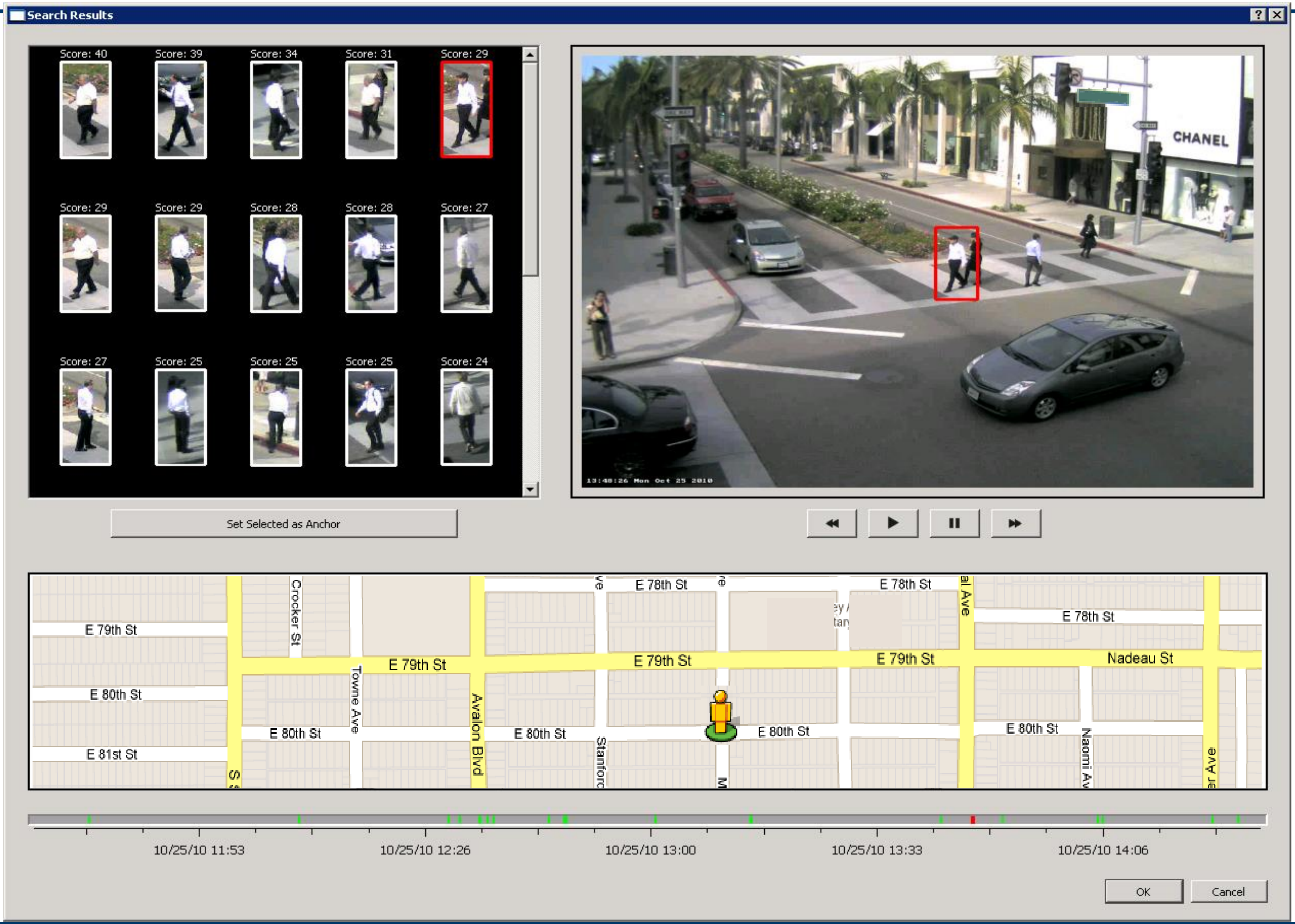
- **Key elements**
  - **Combining contour characterizations, foreground detection, and perspective information in order to get robust moving person detection**
  - **Defining flexible hierarchical probabilistic models for the appearance of persons and their sub-components**

Example : Search Input



- Gender
- Hair / hat color
- Torso clothing color
- Lower body clothing color
- Number and type of bags
- Bag color
- Search time window
- Cameras to search
- Search regions within camera views

Example : Search Results

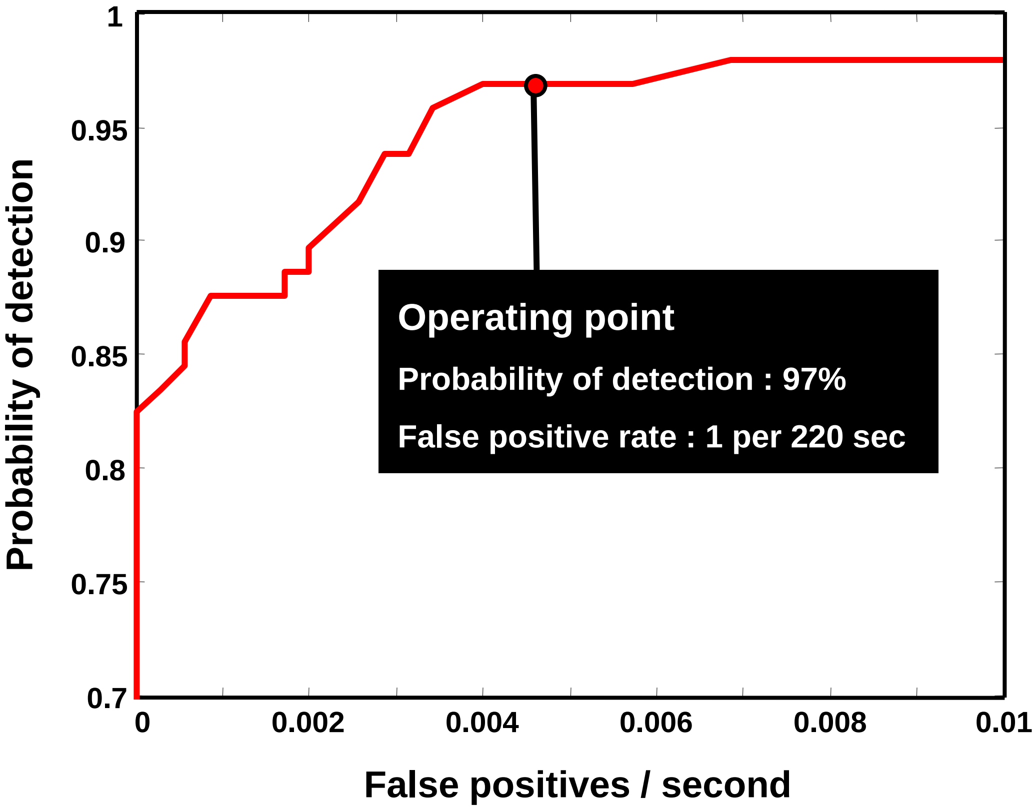


Person Detection Performance

Example surveillance region

- **Unobstructed view**
- **Average person height: ~100 pixels**
- **Moderate pedestrian flow**

ROC curve for person detection

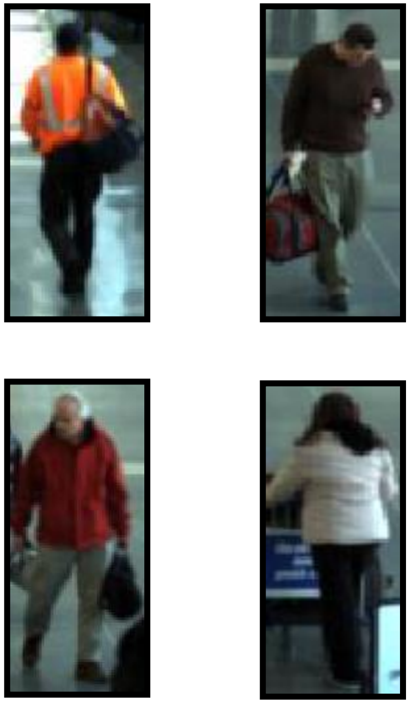


- **Person detection works very well for indoor views with moderate crowd flow and sufficient pixels-on-target**

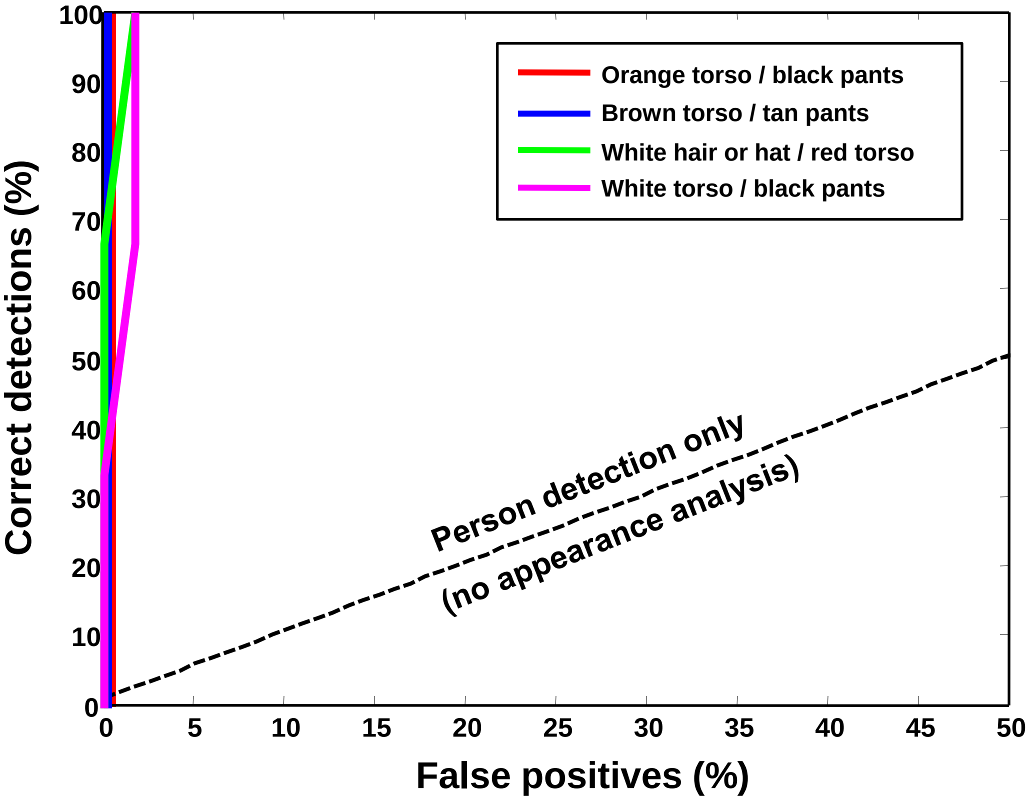
Attribute Scoring Performance

- **Label attributes for all persons who pass through one area of terminal**
- **Execute sample searches, then count missed detections or false positives**

Sample search targets



ROC curves



Summary

- **Built high-resolution 360 degree video surveillance system for critical infrastructure protection**
  - **Demonstrated system at Logan Airport and other public spaces**

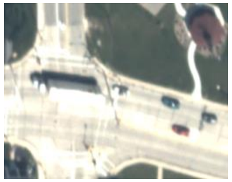
- **Developed attribute-based search application**
  - **One of many video analytics applications from MITLL**

This material is based upon work supported by the Department of Homeland Security under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Department of Homeland Security.

Backups

IIS System Goals

- **Capture high resolution imagery to detect and identify vehicles and people**

	Detection	Identification
Vehicles	**50 cm resolution** 	**13 cm resolution** 
People	**20 cm resolution** 	**1.2 cm resolution** 

240 million pixels for 360º coverage within 100m

- **Provide wide-area persistent camera coverage (240 Mpixels) - Challenges**
  - **High pixel count sensor**
  - **Efficiently manage and store data**
  - **Provide useful forensic and tactical tools to assist the user**
  - **Potential to be cost effective**

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Program Information

- **Immersive Imaging System (IIS)**

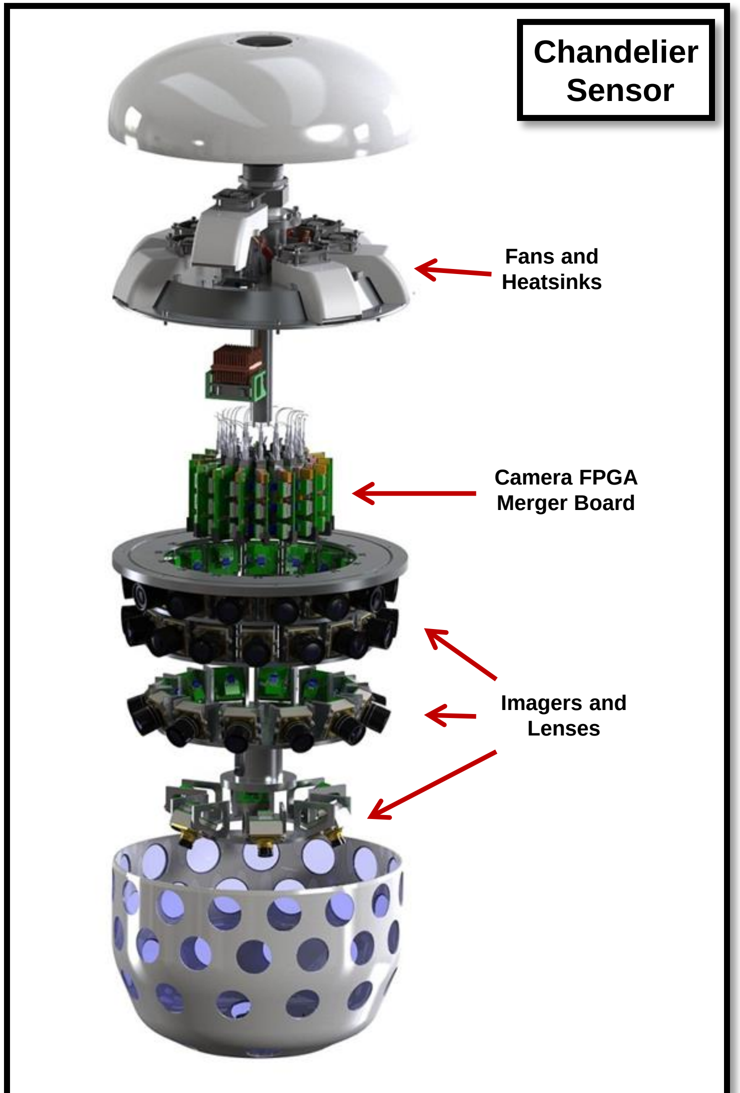
- **Sponsored by DHS S&T**
  - **POC: John Fortune**
- **Partners: Pacific Northwest National Labs**
  - **POC: Douglas MacDonald**
- **Patents Issued:**
  - **US Patent No. 9007432: Imaging Systems and Methods for Immersive Surveillance**
  - **US Patent No. 9036001: Imaging System for Immersive Surveillance**

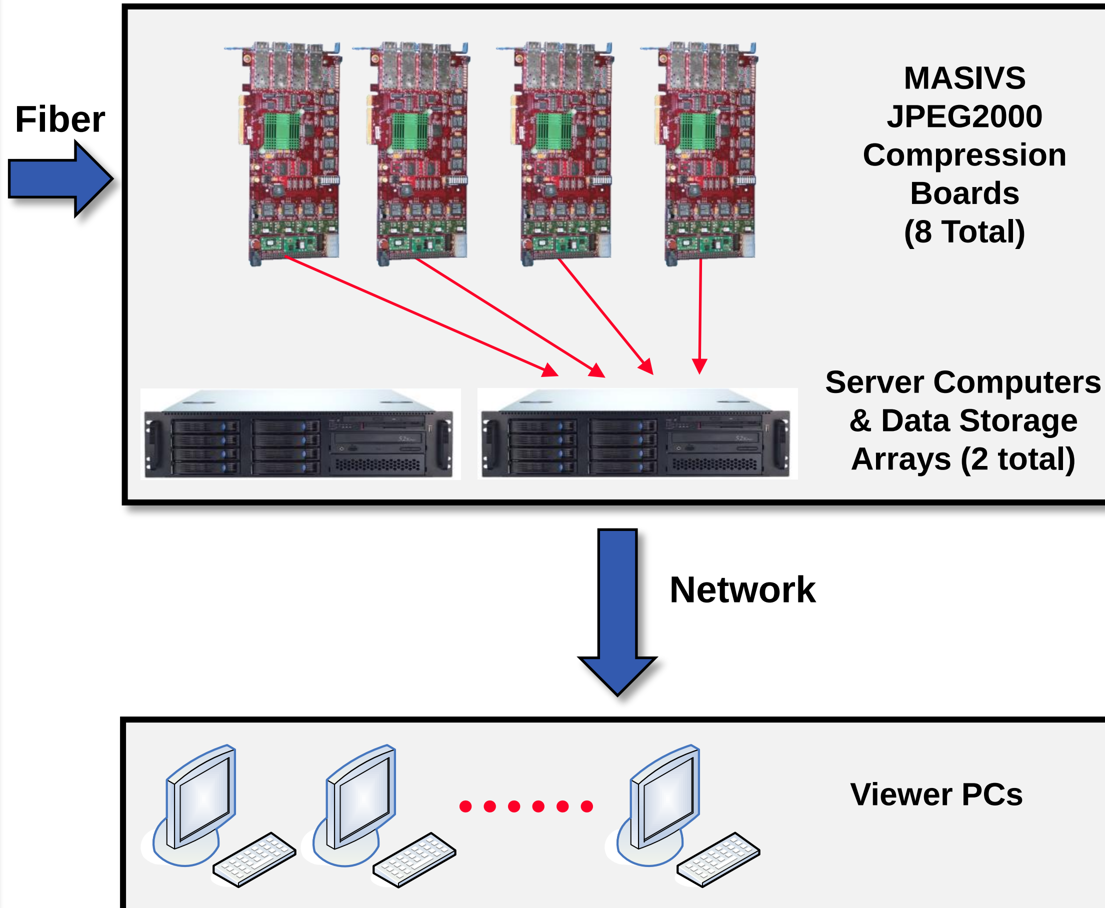
Homeland Security
Science and Technology

Pacific Northwest
NATIONAL LABORATORY

R&D Magazine Top 100 | **Popular Science** | **C4ISR Top 5 Award**

IIS Spiral 3 System Components





Seam Calibration

- **For each camera pair**
  - **Using nominal parameters approximate overlap**
  - **Extract common features**
  - **Find matching features in feature space**
  - **Refine matches using RANSAC algorithm assuming a projective transform**

- **Optimize over all parameters using error in projected features**

SIFT Features

Unprocessed IIS Image

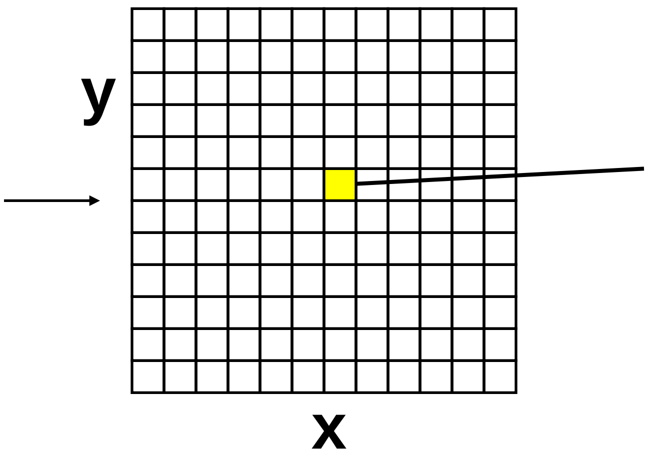
Processed IIS Image

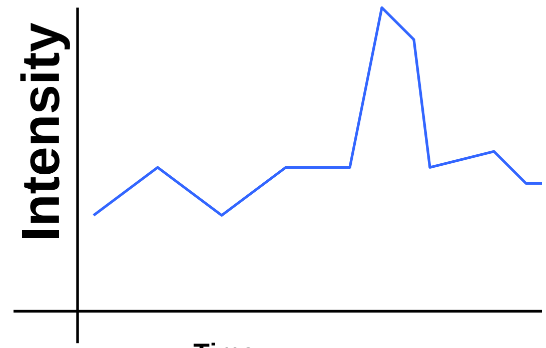
Mover Detection

- **Leverage detection algorithms from Constant Hawk**
  - **Adaptive statistical background modeling**

Rendered Image

Bin to increase SNR





Online computation of Mean (M) and Variance (σ^2)

α **= learning rate = 0.2**

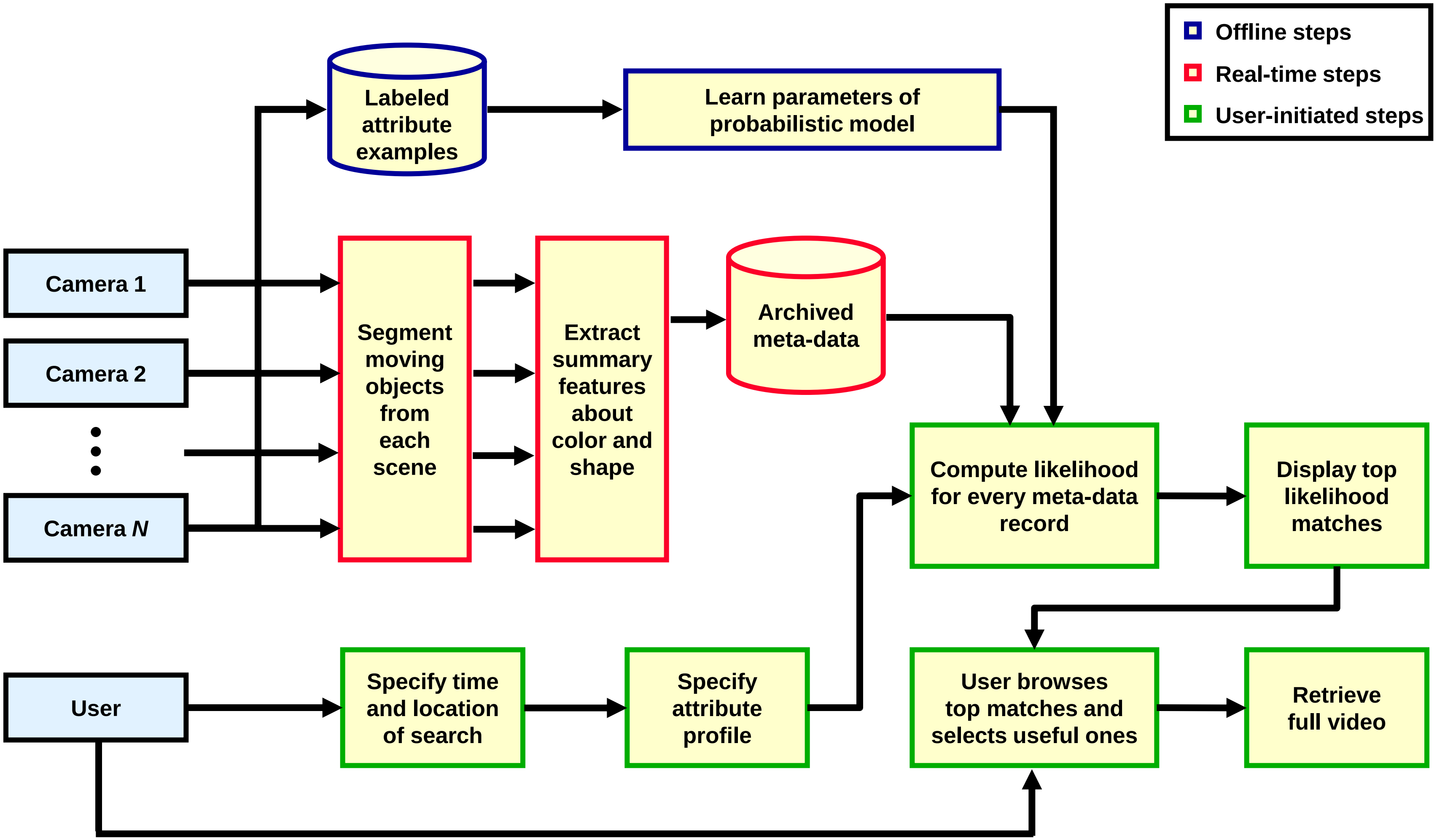
$$M_t(x,y) = \alpha I_t(x,y) + (1-\alpha) M_{t-1}(x,y)$$

$$\sigma_t^2(x,y) = \alpha \sigma_{t-1}^2(x,y) + (1-\alpha)\left(I_t(x,y) - M_t(x,y)\right)^2$$

Test: if I(x,y) > N_{std}*σ, I(x,y) is a foreground object **N_{std} = 5.5**

MIT LL Approach

Algorithm flow chart:

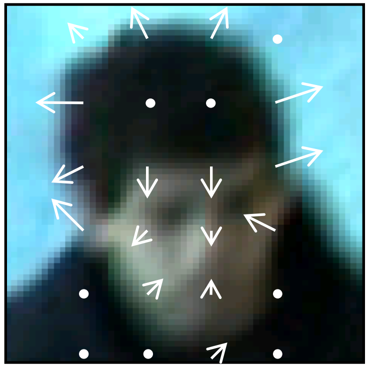


Robust Person Detection

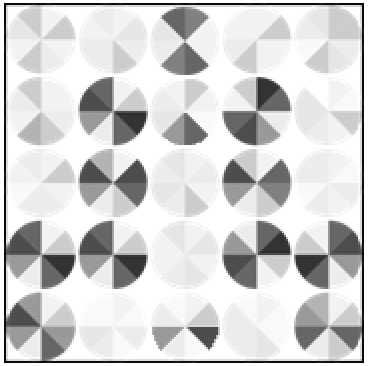
Homeland Security
Science and Technology

Contour analysis

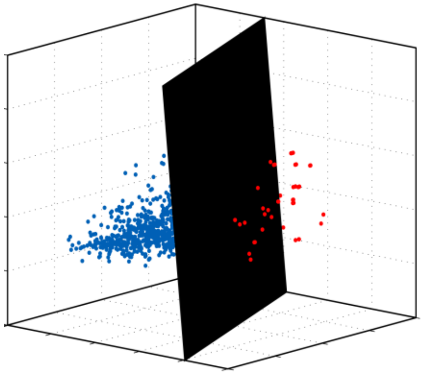
Compute image gradients



Compute local oriented gradient histograms



Apply SVM classifiers (trained on sample data)

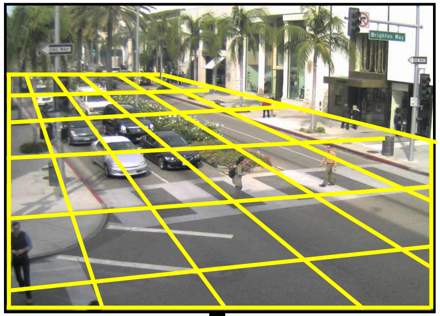


Male vs. female classification

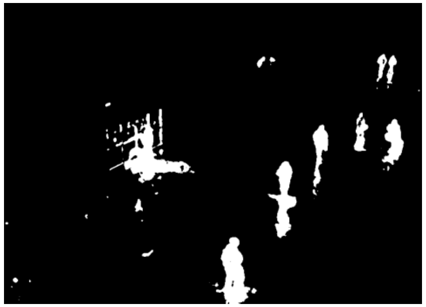
Person vs. non-person classification

Complete detection process

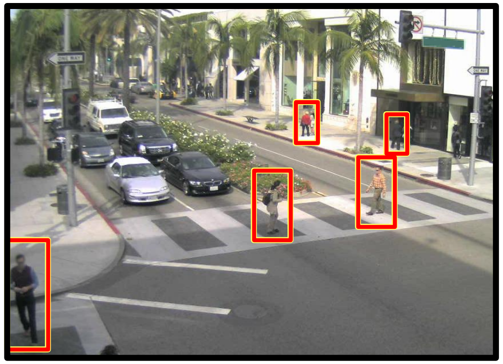
Ground plane estimation



Foreground estimation



Fused detections



LINCOLN LABORATORY
Massachusetts Institute of Technology

Gender Classification

Homeland Security
Science and Technology

Male Samples

Female Samples

- **Retrain local gradient classifier for gender classification based on labeled examples**
- **Resulting classifier relies on contour characteristics (frame, hair, clothing, etc.)**

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Image Interpretation

Model concept

- Given a set of attributes, appearance of person is result of a hierarchical generative process
- At each level of hierarchy, there is a defined probability of generating observations at next level

Generative model structure

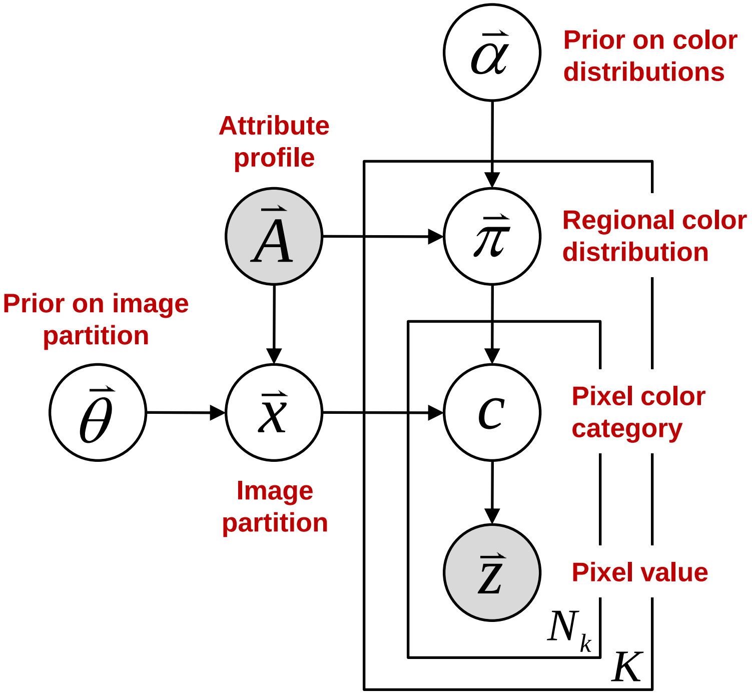


Illustration of hierarchy