Background Models

Slides adapted from Robert Pless

Object Detection

- Images
  - “Find some object”
- Video
  - “Find some (moving) object”
- What’s the difference?

Background Subtraction

- Usually a pre-processing step
  - Object/Anomaly Detection
  - Object Tracking
- AKA Foreground Segmentation

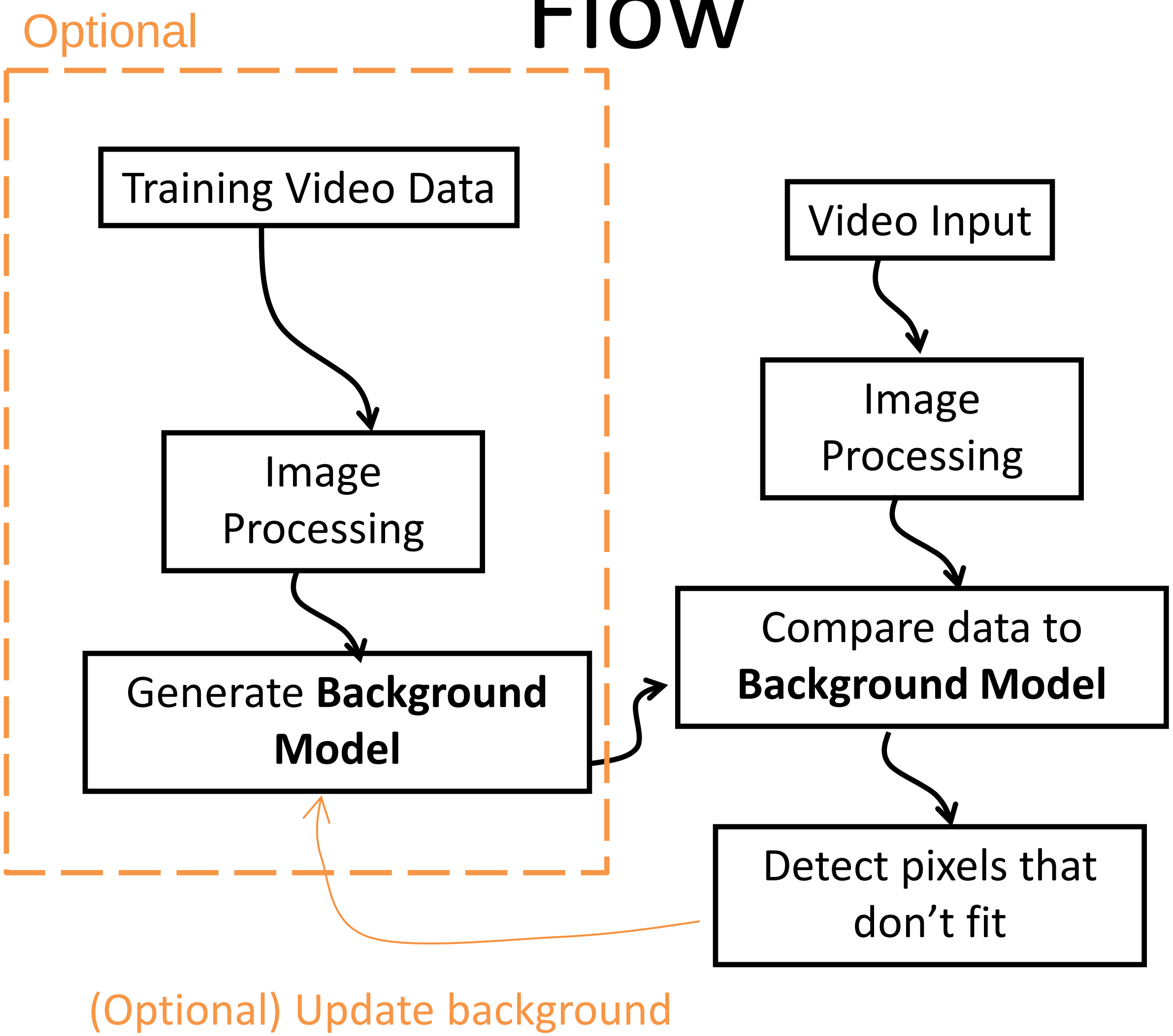
Original Video Frame

Detected Foreground

Background Subtraction

- Background subtraction requires:
  - Representation of the background
    - Background Model
  - Method for learning/updating the background
    - Use recent images
    - Average of training data
      - perhaps incremental averaging

Flow


Optional
Training Video Data
Image Processing
Generate **Background Model**
Video Input
Image Processing
Compare data to **Background Model**
Detect pixels that don't fit
(Optional) Update background

Example Video

Simple Approach

- Assume difference between frames t and t-1 represent motion
- Small differences may be due to imaging noise
  - Use a threshold
- Called *adjacent frame differencing*

Adjacent Frame Differencing

Adjacent Frame Differencing (T = 5)

Adjacent Frame Differencing (T = 20)

Adjacent Frame Differencing (T = 100)

(Another) Simple Approach

- Use a training phase to learn the “average” image
  - Average value at each pixel

- Again, use a threshold (2 options)
  - Fixed (as before)
  - Learned (some pixels vary more than others)

Mean & Difference
(T = 20)

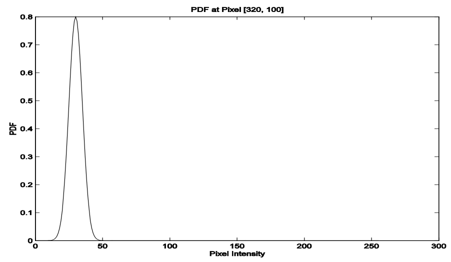
Mean & Difference
(T = 3σ)

Questions

- What are the free parameters to these “simple” approaches?
- When is a background *image* insufficient*?*
- What else could we do (without getting too crazy)?

Statistical Background Image Modeling

Probability Density Function (PDF) of the single pixel's intensity over time

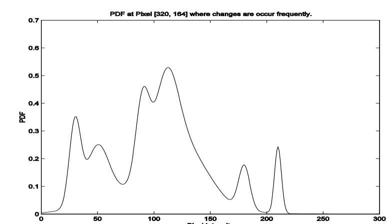


Intensity of the pixel

Sometimes, the PDF of many pixel intensities could be reasonably approximated by a simple model (like a Gaussian, represented by a mean and a variance)

Statistical Background Image Modeling

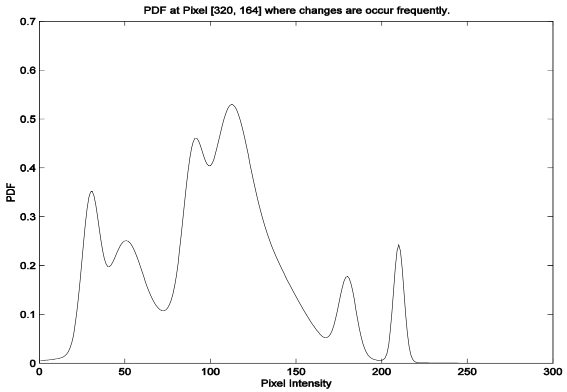
- Some pixels have more complicated appearances.
  - PDFs could not be well modeled by the a single Gaussian.
- Subtracting the “mean” of this distribution would give you a number,
  - Not really meaningful
- Any Ideas?



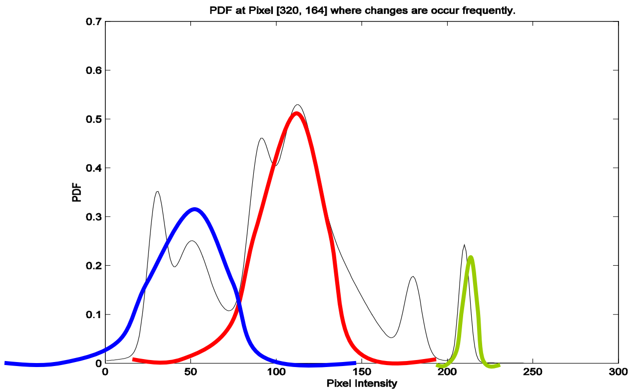
PDF of the single pixel’s intensity over time

Idea #1

- If one Gaussian doesn't work, perhaps 2 (or 3) do.



~



- Known as a *Gaussian mixture model*
  - Can approximate any PDF with enough Gaussian components.
  - We need strategies for updating these models…

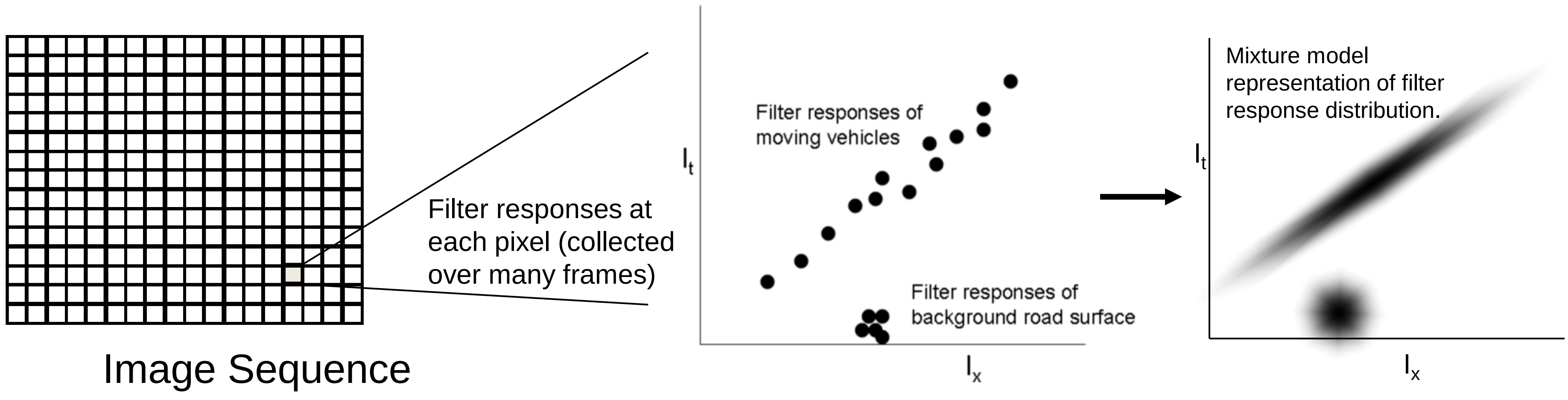
Representing Probability Distributions

- **Easy (in the context of streaming video):**
  - Solve for best fitting (arbitrarily oriented, multi-dimensional) Gaussian.

- **Harder to represent distribution as mixture model**
  - Solving for multiple models with EM (an iterative process which requires maintaining all the data)
  - Leads to adaptations of EM to streaming data:
    - Adaptive Mixtures
    - Cascading Models

Multi-modal models

- If the backgrounds are moving, are there properties that make their distribution easier to represent?

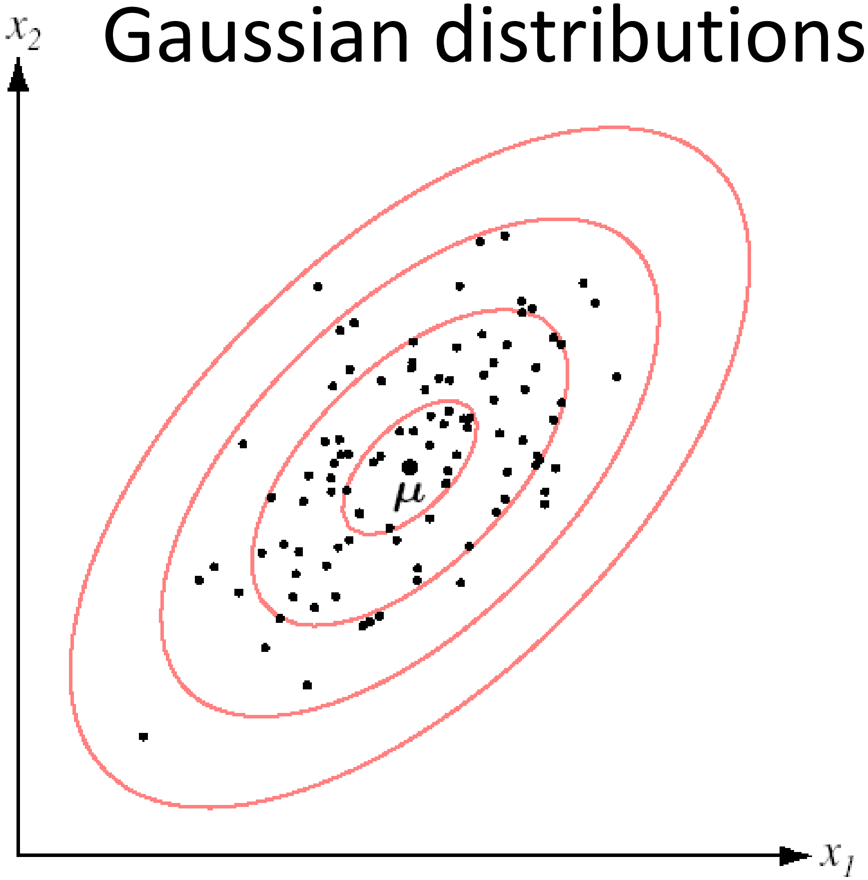


Adaptive mixture models

- Have several Gaussians (μ,Σ)
- Given a new measurement m:
- How likely is this measurement?
  - $\text{Score}_i=(m-\mu_i)\Sigma_i^{-1}(m-\mu_i)$
  - If for all i, score_i is large then unlikely to fit any model (mark as independent)
  - Otherwise, update just the model that has smallest score

Covariance

- Implies ellipsoids of constant probability for Gaussian distributions



2-D Gaussian

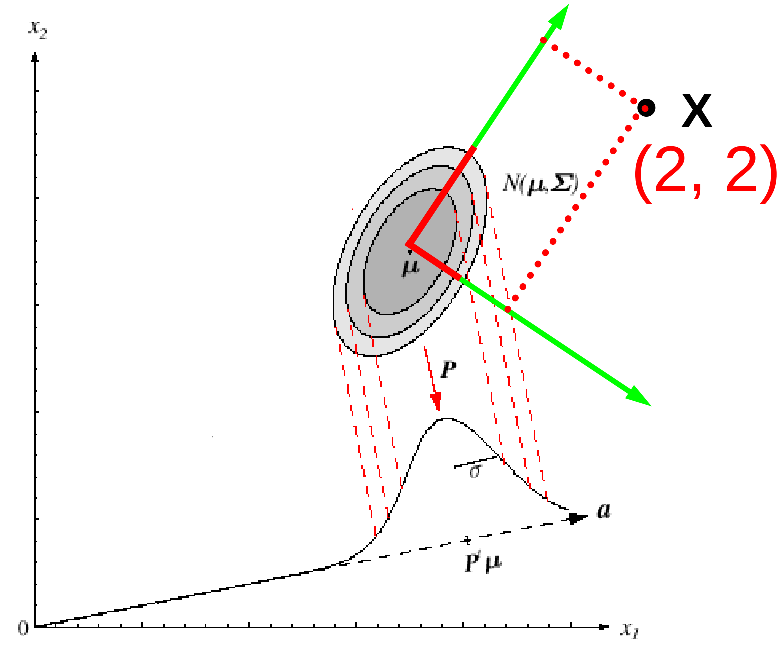
from Duda *et al.*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

1-D Gaussian

Mahalanobis Distance

- Distance of point from distribution
  - Along axes of fitted
  - In units of standard scaled)

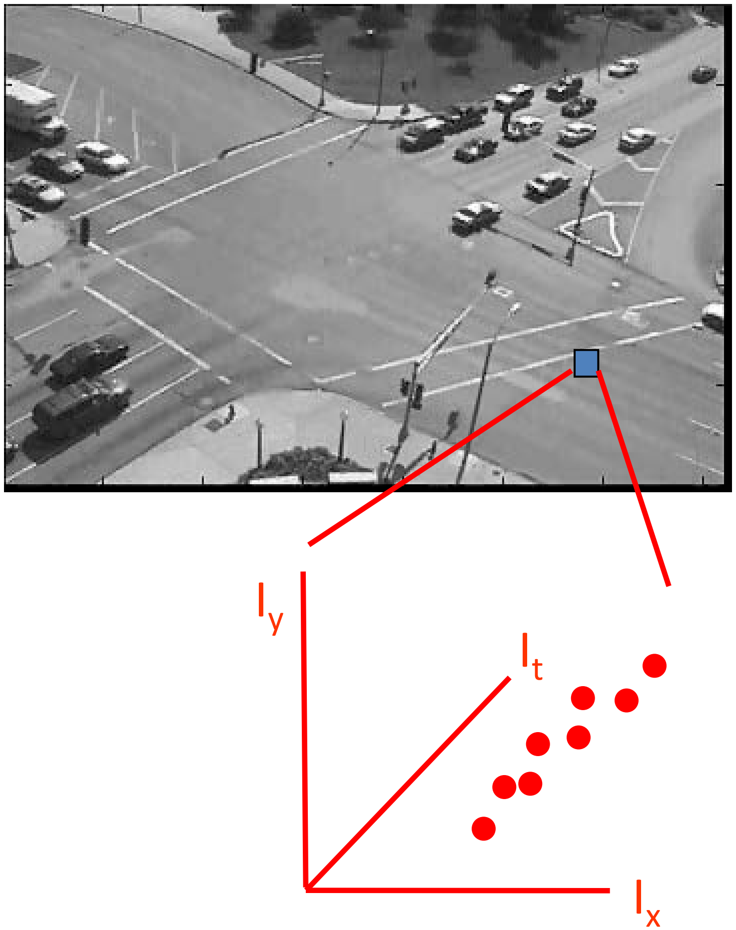


adapted from Duda & Hart

covariance matrix

$$d^2 = (\mathbf{x} - \mu)^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \mu)$$

Spatio-temporal Background Model

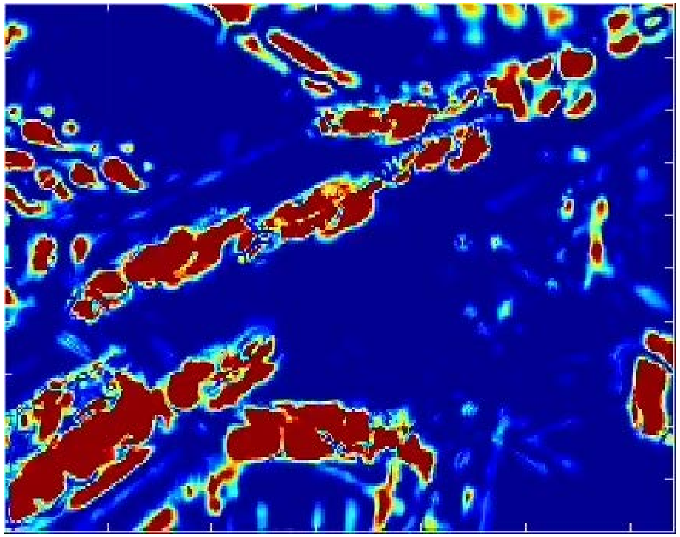


- Instead of intensity, use 3D measurement of I_x, I_y, I_t.
  - Estimate probability distribution of these measurements at each pixel.
- Identify measurements that are "unlikely" to come from background as independent motion

Building the Background Model

- In areas where the background motion is consistent, there is a more specific background model
  - This will allow more independent motions to be correctly marked

St. Louis intersection



Specificity of Background Model

Video Surveillance…

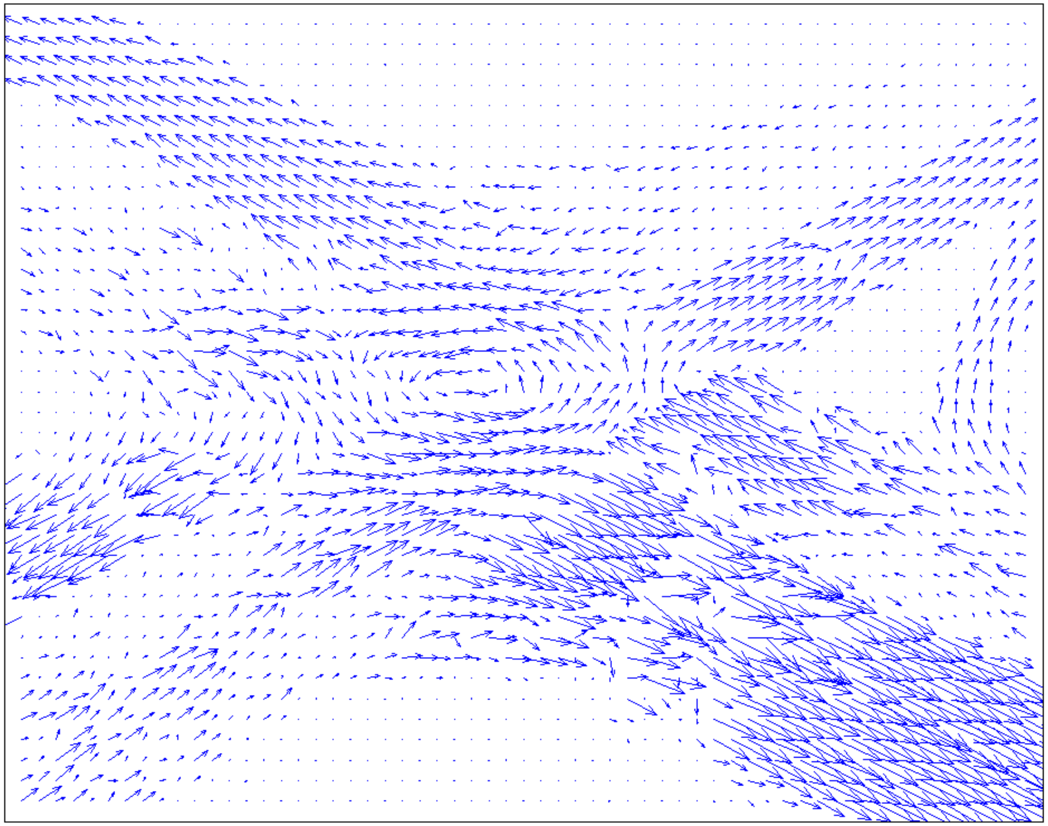
Detection

- During normal traffic patterns, build up background models

Intuitive Background Model…

- Build Model of Typical Background Motion

- Mark areas of video that don't fit that model



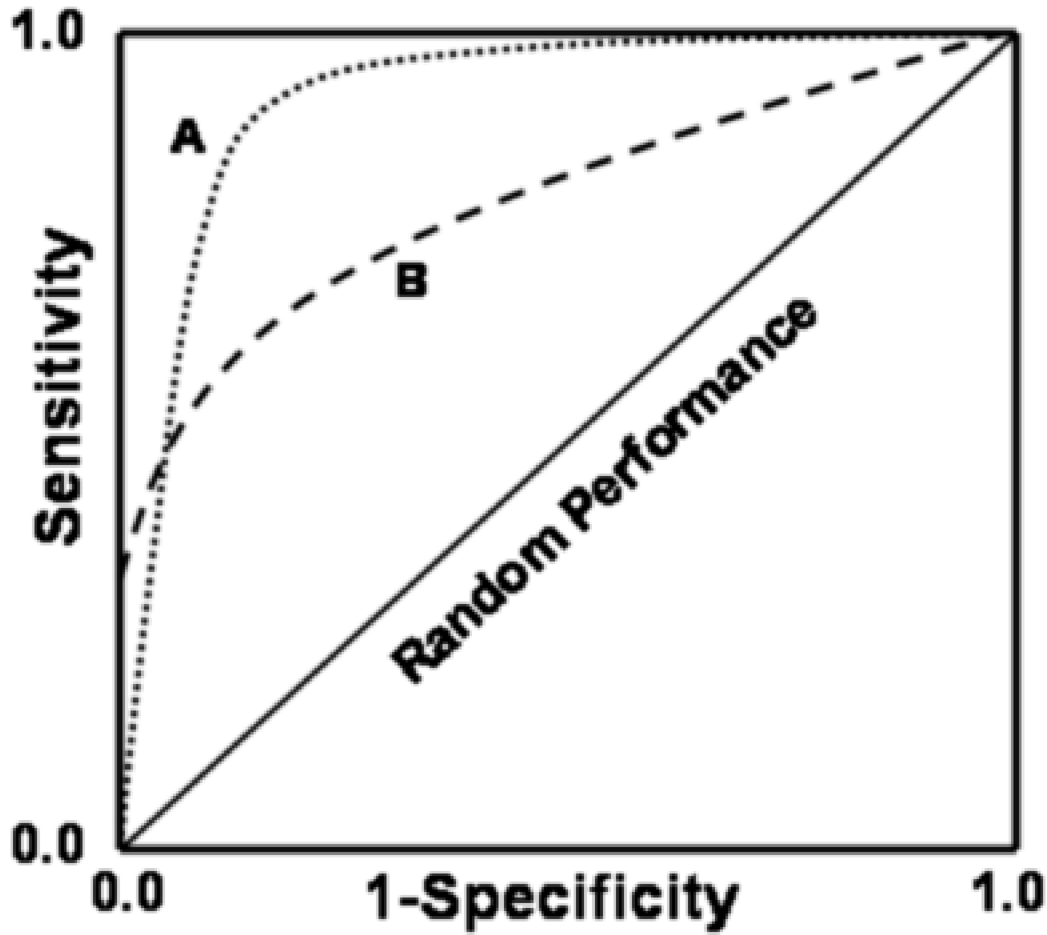
Unusual Traffic Motion

More Dynamic Background Subtraction

Training

Measuring Performance

- ROC Curves
  - Receiver Operator Characteristic (ROC) Plots
  - Depicts how well a particular score function can be used to classify measurements as foreground or background



Summary

- Background Subtraction
  - Model the background
  - Anything that doesn't fit the model is foreground
  - Usually a pre-processing step for *something else*
- Background Models
  - Simple: Previous frames, mean value
  - Most common: Gaussian Mixture Model (GMM)
  - Beyond pixel intensity: motion-based model