Change Detection, Skin Detection

Lecture-10

Motivation

- Detection of interesting objects in videos is the first step in the process of automated surveillance.
- Focus of attention method greatly reduces the processing time required for tracking and activity recognition.

Introduction

- Objectives:
  - Given a sequence of images from a stationary camera identify pixels comprising 'interesting' objects.
  - All independently moving objects are interesting!
- General Solution
  - Model properties of the scene (e.g. color, texture e.t.c) at each pixel.
  - Significant change in the properties indicates an interesting change.

Introduction

Problems in Realistic situations:

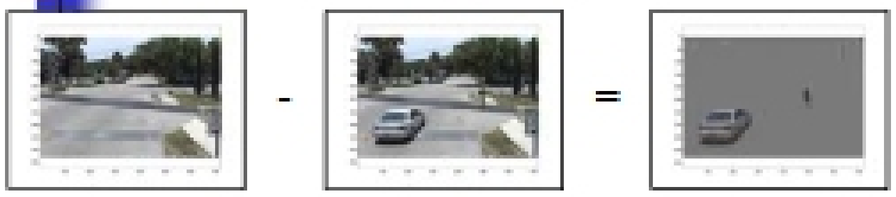
- Moving but uninteresting objects
  - e.g. trees, flags or grass.
- Long term illumination changes
  - e.g. time of day.
- Quick illumination changes
  - e.g. cloudy weather
- Shadows
- Other Physical changes in the background
  - Dropping or picking up of objects
- Initialization

Segmenting Background

Difference Pictures

- Jain, R. and Nagel, H. 1979. ``On the analysis of accumulative difference pictures from image sequences of real world scenes". *IEEE Trans. on Pattern Analysis and Machine Intelligence 1,* 2, pp 206-214.

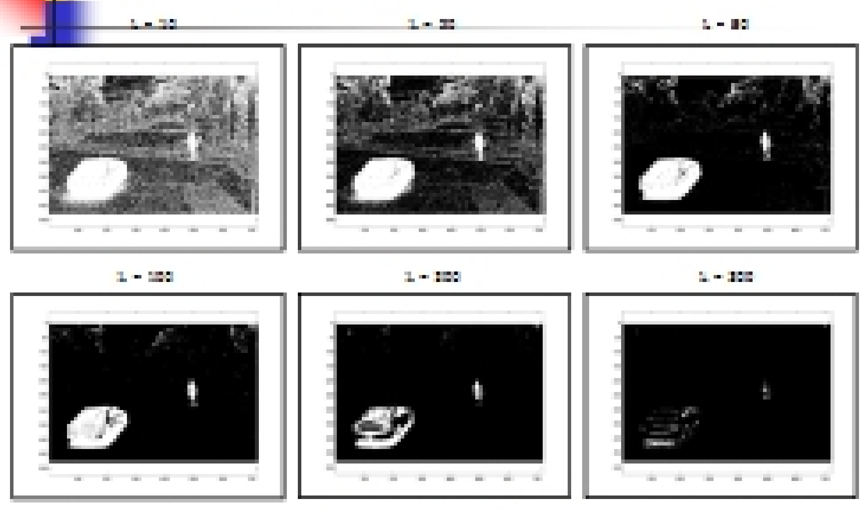
Background Subtraction



- Problem: Choosing a threshold
  - Pixel is foreground if $I_1(x,y)-I_2(x,y) \leq \lambda$ otherwise background?
  - What is the correct value of λ ?

UCF Computer Vision Lab. 7

Setting a Threshold



UCF Computer Vision Lab. 8

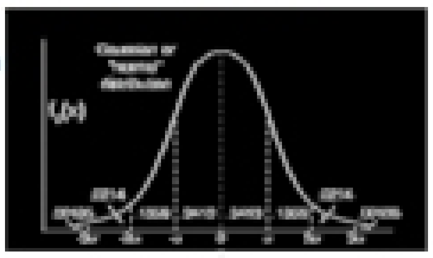
MODELING PIXEL INTENSITIES WITH A NORMAL DISTRIBUTION

Each pixel intensity can be modeled by a Normal Distribution, defined in terms of a mean μ and variance σ^2, as $N(\mu,\sigma^2)$. μ and σ are called parameters.

Useful when you wish to establish membership of a pixel to one of several models.

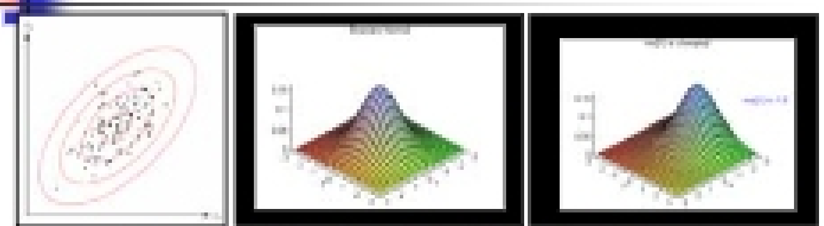
$N(\mu,\sigma^2)$ is a probability distribution function defined by:

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



UCF Computer Vision Lab. 9

Bi-variate Normal Distribution

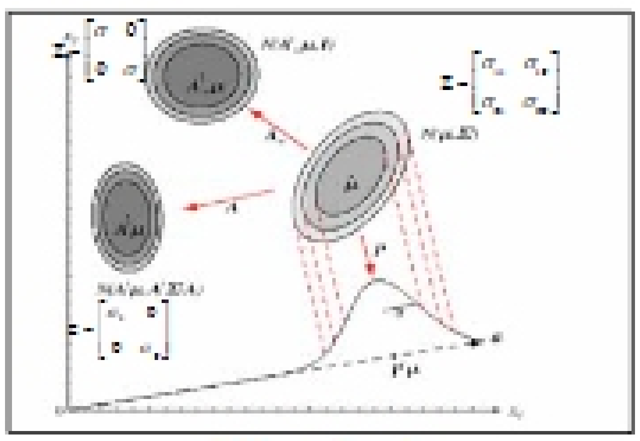


- If we were interested in *r-g*, or *g-b*, or *r-b*...
- The mean can be updated over time simply as

$$\mu_t = (1-\rho)\mu_{t-1} + \rho X_t$$
$$\sigma_t^2 = (1-\rho)\sigma_{t-1}^2 + \rho(X_t-\mu_t)^T(X_t-\mu_t)$$

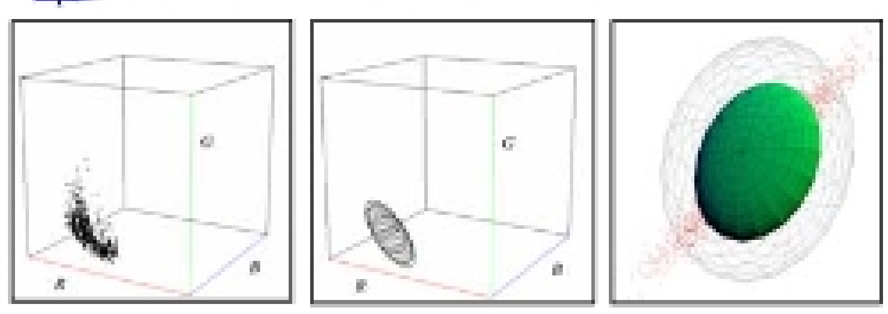
UCF Computer Vision Lab. 10

Covariance



UCF Computer Vision Lab. 11

Tri-variate Normal



UCF Computer Vision Lab. 12

Method I: Pfinder

- Pfinder (Person Finder) - Wren *et al* of MIT (1997)
- C. Wren, A. Azarbayejani, T. Darrel, and A. Pentland, "*Pfinder: Real time Tracking of the Human Body*," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997.
- Color of each pixel modeled as a three-dimensional Gaussian.
- Big Advantage: adaptivity, pixel-wise 'threshold'.

The Bottom Line

- Model each pixel color as a three dimensional normal distribution
- Adapt the color means and variances over time
- Slowly changing illuminations are handled
- Changes to background are eventually learnt
- Relocation and initialization problems are eventually learnt too.

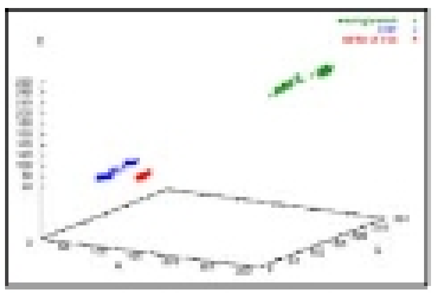
Limitations

- Unfortunately, this method has limitations
- Due to dynamic nature of real-world scenes modeling pixels with single Gaussian distributions is inaccurate
- Quick illumination changes are not handled
- Good for indoor scenes

Method II: Mixture of Gaussians

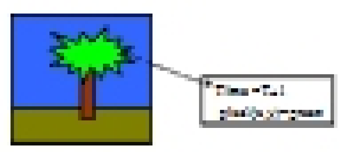
- *MoG* (Mixture of Gaussians) - Stauffer *Grimson* of MIT (2000)
- C. Stauffer, E. Grimson, "*Learning Patterns of Activity using Real-time Tracking*," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000
- Color of each pixel modeled as a *mixture* of three-dimensional Gaussian
- Big Advantage: Handles Multimodality

Multimodality



A Background Subtraction Method by Stauffer and Grimson

In realistic scenarios multiple Processes are generating color '*x*' at each pixel, where $x=[R,G,B]^T$



- A method is required that can incorporate multiple colors in the background model.