SANDIA REPORT
SAND2004-1253
Unlimited Release
Printed March 2004

Automated Video Screening for Unattended Background Monitoring in Dynamic Environments

Jeffrey J. Carlson

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2004-1253
Unlimited Release
Printed March 2004

Automated Video Screening for Unattended Background Monitoring in Dynamic Environments

Jeffrey J. Carlson
Intelligent Systems and Robotics Center

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1003

Abstract

This report addresses the development of automated video-screening technology to assist security forces in protecting our homeland against terrorist threats. A threat of specific interest to this project is the covert placement and subsequent remote detonation of bombs (e.g., briefcase bombs) inside crowded public facilities. Different from existing video motion detection systems, the video-screening technology described in this report is capable of detecting changes in the static background of an otherwise, dynamic environment -- environments where motion and human activities are persistent. Our goal was to quickly detect changes in the background – even under conditions when the background is visible to the camera less than 5% of the time. Instead of subtracting the background to detect movement or changes in a scene, we subtracted the dynamic scene variations to produce an estimate of the static background. Subsequent comparisons of static background estimates are used to detect changes in the background. Detected changes can be used to alert security forces of the presence and location of potential threats. The results of this research are summarized in two MS Power-point presentations included with this report.

Acknowledgements

The following individuals are recognized for their contributions to this project and this report:

Jason Neely, Department 15211, and

Denise Padilla, Department 15211.

Table of Contents

Introduction

The problem we are addressing is the development of automated video-screening technology to assist security forces in protecting our homeland against terrorist threats. A prevailing threat is the covert placement and subsequent remote detonation of bombs (e.g., briefcase bombs) inside crowded public facilities. These locations are ideal for terrorist attacks because 1) many facilities do not screen people entering or exiting the facility due to high costs and intense manpower requirements, 2) it is relatively easy to place a bomb unnoticed because of all the surrounding activity, 3) there is a high potential for large numbers of casualties, and 4) the idea of someone bombing a crowded public facility strikes fear into the hearts of nearly all Americans.

Background

Although video surveillance systems are increasingly more common in public facilities throughout the U.S., current systems are unable to detect the placement of bombs. The mere presence of surveillance cameras is assumed to provide some degree of deterrence. It is also unlikely that security personnel could detect a bomb or, someone placing a bomb, by observing live video from surveillance cameras. The problems lie in the large number of cameras required to effectively monitor a large area, the limited number of security personnel employed to protect these areas, and the intense diligence required to effectively screen live video from even a single camera. Automated video motion detection and tracking systems currently exist for detecting intrusions into a monitored, or protected, area (e.g., the perimeter surrounding a nuclear facility). One of the basic underlying assumptions used by algorithm designers of these systems is that the background is free of targets, or motion, most of the time. That is, the camera mostly observes a relatively static background. The performance of these systems is poor in extremely dynamic environments where motion and human activity are persistent (e.g., inside a subway station, airport, or bus depot).

Technical Approach

Our approach was to develop an automated video-screening technology that is capable of *quickly* detecting changes in the static background of an otherwise, dynamic environment. Different from existing video-detection systems designed to operate in static environments, the video-screening technology is capable of detecting changes in the static background of a dynamic environment: environments where motion and human activities are *persistent*. Our goal was to quickly detect background changes, even if the background is visible to the camera less than 5% of the time. Our approach employs statistical scene models based on mixture densities. We hypothesized that the static-background component of the mixture has a small variance compared to dynamic components. Our initial experiments show this is true about 90% of the time. We have identified extensions to these models that will enable accurate estimation of the static background over 99.9 % of the time. This requirement is based on manpower estimates for response. We have demonstrated robust-threat-detection capabilities using subsequent comparisons of static-background estimates to detect changes and to alert security to the presence and location of potential threats (e.g., the placement of a briefcase bomb next to a trash can). A guard can then make a visual assessment of any potential threat and plan an appropriate response.

Results
The following power-point slides summarize the results of this research.

Automated Video Screening for Unattended Background Monitoring in Dynamic Environments

PI: Jeff Carlson

Team Members: Jason Neely, Denise Padilla

Funding: $100K

Duration: 6 months

Sandia National Laboratories

Objectives

- **Develop automated video screening technology to detect background changes in a dynamic scene environment**
- **Assist security forces in protecting our homeland against terrorist threats**
- **Defeat a prevailing threat: covert placement and subsequent remote detonation of bombs inside crowded public facilities (e.g. airport, subway station, bus depot)**

Sandia National Laboratories

Crowded Public Facility Vulnerabilities

- Many facilities do not screen incoming and outgoing traffic due to high costs and intense manpower requirements

- Surrounding activity makes it relatively easy to place a bomb unnoticed

- There is a high potential for a large number of casualties

- Ideal location for a terrorist attack because the idea of the bombing of a crowded public facility strikes fear into the hearts of nearly all Americans

Sandia National Laboratories

Limitations of Existing Video Surveillance Systems

- Current systems are unable to detect the placement of bombs

- Security personnel are unlikely to detect a bomb or the placement of a bomb by observing live video from surveillance cameras because:
  - A large number of cameras are required to effectively monitor a large area
  - There are a limited number of security personnel employed to protect these areas
  - Intense diligence is required to effectively screen live video from even a single camera

- Existing video motion detection and tracking systems used for detecting intrusions into a monitored area assume the background is free of targets, or motion, most of the time

Sandia National Laboratories

Approach

- Unlike existing video motion detection systems that operate in primarily static environments, the proposed technology will be capable of detecting changes in the static background of a dynamic environment (environments where motion and human activities are commonplace)
- Instead of subtracting the background to detect movement or changes in a scene, we subtract the dynamic scene variations to produce an estimate of the static background
- Subsequent comparisons of static background estimates are used to detect changes in the background (e.g. , the placement of a briefcase bomb next to a trash can)

Sandia National Laboratories

Approach

- Detected changes will be used to alert security forces of the presence and location of potential threats
- Security forces can then make a visual assessment of any potential threats and plan an appropriate response

Sandia National Laboratories

Work in FY03

- **Feature Discovery – We investigated features that are easily extracted from video data and that are useful for background extraction in dynamic scene environments**

- **Algorithm Development – Algorithms were developed using these features to extract the static background of a dynamic scene**

Feature Discovery

- **The pdf of each pixel can be modeled as a Gaussian mixture:**

$$f(x) = P_T \cdot f(x \mid T) + P_B \cdot f(x \mid B)$$

Where

$$f(x \mid B) \sim N(\mu_B, \sigma_B)$$
$$f(x \mid T) \sim N(\mu_T, \sigma_T)$$

(B = Background, T = Target)

And we assume

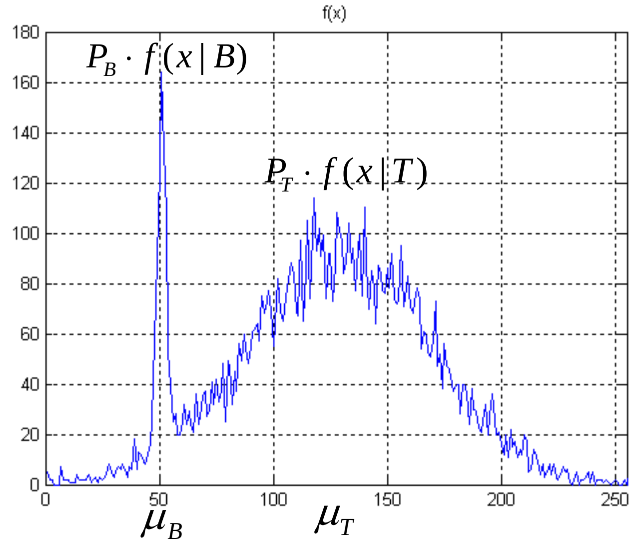
$$\sigma_T >> \sigma_B \quad , \quad P_T >> P_B$$

- **The background estimate is determined from an estimate of** μ_B

- **The statistic we use for estimating the mean background level is the sample mode of the mixture density**

Feature Discovery

- **An example mixture showing the sample mode close to the mean background pixel level** (μ_B)



Feature Discovery

- **Current video motion detection and tracking systems assume the background is free of motion most of the time**

- **This assumption allows pixel averaging to be a good estimate of** μ_B

- **Averaging is not a good estimate of** μ_B **in dynamic scene environments**

Constraint

If we calculate

$$P_T \cdot f(x|T) = \frac{P_T}{\sqrt{2\pi}\sigma_T} \cdot e^{-\frac{(x-\mu_T)^2}{2\sigma_T^2}} \qquad \textbf{At} \qquad \mu_T : \frac{P_T}{\sqrt{2\pi}\sigma_T}$$

And

$$P_B \cdot f(x|B) = \frac{P_B}{\sqrt{2\pi}\sigma_B} \cdot e^{-\frac{(x-\mu_B)^2}{2\sigma_B^2}} \qquad \textbf{At} \qquad \mu_B : \frac{P_B}{\sqrt{2\pi}\sigma_B}$$

A sufficient condition for the sample mode to be a good estimate of μ_B is

$$\frac{P_B}{\sigma_B} > \frac{P_T}{\sigma_T}$$

Constraint

If we substitute $P_T = 1 - P_B$

Then

$$P_B > \frac{1}{1 + \frac{\sigma_T}{\sigma_B}}$$

Let $\sigma_T = k \cdot \sigma_B$

$$\boxed{P_B > \frac{1}{1+k}}$$

If k = 10, we must see the background ~10% of the time in order to be able to estimate it using the sample mode

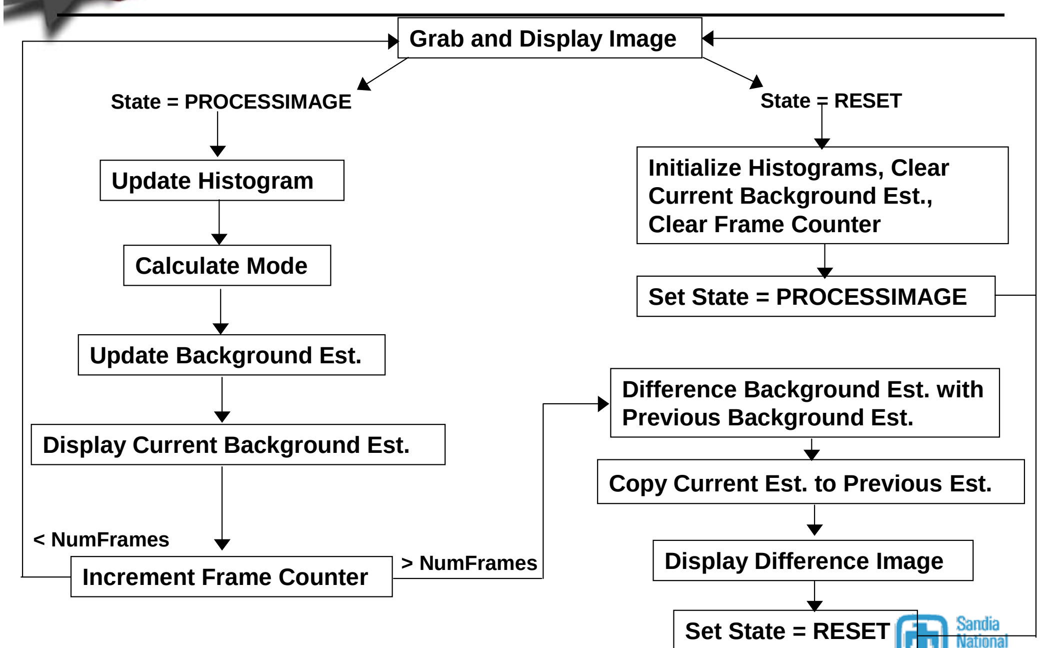
Algorithm

- The algorithm generates and displays 4 images:

  - Current Input Image
  - Current Background Estimate
  - Previous Background Estimate
  - Discrete Difference Image

    (Current Background Est. – Previous Background Est.)

Algorithm



Performance Metrics

- **Probability of Detection (PD) - Given a set number of background pixels that have been changed physically, this is a measure of the percentage of those pixels that are detected as a change in the background image**

- **Probability of False Alarm (PFA) - This is the number of false detections**
  - **Best represented as a rate (avg # false pixel detections/frame OR avg # false detections/day)**
  - **False detections can be declared as either pixel-by-pixel or frame-by-frame**

Sandia National Laboratories

Performance Metrics

- **Detection Lag (DL) - Measure of time that elapses between a physical change in the background and the time at which the difference image displays this change**

- **Extinguish Lag (EL) - This is a measure of how quickly the system assumes any new background changes into the "old background"; it is the time that elapses between the point of detection and the earliest point at which no pixels are shown in the difference image**

- **Both Detection Lag and Extinguish Lag can be determined analytically**

Sandia National Laboratories

Future Work

- **Investigate statistics other than the sample mode to estimate μ_B**

  Ex: Consecutive duration of a particular grey level

- **Histogram Updating (Time Weighted Averaging)**
  - **Histogram values are calculated based on the weighted sum of old values and new values**

$$Hist_{new} = (1-\alpha)\cdot Hist_F + \alpha \cdot Hist_{\alpha \cdot F} \qquad (0 \le \alpha \le 1)$$

- **Histogram Discrete Filtering**
  - **Used to detect a narrow peak (background) that has less amplitude than the peak attributed to targets (people)**
  - **Allows background estimation even if** $\boxed{P_B < \frac{1}{1+k}}$

Video Background Extraction

Department 15212:

Jeff Carlson

Denise Padilla

Jason Neely

Video Background Extraction

Introduction

- Statistical Modeling
- Method1
  - Analysis
  - Simulation
- Method2
  - Analysis
  - Simulation
- Conclusions

Sandia National Laboratories

Statistical Modeling

- By recognizing that a video pixel represents either motion or background, we can model it's probability density function using a mixture:

$$f_x(x) = P_B f_x(x|B) + P_M f_x(x|M)$$

- By quantizing x into discrete values (e.g., k = 0, 1, …, 255), we can use histograms to estimate the density function

$$f_k(k) = P_B f_k(k|B) + P_M f_k(k|M)$$

- The task of background extraction is reduced to the following problem:

Determine: $\hat{\mu}_B = E(f_k(k \mid B))$ from: $f_k(k)$

Sandia National Laboratories

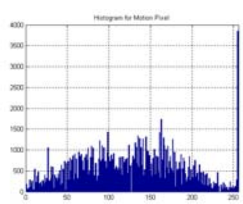
Statistical Modeling

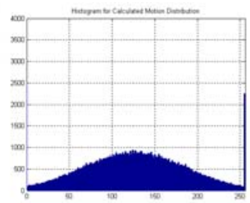
<u>Empirical Results</u>

For a given pixel:

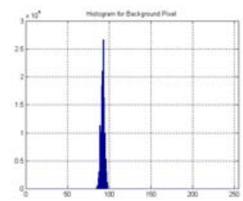
Measured | Generated

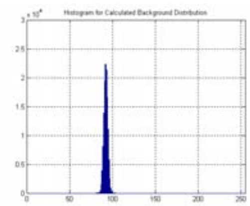
"Motion" characterized by *Gaussian* distribution with *large sigma* and *centered mean*





"Background" characterized by *Gaussian* distribution with *small sigma*





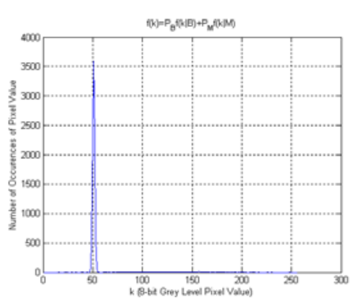
Key Observation: $\sigma_M >> \sigma_B$

Statistical Modeling

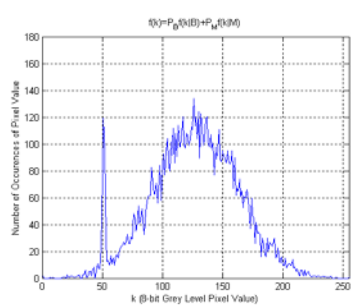
$$f_x(x) = \frac{P_B}{\sqrt{2\pi}\sigma_B} e^{\frac{-(x-\mu_B)^2}{2\sigma_B^2}} + \frac{1-P_B}{\sqrt{2\pi}\sigma_M} e^{\frac{-(x-\mu_M)^2}{2\sigma_M^2}}$$

2 examples of mixture densities generated in MATLAB with $\sigma_M >> \sigma_B$

$P_B >> P_M$



$P_M >> P_B$



Method 1-Analysis

- Method 1 involves using *'sample mode'* to extract background

- Sample mode = $mode(f_k(k))$, $f_k(k)$ = histogram estimate of $f_x(x)$

- We expect: $mode(f_k(k)) \approx \mu_B$ or $mode(f_k(k)) \approx \mu_M$

- condition for *'mode'* to identify background:

$$f_x(\mu_B) > f_x(\mu_M)$$

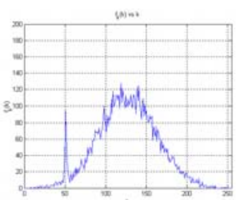
$$\downarrow$$

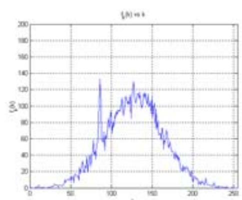
$$P_B > \frac{1}{1 + \frac{\sigma_M}{\sigma_B} K} \quad \text{where} \quad K = \frac{1 - e^{\frac{-(\mu_M - \mu_B)^2}{2\sigma_B^{\,2}}}}{1 - e^{\frac{-(\mu_M - \mu_B)^2}{2\sigma_M^{\,2}}}}$$

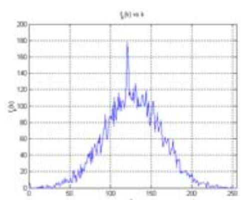
Method 1-Analysis

- The constant K corrects for overlapping probability densities

- As μ_B approaches μ_M, K becomes large $\lim_{\mu_B \to \mu_M} K = \left(\frac{\sigma_M}{\sigma_B}\right)^2$ and required P_B becomes small





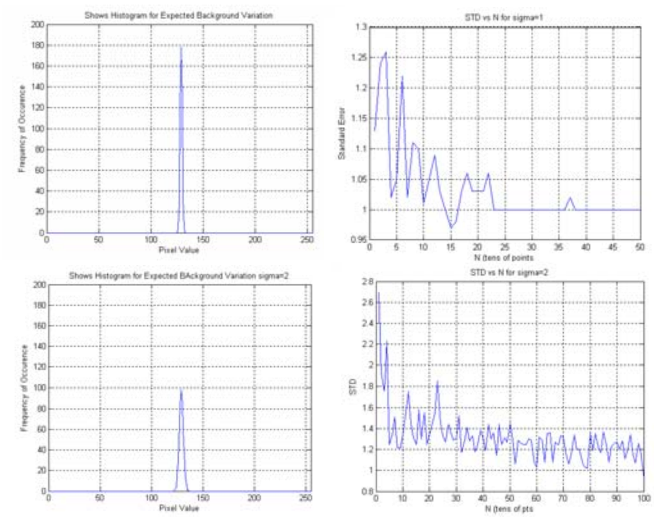


- For worst case, probability density functions are disjoint, K=1

- The sufficient condition is thus: $P_B > \frac{1}{1 + \frac{\sigma_M}{\sigma_B}}$

Method 1-Analysis

How many samples do we need to use the '*sample mode' to estimate μ_B ?*

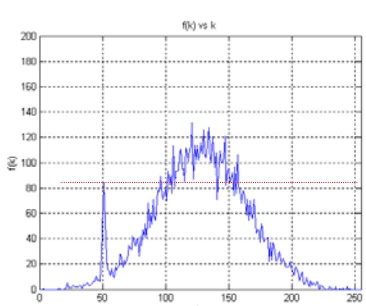


Sandia National Laboratories

Method 1-Analysis

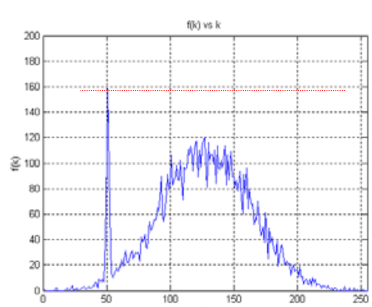
$$P_B < \frac{1}{1 + \frac{\sigma_M}{\sigma_B} K}$$

$$\text{mode}(f_k(k)) \approx \mu_M$$



$$P_B > \frac{1}{1 + \frac{\sigma_M}{\sigma_B} K}$$

$$\text{mode}(f_k(k)) \approx \mu_B$$



Sandia National Laboratories

Method 1-Simulation

• Statistical simulation may be used to check the results of analysis

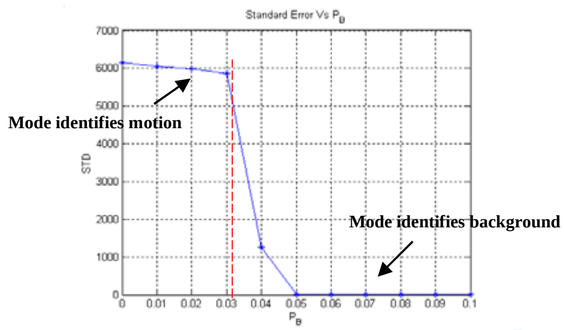
Example System:

μ_B=50, σ_B=1, μ_M=127, σ_M=30, 100 pixels, N=10,000 pts

Simulation varies P_B and determines STD, given *error* = *mode*($f_k(k)$) - μ_B for each

$$P_B > \frac{1}{1+\frac{\sigma_M}{\sigma_B}K}$$

$$P_B > 0.0311$$



Method 1-Analysis

Effect of *N* on the Standard Error *(STD)*:

• The binomial PDF is used to determine the dependency of *STD* on *N*

• *STD*, The standard error is the standard deviation of the error: *error* = *sample mode*-μ_B

$$P(f_k(\mu_B)=h_B)=\binom{N}{h_B}(P_B f_k(\mu_B \mid B))^{h_B}(P_M f_k(\mu_B \mid M))^{N-h_B}$$

$$P(f_k(\mu_M)=h_M)=\binom{N}{h_M}(P_B f_k(\mu_M \mid B))^{h_M}(P_M f_k(\mu_M \mid M))^{N-h_M}$$

$$P(f_k(\mu_B)>f_k(\mu_M))=P(f_k(\mu_B)=f_k(\mu_M)+\alpha) \text{ for all } \alpha \geq 1$$

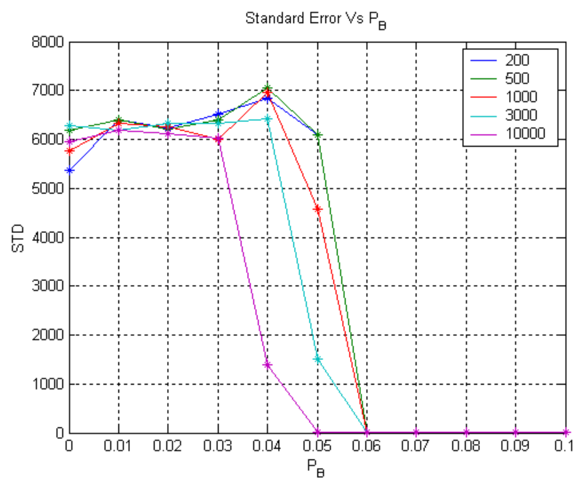
$$P(f_k(\mu_B)>f_k(\mu_M))=\sum_{h_m=1}^{N}\sum_{\alpha=1}^{N-h_m}P(f_k(\mu_B)=h_M+\alpha)P(f_k(\mu_M)=h_M)$$

Method 1-Simulation

Example System: **for variable *N***

μ_B=50, σ_B=1, μ_M=127, σ_M=30, 100 pixels

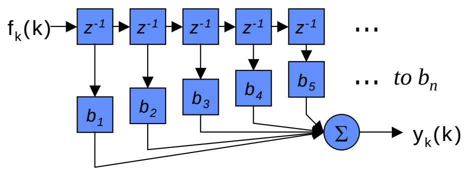
N is varied between 200 and 10,000 points

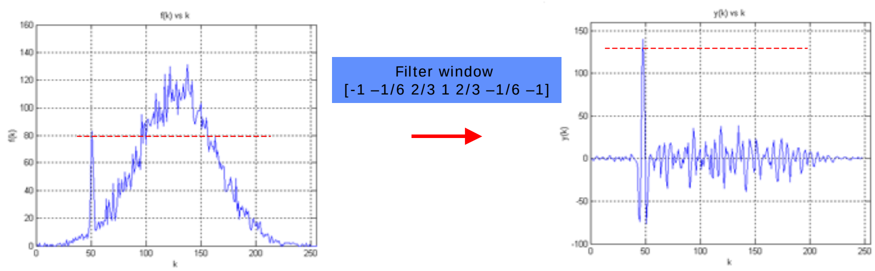


Method 2-Analysis

- Method 2 involves applying a discrete filter to the histogram before using *'mode'*
- Given: $f_k(k) = P_B\, f_k(k \mid B) + P_M\, f_k(k \mid M)$
- A discrete filter is made by allowing $g_k(k)$ to be the windowed version of $f_k(k|B)$ on an interval centered at μ_B
- Thus, $y_k(k) = g_k(k) \otimes f_k(k)$
- We expect $mode(y_k(k)) = \mu_B$

Method 2-Simulation



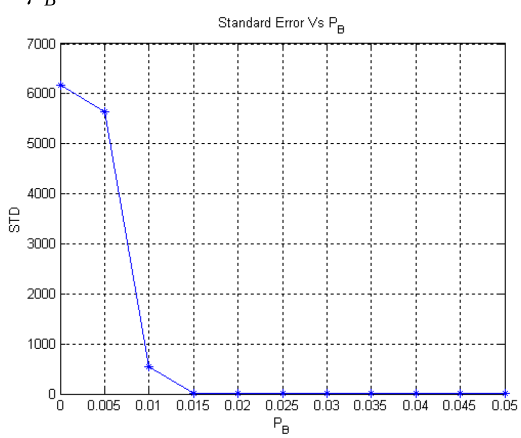


Method 2-Simulation

Example System:

μ_B=50, σ_B=1, μ_M=127, σ_M=30, 100 pixels, N=10,000 pts

Simulation varies P_B and determines the squared error (STD) between *mode($y_k(k)$)* and μ_B for each

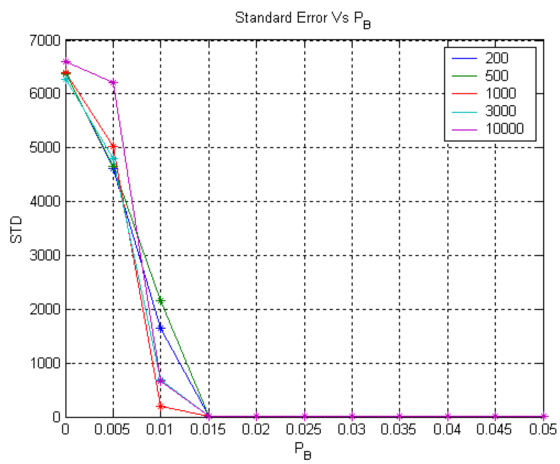


Method 2-Simulation

Example System: **for variable *N***

μ_B=50, σ_B=1, μ_M=127, σ_M=30, 100 pixels

N is varied between 200 and 10,000 points

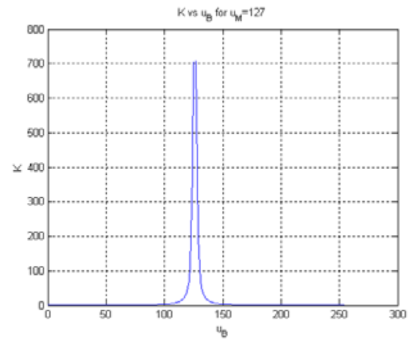


Conclusions

- Pixel values may be modeled as a mixture of $P_B f_x(x|B)$ and $P_M f_x(x|M)$ probability density functions
- Empirical evidence suggests $\sigma_M >> \sigma_B$
- Given a sample of pixel values $f_k(k)$, operations on $f_k(k)$ may be done to determine an estimate of μ_B
- Analysis and simulation suggest that the background may be extracted successfully even when $P_B << P_M$

Ancillary

Plot of K vs μ_B



References

1. **Van Trees**, H., Detection, Estimation, and Modulation Theory Part I (J. Wiley and Sons, New York, 1971).

2. Athanasios **Papoullis**, “Probability, Random Variables, and Stochastic Processes”, Third Edition, McGraw Hill.

Distribution

1	MS 0323	D. L. Chavez 01011
1	MS 1002	S. C. Roehrig 15200
1	MS 1003	J. T. Feddema 15211
1	MS 1003	J. J. Carlson 15211
1	MS 1003	J. Neely 15211
1	MS 1003	D. Padilla 15211
6	MS 1003	B. Ewing 15211
1	MS 9018	Central Technical Files, 8945-1
2	MS 0899	Technical Library, 9616