SIGNAL & IMAGE PROCESSING LAB

LOCAL-GLOBAL BACKGROUND MODELING FOR ANOMALY DETECTION IN HYPERSPECTRAL IMAGES

Eyal Madar

Supervisors:

Prof. David Malah

Dr. Meir Bar-Zohar

7/12/2010

Outline

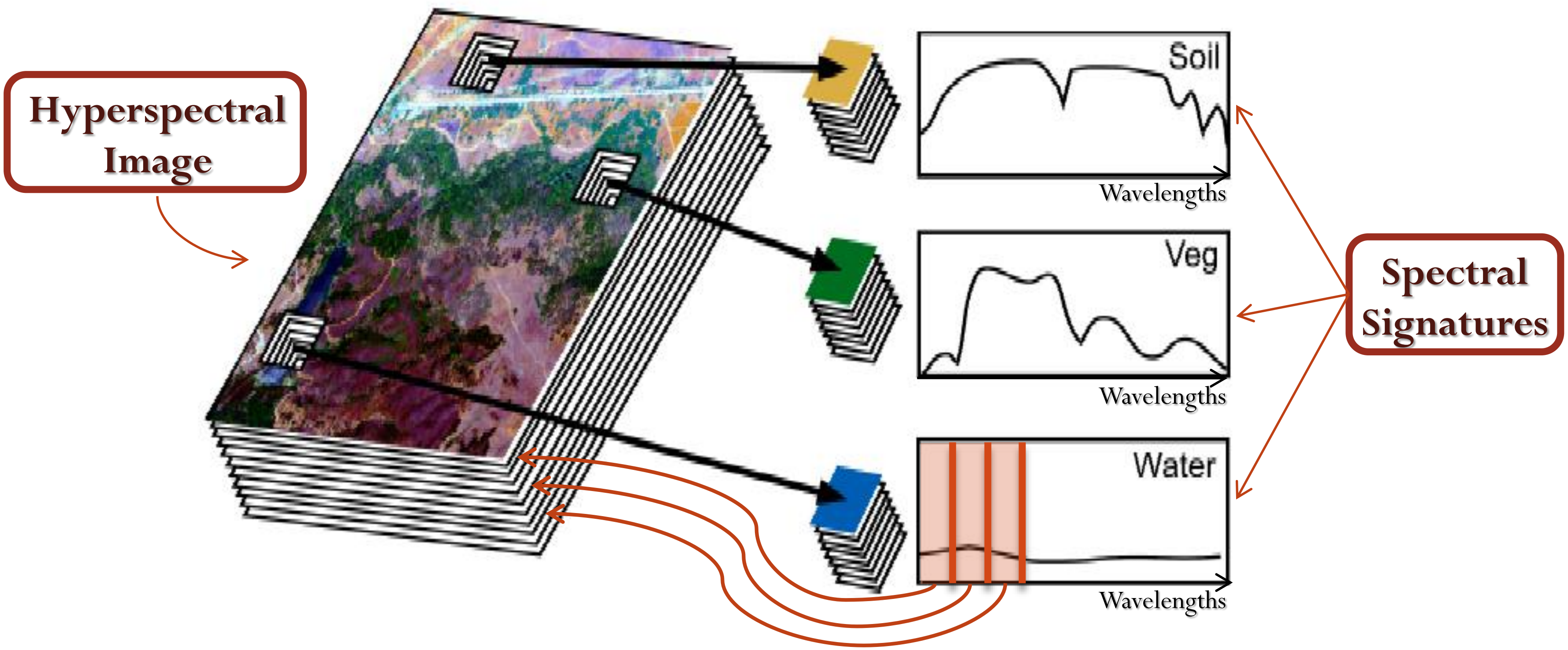
- Hyperspectral Imaging
  - Image Acquisition
  - Anomaly Detection
- Statistical Background Modeling
  - Local Approach
  - Global Approach
- Combined Local/Global Approach – Proposed Algorithm
  - Local Part
  - Global Part
- Improvements of the Proposed Algorithm
  - Spectral Clustering
  - Non-Gaussian Fitting
- Summary and Future Work

Outline

- **Hyperspectral Imaging**
  - **Image Acquisition**
  - **Anomaly Detection**
- Statistical Background Modeling
  - Local Approach
  - Global Approach
- Combined Local/Global Approach – Proposed Algorithm
  - Local Part
  - Global Part
- Improvements of the Proposed Algorithm
  - Spectral Clustering
  - Non-Gaussian Fitting
- Summary and Future Work

Hyperspectral Imaging

Spectral Signature

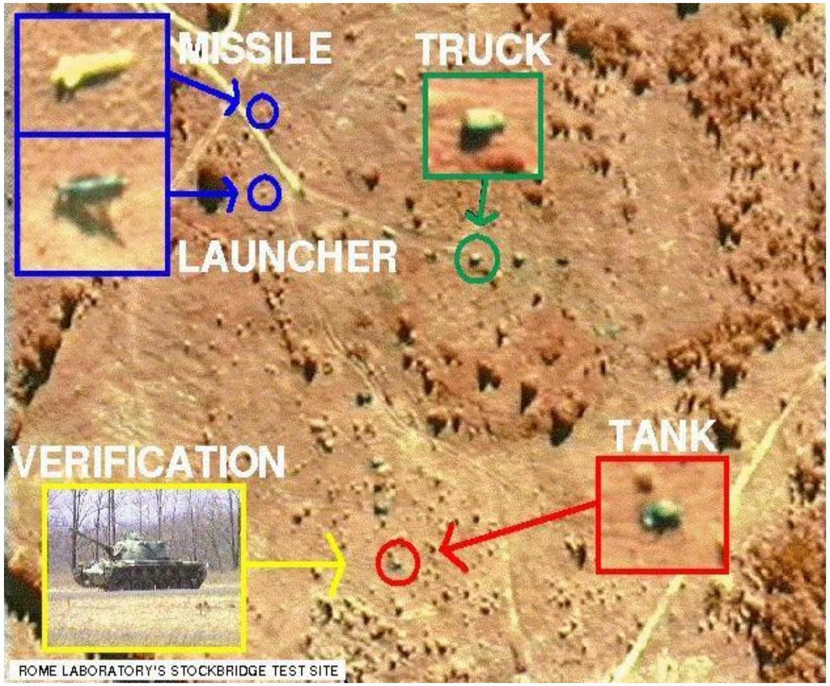


Hyperspectral Imaging
Image Acquisition

- Obtain a continuous spectrum of electromagnetic radiation reflected from the surface of the earth
- Each pixel corresponds to a spectral signature (spectrum) reflected from the pixel location
- Each material can be (uniquely) characterized by its spectrum

Detection in Hyperspectral Imaging

- **Military Interest:**
  - Detect camouflaged man-made objects.
  - Using the spectral signature to discriminate targets from the background



- **Automatic Detection:**
  - *Supervised Algorithms*

    Based on prior knowledge on the target spectral signature
  - *Unsupervised Algorithms*

    No knowledge on the target spectral signature.

Anomaly Detection

- **Unsupervised Detection**
  - Prior anomaly signatures are unknown

- **Anomaly detection methods**:
  1. Model the background
  2. Detect anomalies by finding pixels that are not well-described by the background model

- **Statistical Background Modeling**
  - Local approach
  - Global approach

Outline

- Hyperspectral Imaging
  - Image Acquisition
  - Anomaly Detection
- **Statistical Background Modeling**
  - **Local Approach**
  - **Global Approach**
- Combined Local/Global Approach – Proposed Algorithm
  - Local Part
  - Global Part
- Improvements of the Proposed Algorithm
  - Spectral Clustering
  - Non-Gaussian Fitting
- Summary and Future Work

Statistical Background Modeling

- **Local Approach:**
  - Background is estimated in a local neighborhood of a tested pixel.
  - An anomaly is a pixel spectrally different from the local background

- **Global Approach:**
  - Background modeling is based on the entire image.
  - An anomaly is a pixel spectrally different from the global background.

Local Approach

RX (Reed-Xiaoli 1990) - 1

- **Assumption**

The background pixels in a local neighborhood of a tested pixel are assumed to be independent, identically distributed, Gaussian random vectors.

$$H_0 = N(\mu, \Sigma) \qquad \text{(Anomaly absent)}$$

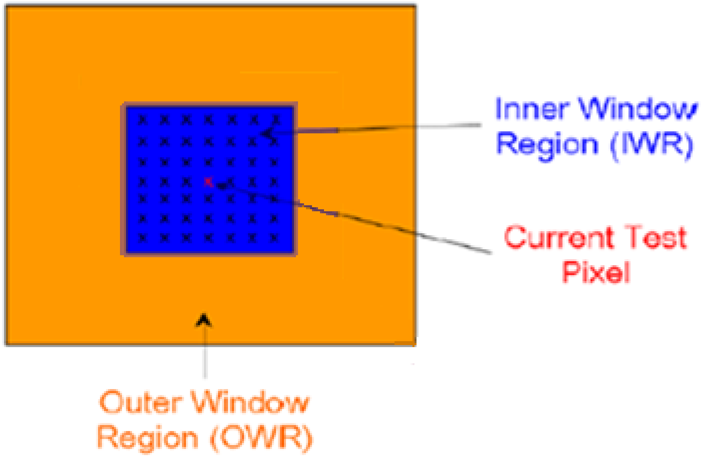
$$H_1 = N(\mu + a\mathbf{s}, \Sigma) \qquad \text{(Anomaly present)}$$

- **Maximum Likelihood Gaussian Statistics Estimation**

$$\mu = \frac{1}{N_b}\sum_{i=1}^{N_b} \mathbf{x}_i$$

$$\Sigma = \frac{1}{N_b - 1}\sum_{i=1}^{N_b} (\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T$$

$N_b =$ Number of pixels in OWR



Local Approach

RX (Reed-Xiaoli 1990) - 2

- **Generalized Likelihood Ratio Threshold (GLRT)**

  The Mahalanobis distance between the tested pixel and the background mean vector is compared to a threshold to detect an anomaly:

$$RX(y) = (\mathbf{y} - \mu)^T \Sigma^{-1} (\mathbf{y} - \mu) > Th$$

Global Approach
GMM-RX

Assumption

The background process is modeled by a linear combination of K Gaussian distributions.

$$f(\mathbf{X}) = \sum_{k=1}^{K} \tau_k f_k(\mathbf{X}|\mu_k, \Sigma_k)$$

$$f_k(\mu_k, \Sigma_k): \quad \text{Gaussian Distribution}$$

$$\tau_k: \quad \text{Probability of cluster k occurence}$$

Anomaly Test

An anomaly is, as in local methods, a pixel that does not fit well to the background process.

$$\text{GMM-RX}(\mathbf{y}) = \sum_{k=1}^{K} \tau_k f_k(\mathbf{y}|\mu_k, \Sigma_k) < P_{th}$$

Local and Global

Summary

Local Approach

- ***Advantage***

  Due to the *many degrees of freedom*, local background models can be tightly fitted to the background data.

- ***Problem***
  - Too high number of degrees of freedom may cause model *overfitting.*
  - Insufficient data for parameters estimations of complex local models

Global Approach

- ***Advantage***

  More resistant to the overfitting problem.

- ***Problem***
  - Limited ability to adapt to all nuances of the background process (*underfitting* problem)
  - Difficult optimization process with a lot of local minima

Outline

- Hyperspectral Imaging
  - Image Acquisition
  - Anomaly Detection
- Statistical Background Modeling
  - Local Approach
  - Global Approach
- **Combined Local/Global Approach**
  - **Local Part**
  - **Global Part**
- Improvements of the Proposed Algorithm
  - Spectral Clustering
  - Non-Gaussian Fitting
- Summary and Future Work

Combined Local-Global Approach

- **Goal**

  Significantly improve detector performance by a proper combination of the local and global background modeling principles.

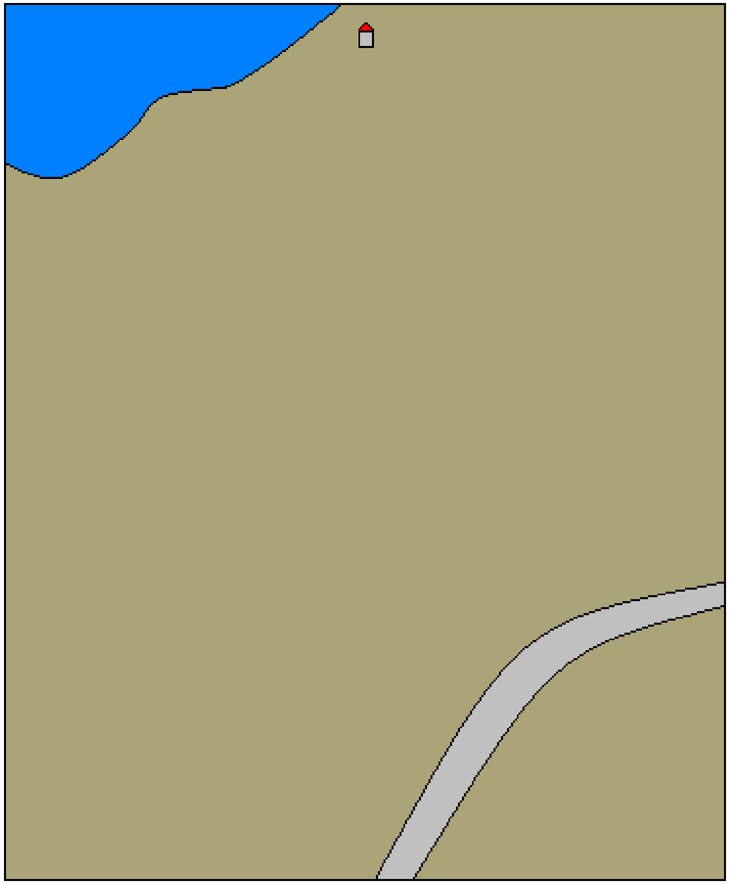
- **Background Extreme Value Analysis - BEVA**

  Local part:
  - Local background model estimation based on Extreme Value Theory

  Global part:
  - Global post-filtering using a "dictionary" of local background model

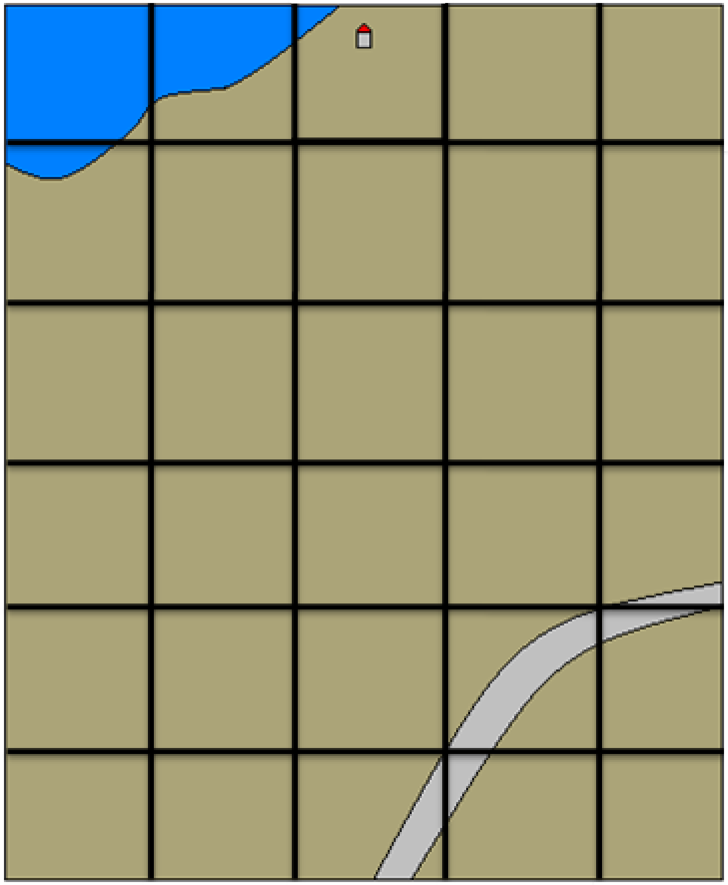
BEVA

Background Extreme Value Analysis



BEVA

Background Extreme Value Analysis

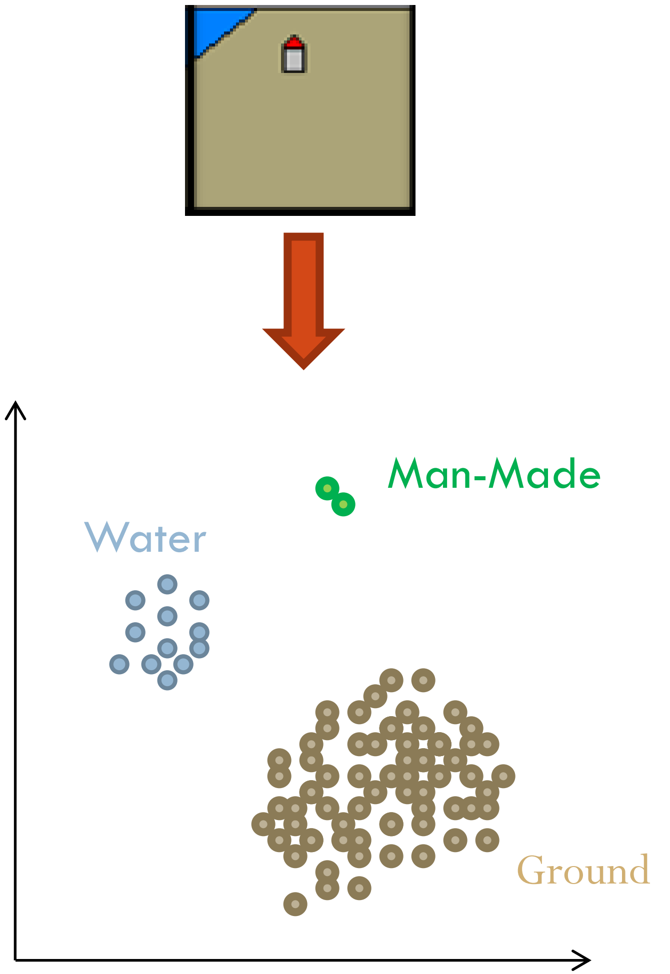


BEVA's Local Part

Background Model Assumption

- **The local background model is :**
  1. *Composed of a small number of distinct clusters up to L*
  2. *Ordered by size*
  3. *Each distributed as a separate Gaussian distribution*

$$\begin{cases} x \in C_k, \quad 1 \leq k \leq L \\ C_k \sim N(\mu_k, \Sigma_k) \\ |C_1| \geq |C_2| \geq \cdots \geq |C_L| \end{cases}$$



Scatter Plot

Background cluster hypothesis test (1)

Goal

An automatic test that isolates pixels belonging to a specific background cluster in a local image area.

Assumption

The cluster statistics (Gaussian mean and covariance) are known

"Background"

Indices of a specific background cluster pixels in a local block

υ - Maximum Mahalanobis distance of B

"Anomaly"

Indices of other background clusters or anomalies pixels in a local block

ξ - Maximum Mahalanobis distance of A

Background cluster hypothesis test (2)

- **Distribution of υ**

  Given by extreme value statistics of maximum-norm Gaussian realizations:

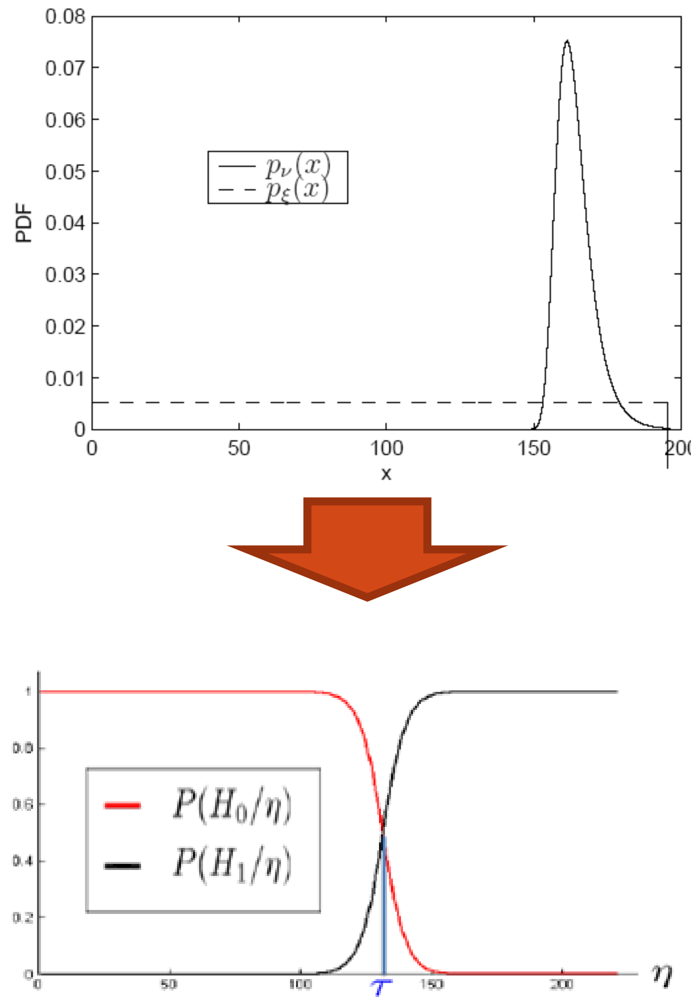
$$P(\upsilon \leq x) = G(a_{p,N}(x - b_{p,N}))$$

  with

$$G(x) = e^{-e^{-x}} \quad \text{(Gumbel distribution)}$$

- **Distribution of ξ**

  Assumed to be uniformly-distributed



BEVA's Local Part

Algorithm

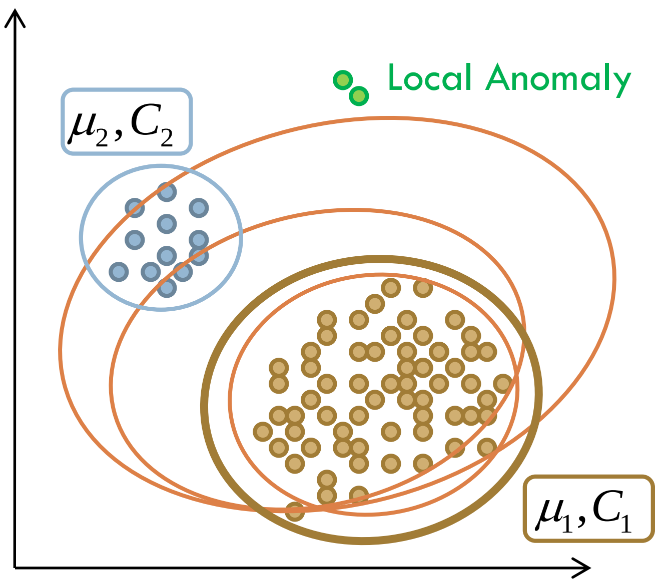
- **<u>Main Loop :</u>**

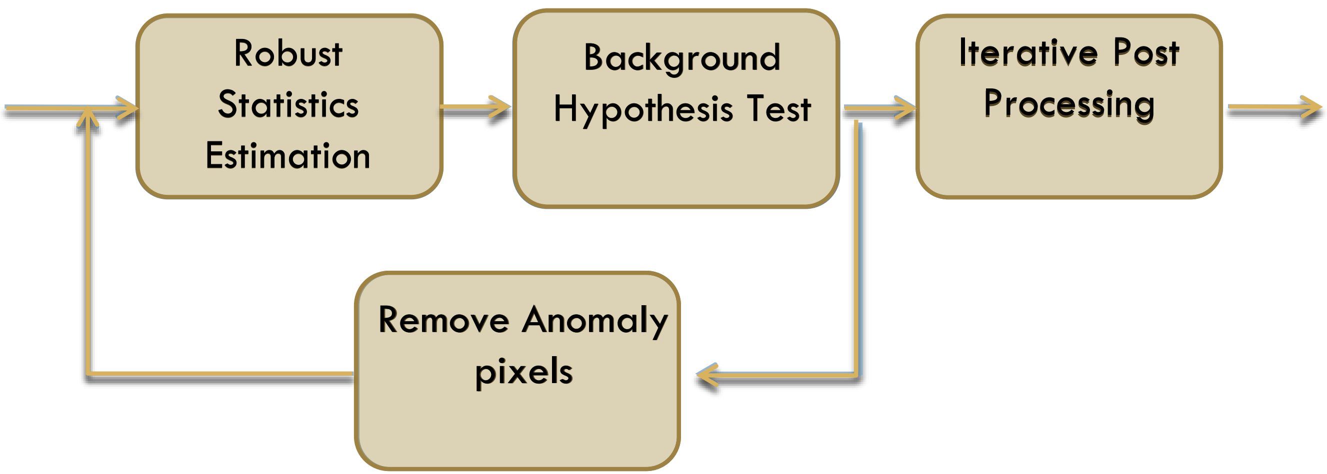
  An intermediate set B_{tmp} exclusively composed of pixels of the dominant background cluster is obtained.

- **<u>Post Processing:</u>**

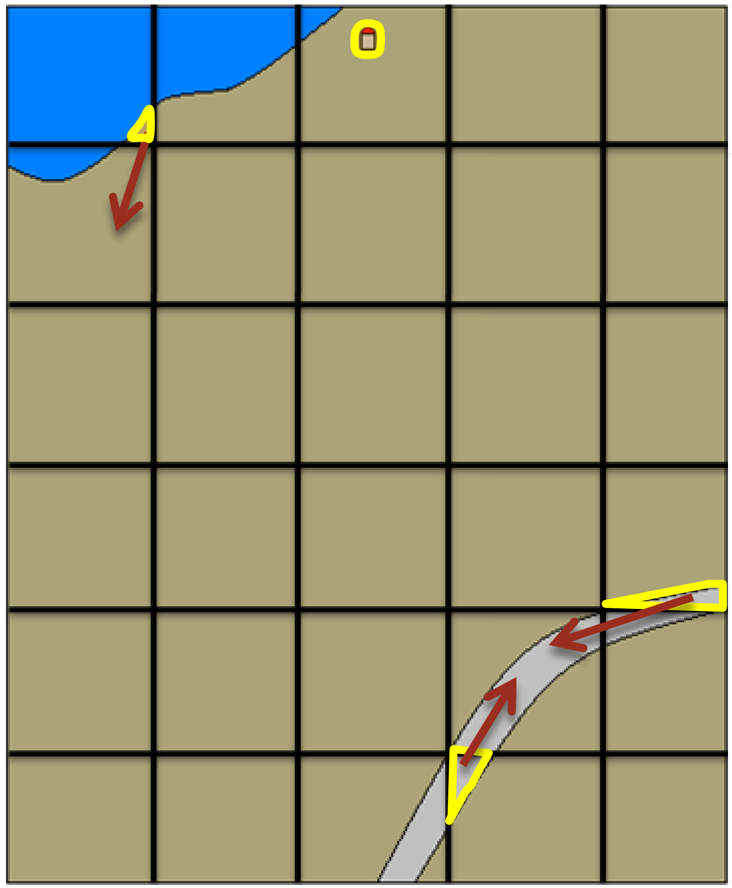
  Re-introduce the excluded pixels back into B_{tmp}

- **<u>Result:</u>**
  - Clusters statistics estimation
  - Pixel classification (local anomalies)





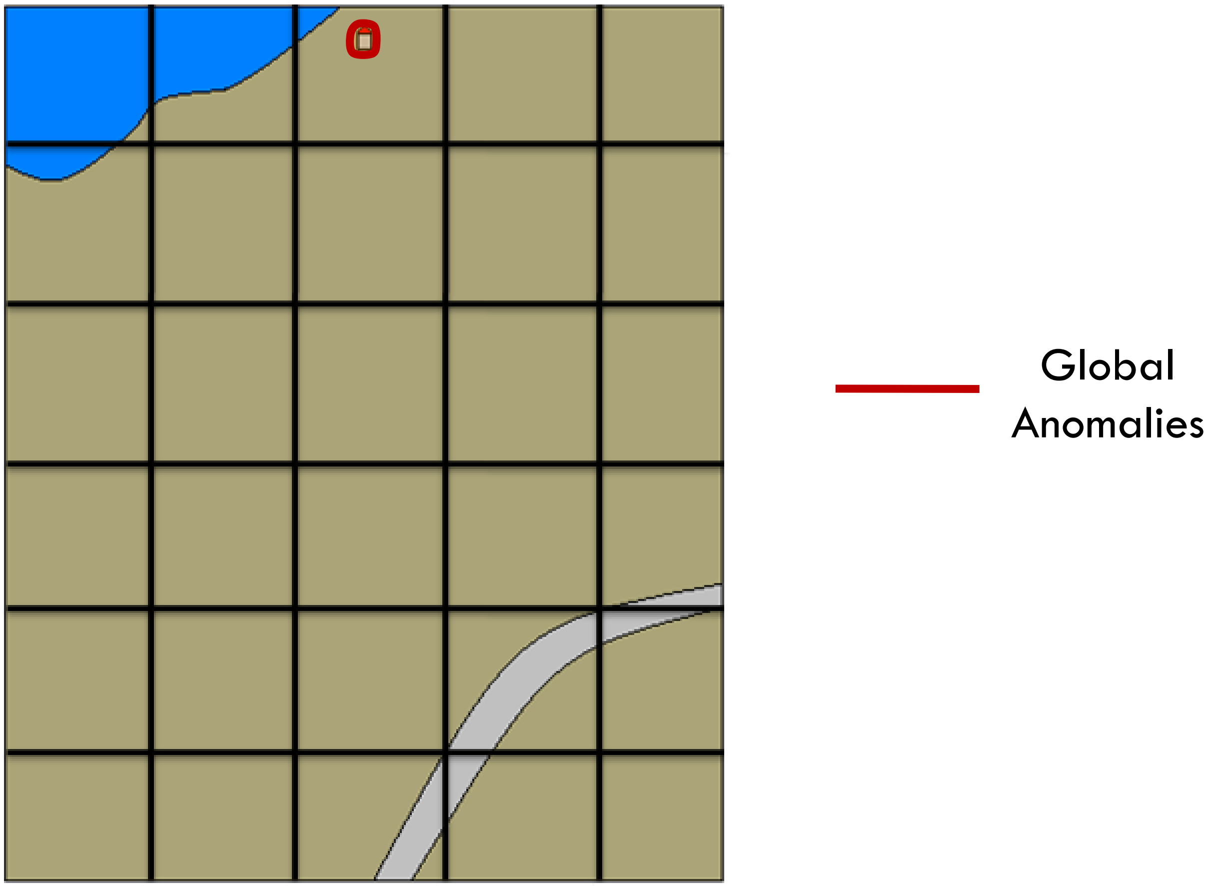
BEVA's Global Part



Anomaly classification to another similar background cluster in a block of a global area

Local Anomaly Detection

BEVA's Global Part



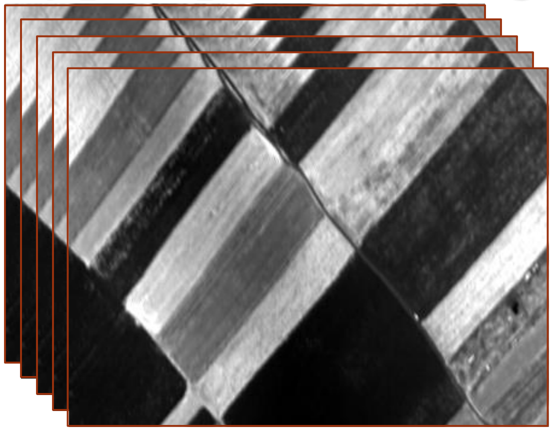
Experimental Results

- **Data:**
  - AISA airborne sensor
  - 5 real hyperspectral image cubes (1.2km^2)
  - 65 spectral bands (400-1000nm)
  - 50 anomalies (vehicles and small constructions)

- **Algorithm:**
  - Local block size 35x35
  - Global block size 525x300

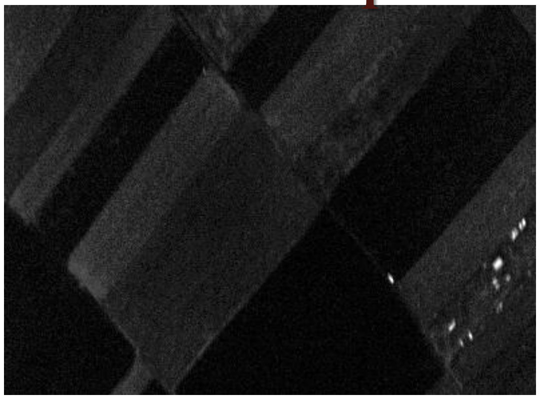
ROC Curve

Hyperspectral Image



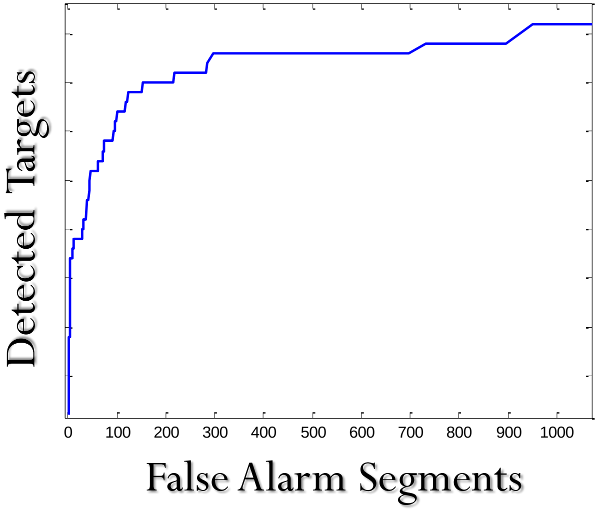
Anomaly Detection

Heat Map



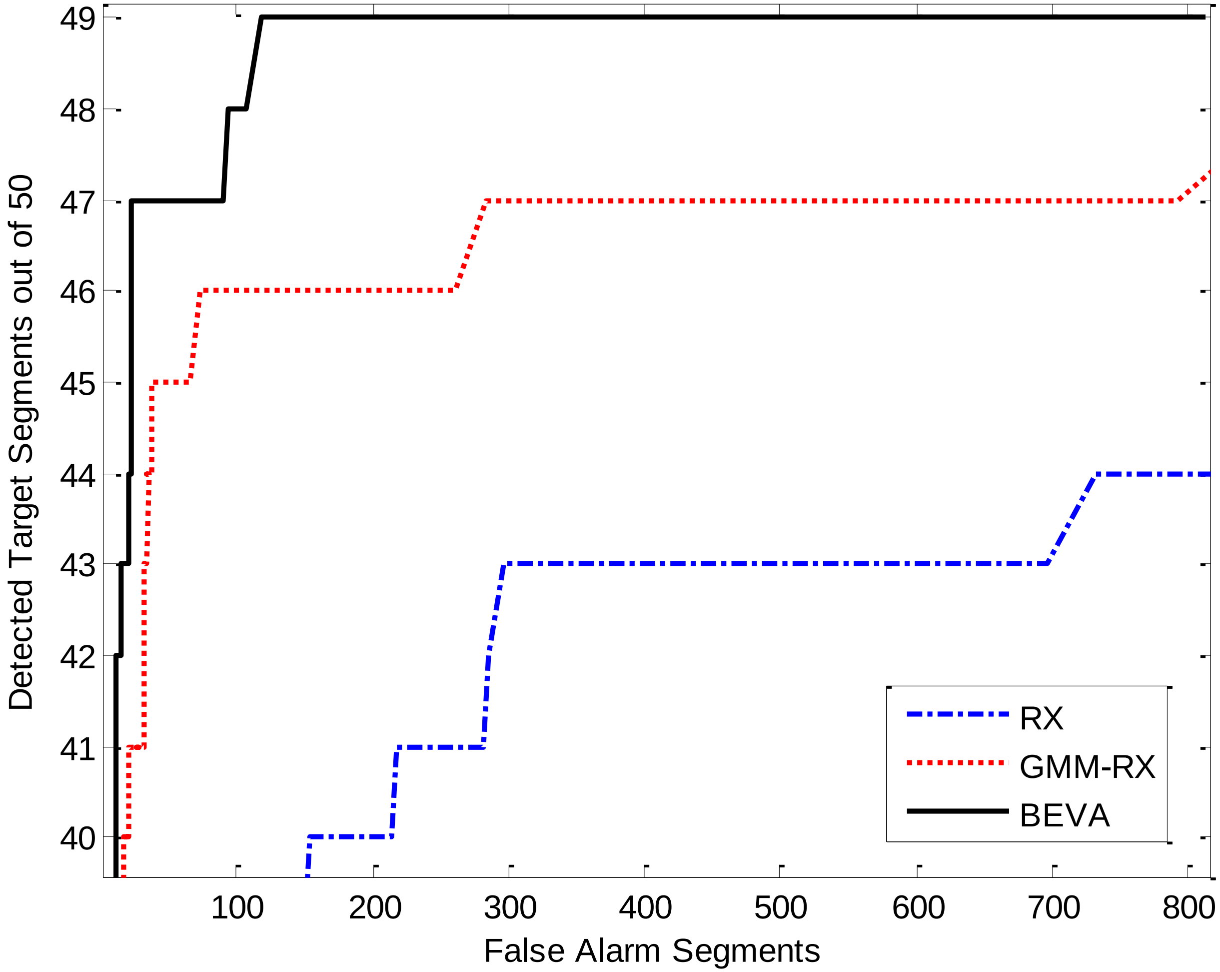
Ground Truth

ROC curve



BEVA

ROC Curve



Outline

- Hyperspectral Imaging
  - Image Acquisition
  - Anomaly Detection
- Statistical Background Modeling
  - Local Approach
  - Global Approach
- Combined Local/Global Approach – Proposed Algorithm
  - Local Part
  - Global Part
- **Improvements of the Proposed Algorithm**
  - **Spectral Clustering**
  - **Non-Gaussian Fitting**
- Summary and Future Work

Spectral BEVA

Motivation

- **BEVA's drawbacks**
  - ***Local Segmentation problem***

    Difficulties in multimodal background pdf estimation. More than 90% local blocks are segmented into just one cluster

  - ***Global Filter problem***

    Spatially dispersed background pixels are wrongly detected in BEVA as anomalies

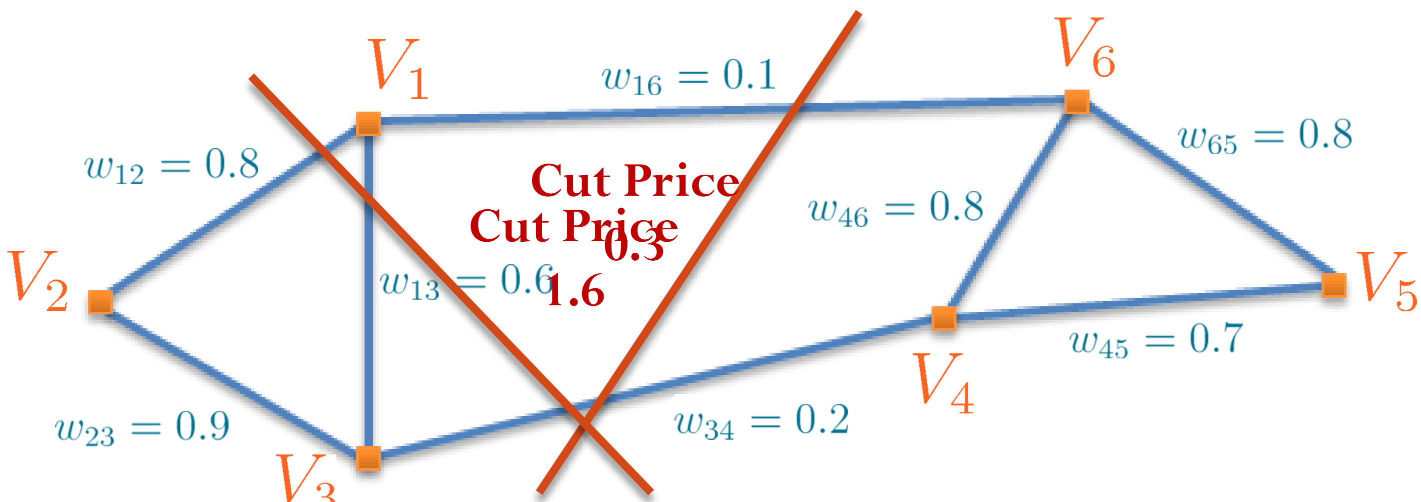
- **Spectral BEVA solution**
  - Segment the local block using Spectral Clustering
  - Estimate background statistics for each local segment
  - Add an auxiliary global background dictionary

Spectral BEVA

Similarity Graph

❑ Graph Theory

- Represent dataset as a weighted graph $G(V, E)$
- Pixels $\{x_i\}_1^N$ as the vertices V
- All pairs of vertices are connected by an edge $E = w_{ij}$
- Large weights mean that the adjacent vertices are very similar; small weights imply dissimilarity.



Spectral BEVA

Normalized Graph Cut

- **Definition**

$$Ncut(A_1, \cdots, A_C) = \sum_{i=1}^{C} \frac{cut(A_i, \bar{A}_i)}{vol(A_i)}$$

$$vol(A) = \sum_{i,j \ in \ A} w_{i,j}$$

$$\mathbf{A}^* = min_{A_1, \cdots, A_C} Ncut$$

- **Why use this criterion?**
  - Segments the vertices of the graph
  - Favors balanced partitions.

Computing an optimal cut is NP-hard

Spectral BEVA

Spectral Clustering Algorithm (1)

1. Define an similarity function between 2 pixels

$$w_{ij} = exp(-\frac{||x_i - x_j||_2^2}{\sigma^2})$$

2. Compute similarity matrix (W), degree diagonal matrix (D) and normalized Laplacian matrix (L)

$$D_{ii} = \sum_{j=1}^{N} w_{ij}$$

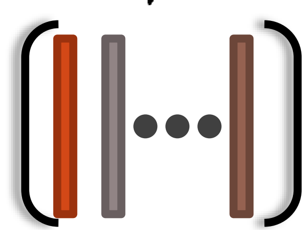
$$L = D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$$

3. Solve: $Lv = \lambda v$

4. Find the C largest eigenvalues of L :

$$V = \{v_1, v_2, \cdots, v_C\} \in \mathbb{R}^{N \times C}$$

V



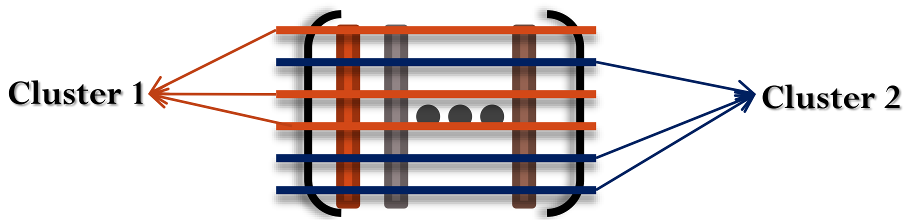
Spectral BEVA

Spectral Clustering Algorithm (2)

5. Re-normalize the rows of V

$$\widetilde{V}_{ij} = V_{ij} / \sqrt{\sum_j V_{ij}^2}$$

6. Treat each row of $\widetilde{V}$ as a point in $\mathbb{R}^C$ and cluster via k-means

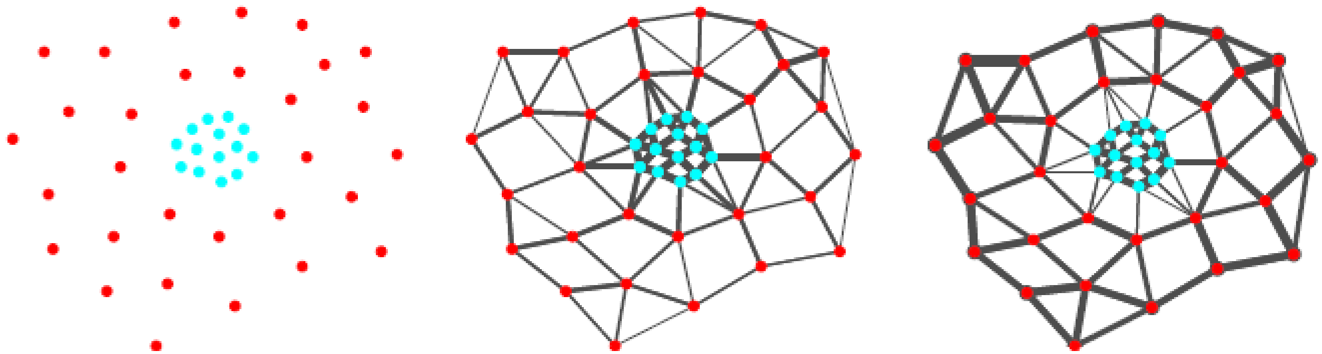


7. Assign the original point x_i to cluster c if and only if the corresponding row i of the matrix $\widetilde{V}$ was assigned to cluster c

Spectral BEVA

Local Scaling (Zelnik-Manor, Perona – 2004)

- **Motivation**



Global Scaling Parameter **Local Scaling**

Calculate a local scaling parameter σ_i for each data point x_i

- **Self Tuning**
  - Using the local statistics of the neighborhood of point x_i
  - $\sigma_i = d(x_i, x_K)$ where x_K is the K'th neighbor of point x_i

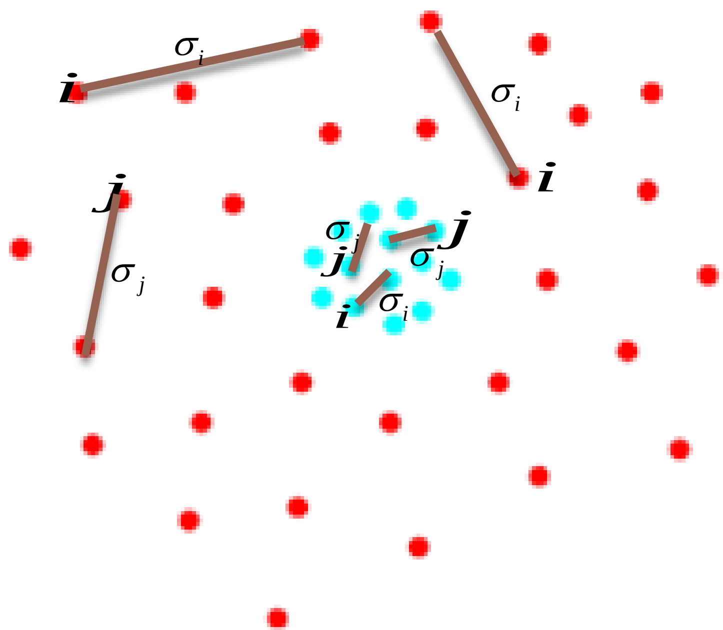
Spectral BEVA

Local Scaling

- **Cases:**
  - Both large scaling parameters, large distance → **High similarity**
  - One small scaling parameter, large distance → **Low similarity**
  - Both small scaling parameters, small distance → **High similarity**

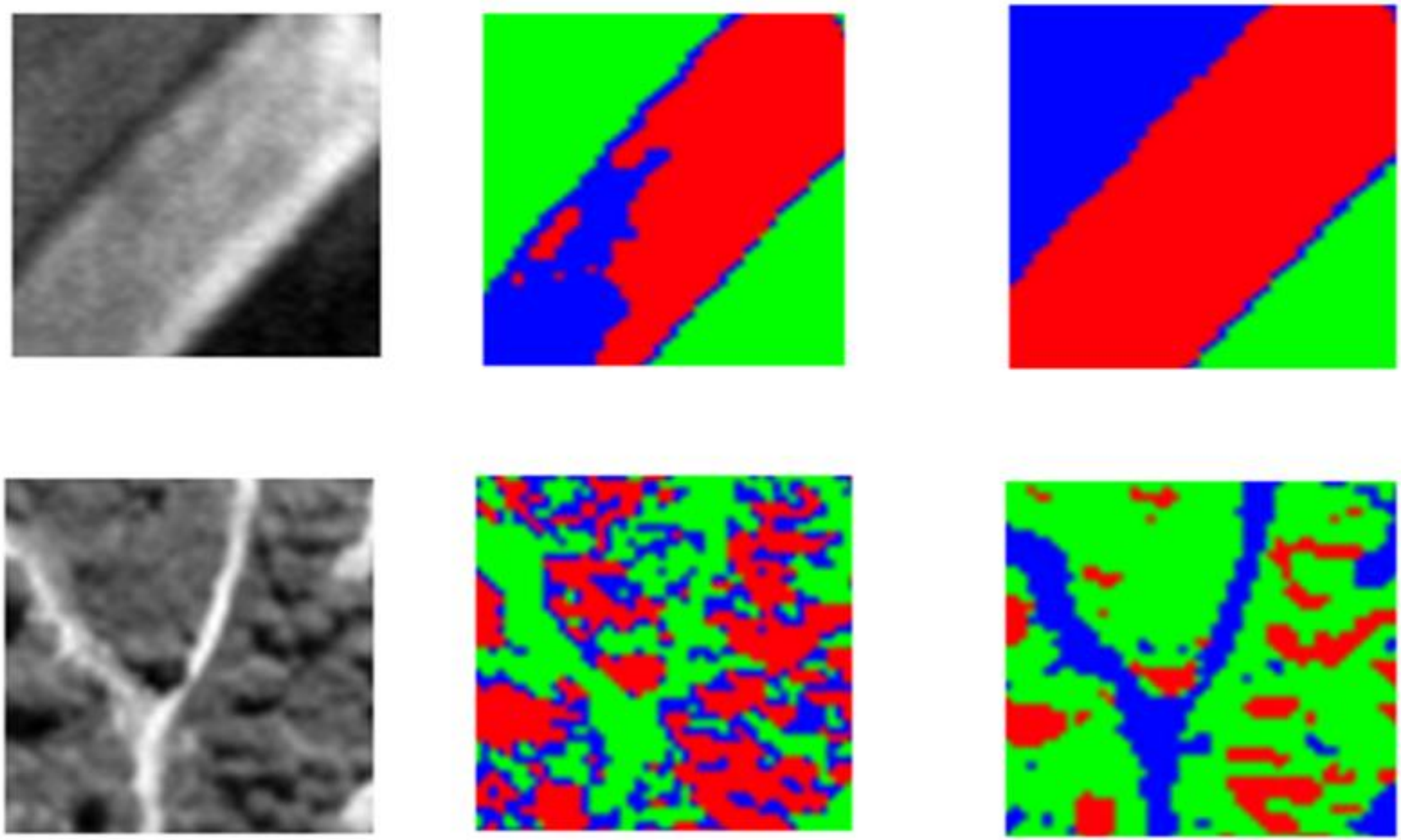
$$w_{ij} = e^{-\frac{||x_i - x_j||_2^2}{\sigma_i \sigma_j}}$$

$$\sigma_i = d(x_i, x_K)$$



Spectral BEVA

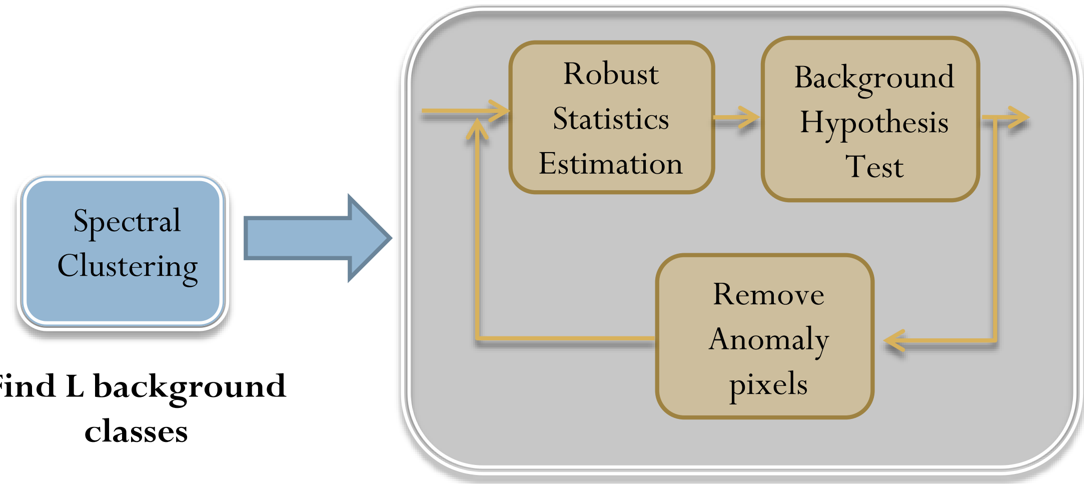
Clustering Results



Global Scaling **Local Scaling**

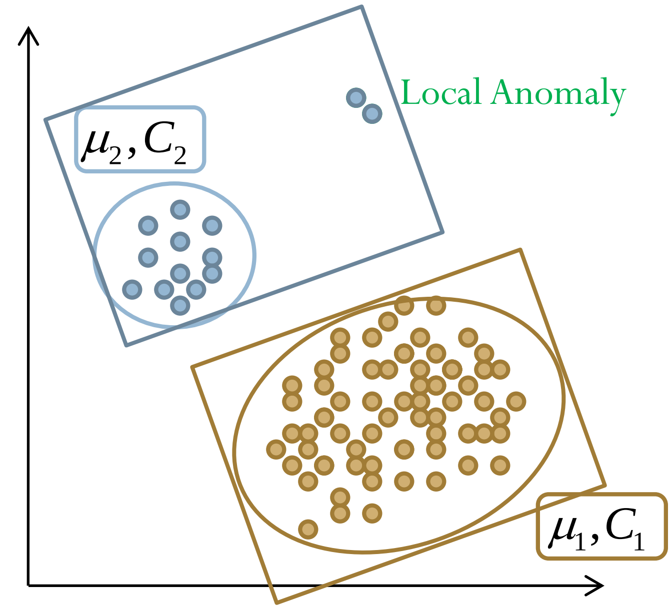
Spectral BEVA

Local Part



Find L background classes

Robust Gaussian Statistics estimation for each background class



Spectral BEVA

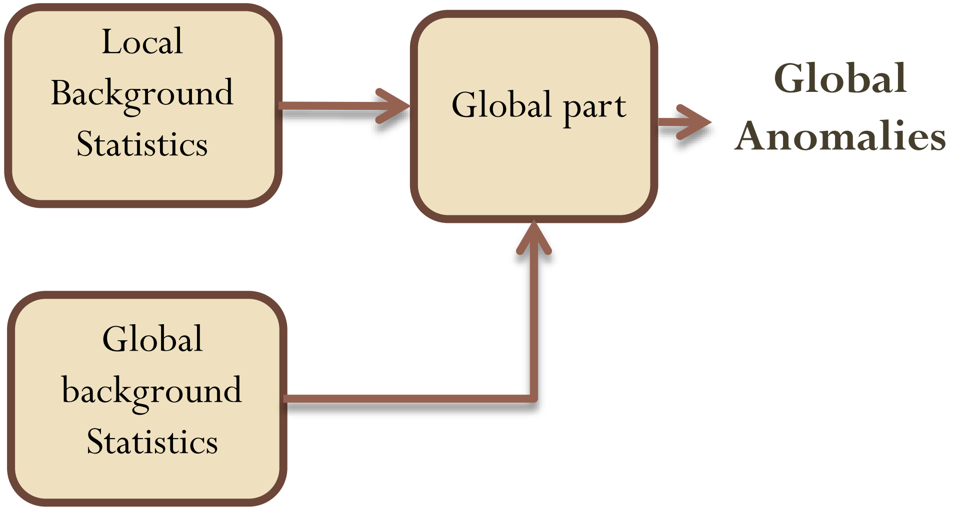
Using an auxiliary dictionary

Local Spectral BEVA

$$\begin{cases} 1 \le k \le L, \quad 1 \le j \le T \\ C_{jk} \sim N(\mu_{jk}, \Gamma_{jk}) \end{cases}$$

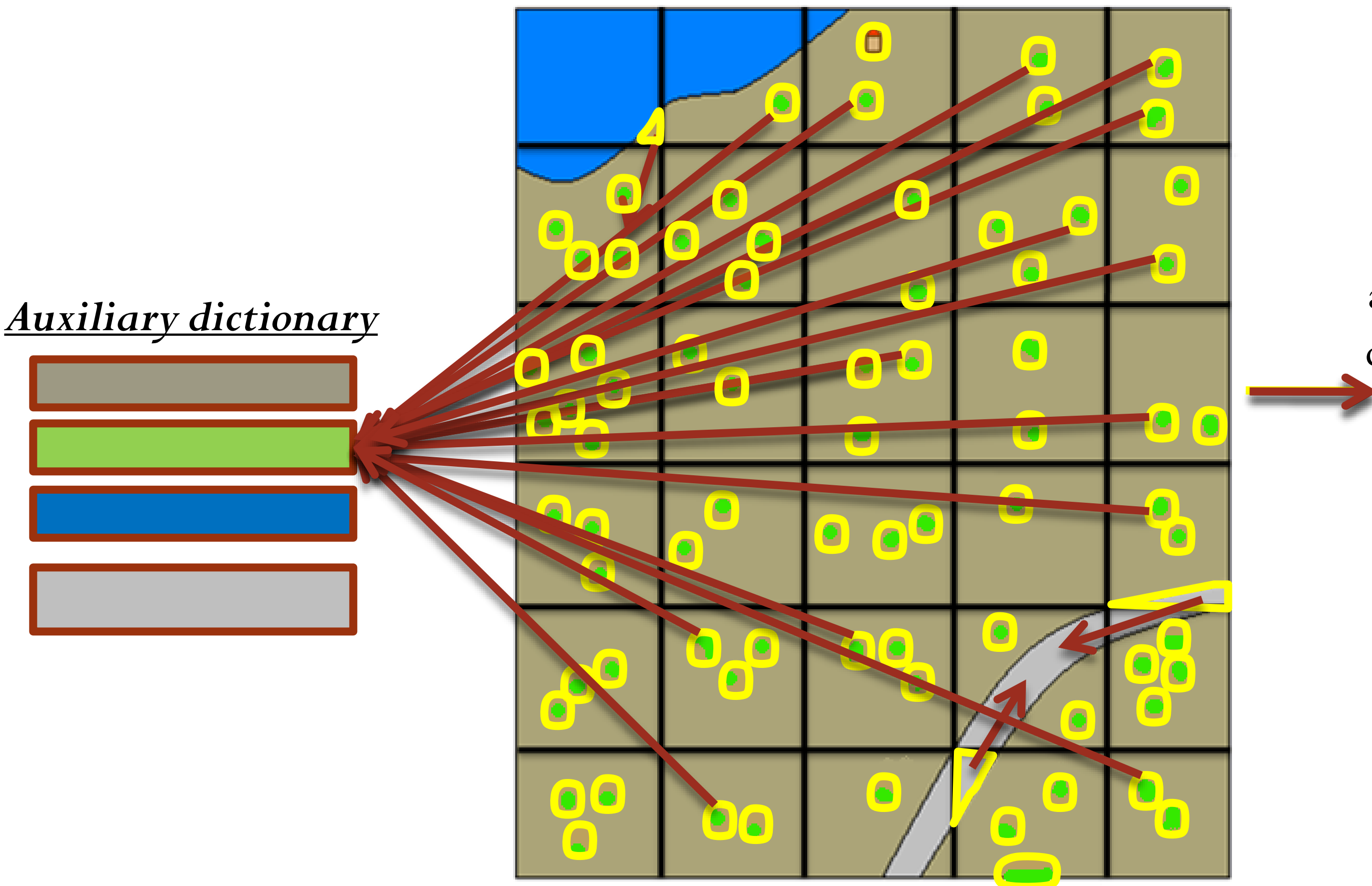
Global GMM:

$$C_j^{gmm} \sim N(\mu_j^{gmm}, \Gamma_j^{gmm}) \quad 1 \le j \le L^{gmm}$$



Spectral BEVA

Global Part

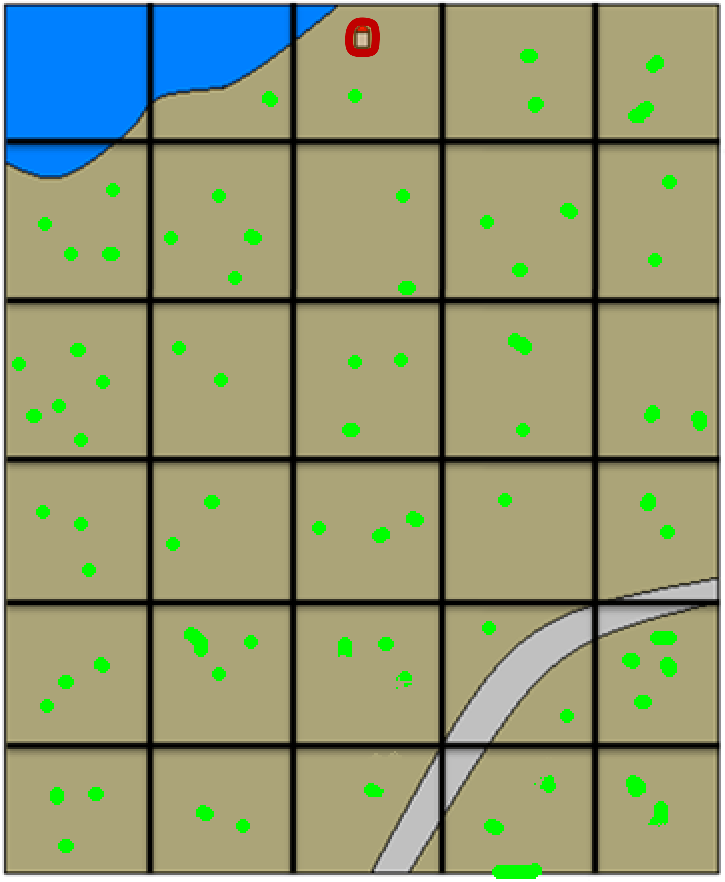


Anomaly classification to another similar background cluster in a block of a global area **or a global cluster of the auxiliary dictionary**

Local Anomaly Detection

Spectral BEVA

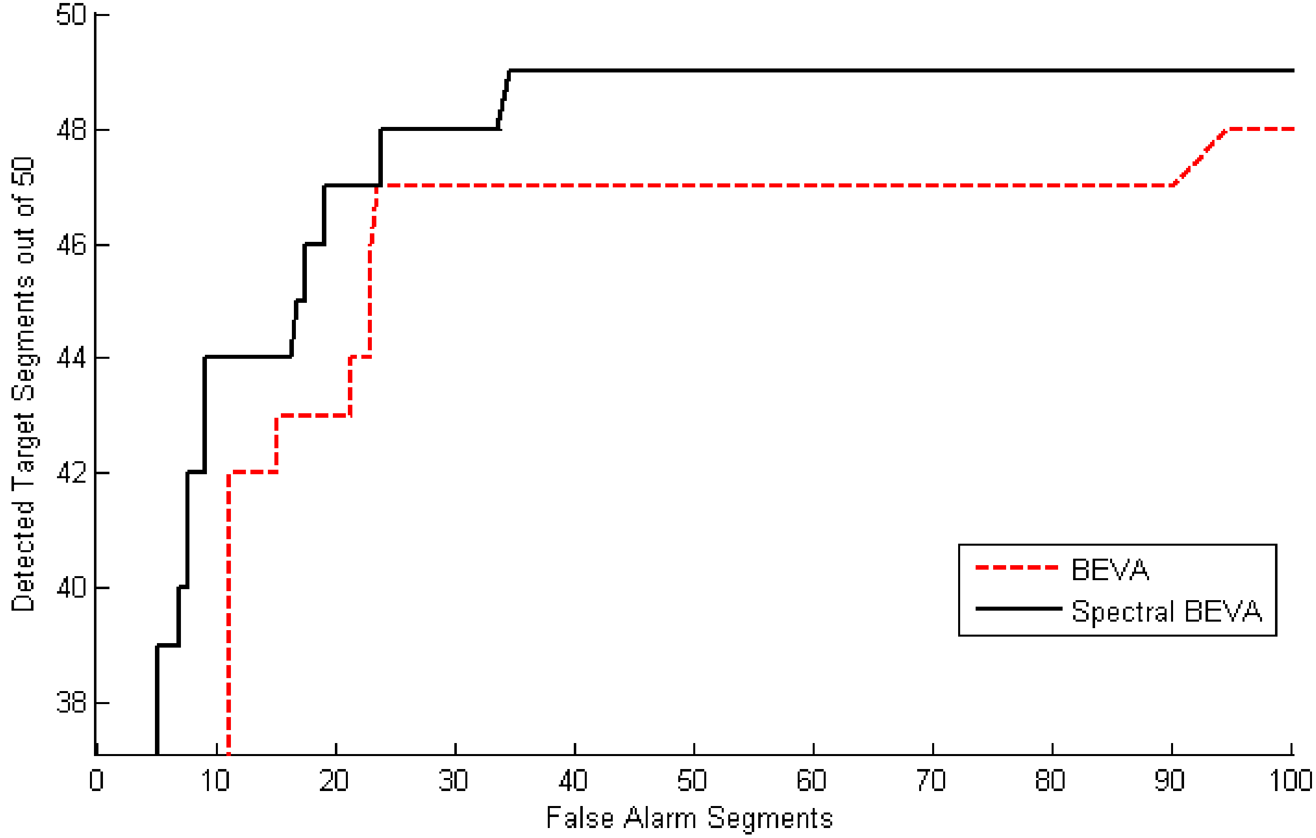
Global Part



Global
Anomalies

Spectral BEVA

ROC



Gaussian Assumption in BEVA

- ***Pros***
  - Efficient processing
  - Mathematically tractability
  - Simplifies the derivation of decision rules

- ***Cons***
  - Not sufficiently adequate to represent the statistical behavior of real hyperspectral background cluster
  - Distributions of hyperspectral data have heavier tails than the Gaussian pdf

➡ **Can lead to an excess number of false alarms**

Non Gaussian BEVA (NG-BEVA)

Local Background Model

N Background Cluster pixels

Mahalanobis Distances

$$d_i = (x_i - \mu)\Sigma^{-1}(x_i - \mu)^T$$

BEVA

$\{d_i\}_1^N \sim \chi^2(p)$ (Chi-squared)

$p =$ number of spectral bands

Background Cluster $\sim N(\mu, \Sigma)$

NG-BEVA

$\{d_i\}_1^N \sim \mathbf{\Gamma}(k, \Theta)$ (Gamma)

NG-BEVA
Gamma Fitting

- Pdf of the Gamma Distribution $\mathbf{\Gamma}(k, \Theta)$:

$$f(u) = \frac{1}{\Theta^k \Gamma(k)} u^{k-1} e^{-u/\Theta} \qquad \text{with} \quad u \geq 0$$

Gamma function

- The Chi-squared distribution is a special case of the Gamma distribution

$$\chi^2(p) \triangleq \mathbf{\Gamma}(k = \frac{p}{2}, \Theta = 2)$$

- Θ and k are estimated using Maximum Likelihood

NG-BEVA

Maximum Likelihood Estimation

- **Iterative estimation of k**

$$k_0 = \frac{3 - s + \sqrt{(s-3)^2 + 24s}}{12s}$$

$$k_{i+1} = k_i - \frac{ln(k_i) - \psi(k_i) - s}{\frac{1}{k_i} - \psi'(k_i)}$$

$$s = ln(\frac{1}{N}\sum_{i=1}^{N} x_i) - \frac{1}{N}\sum_{i=1}^{N} ln(x_i)$$

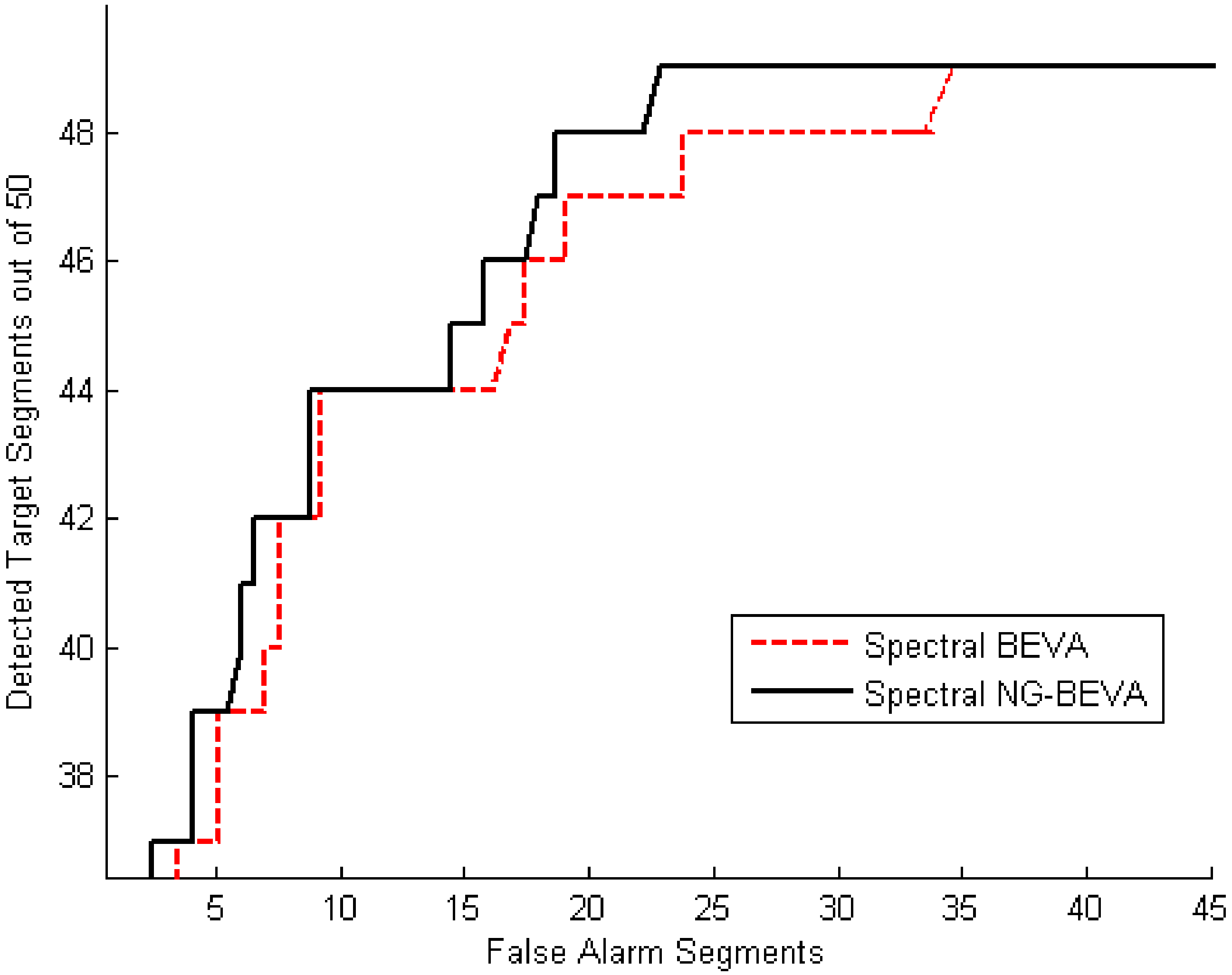
$$\psi(k) = \frac{\Gamma'(k)}{\Gamma(k)}$$

- **Estimation of Θ**

$$\hat{\Theta} = \frac{1}{kN}\sum_{i=1}^{N} x_i$$

NG-BEVA

ROC



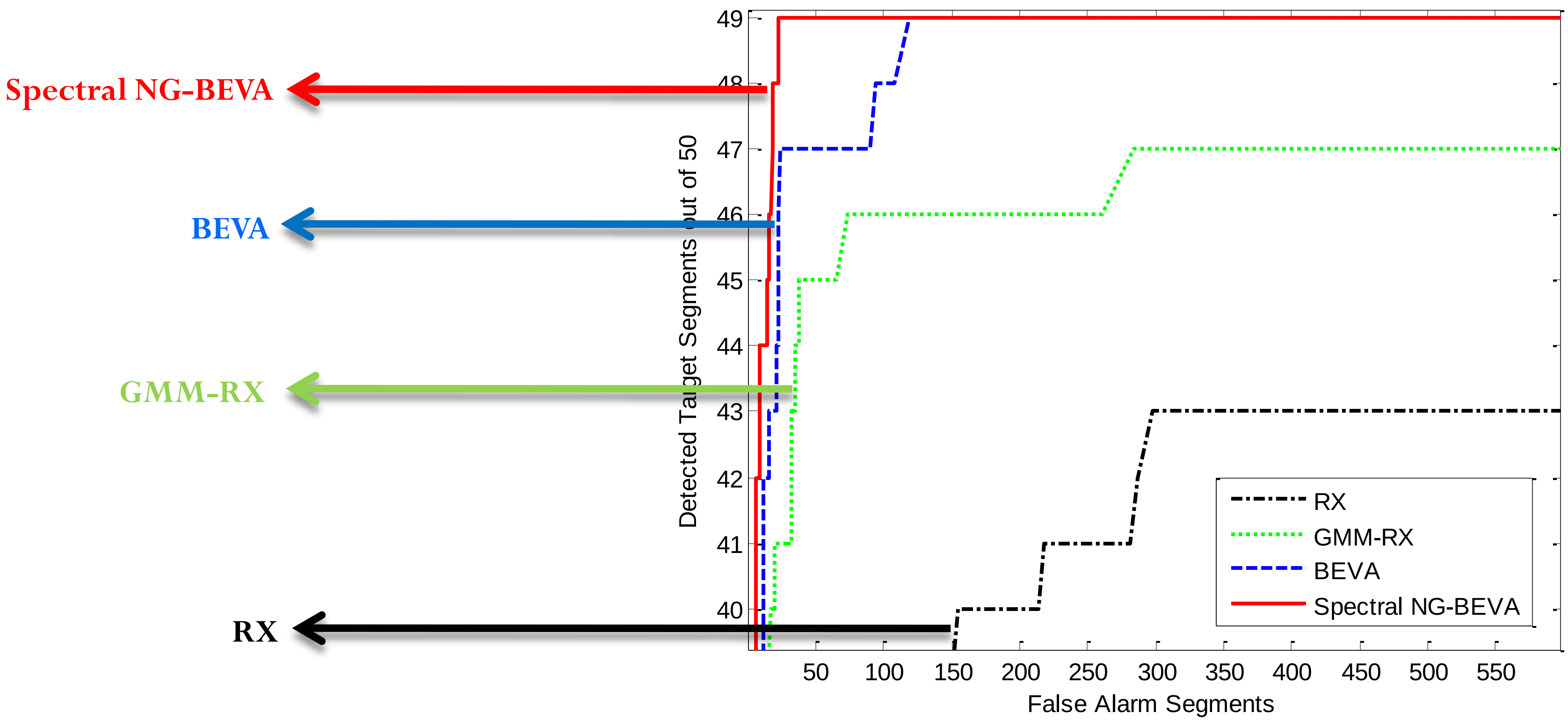
Detected Target Segments out of 50
48
46
44
42
40
38
5
10
15
20
25
30
35
40
45
False Alarm Segments
Spectral BEVA
Spectral NG-BEVA

Outline

- Hyperspectral Imaging
  - Image Acquisition
  - Anomaly Detection
- Statistical Background Modeling
  - Local Approach
  - Global Approach
- Combined Local/Global Approach – Proposed Algorithm
  - Local Part
  - Global Part
- Improvements of the Proposed Algorithm
  - Spectral Clustering
  - Non-Gaussian Fitting
- **Summary and Future Work**

Summary

ROC



Spectral NG-BEVA
BEVA
GMM-RX
RX
49
48
47
46
45
44
43
42
41
40
Detected Target Segments out of 50
50
100
150
200
250
300
350
400
450
500
550
False Alarm Segments
RX
GMM-RX
BEVA
Spectral NG-BEVA

Summary

- ***Combined Local-Global proposed algorithm***
  - *Local* – Spectral Clustering, greedy sequential estimation process and Gamma distribution fitting
  - *Global* – Filtering using large image area statistics and an auxiliary dictionary

- ***Pros***
  - Reduces the vast number of degrees of freedom while retaining the ability to be locally adjusted to the background.
  - Outperforms both standard local and global algorithms.

Computation Time

- ***Data***
  - 350x350
  - 65 bands
- ***BEVA***
  - Local part 35x35
  - Global part 350x280
- ***Computer***
  - Intel Core 2 duo 2Ghz
  - 2 GB Ram
  - Environment - Matlab

RX	**770 sec**
GMM-RX	**27 sec**
BEVA	**61 sec**
Spectral BEVA	**413 sec**
Spectral NG BEVA	**491 sec**

Future Work

- Non-linear local-global algorithm based on a kernelization using the similarity map obtained by Spectral Clustering

- Automatically select the proper number of clusters for each local block

- Dimensionality reduction as a preprocessing stage of the BEVA algorithm


Thank You!