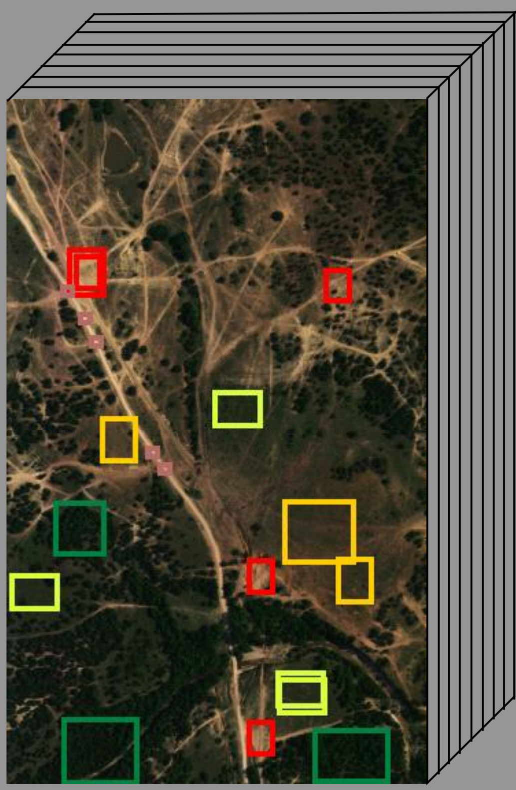
# Locating the Most Relevant Identification Information in Spectral Cubes and Similar Large Data Sets




Robert K. McConnell
WAY-2C
Arlington, Massachusetts
rkm@way2c.com
http://www.way2c.com


WAY-2C

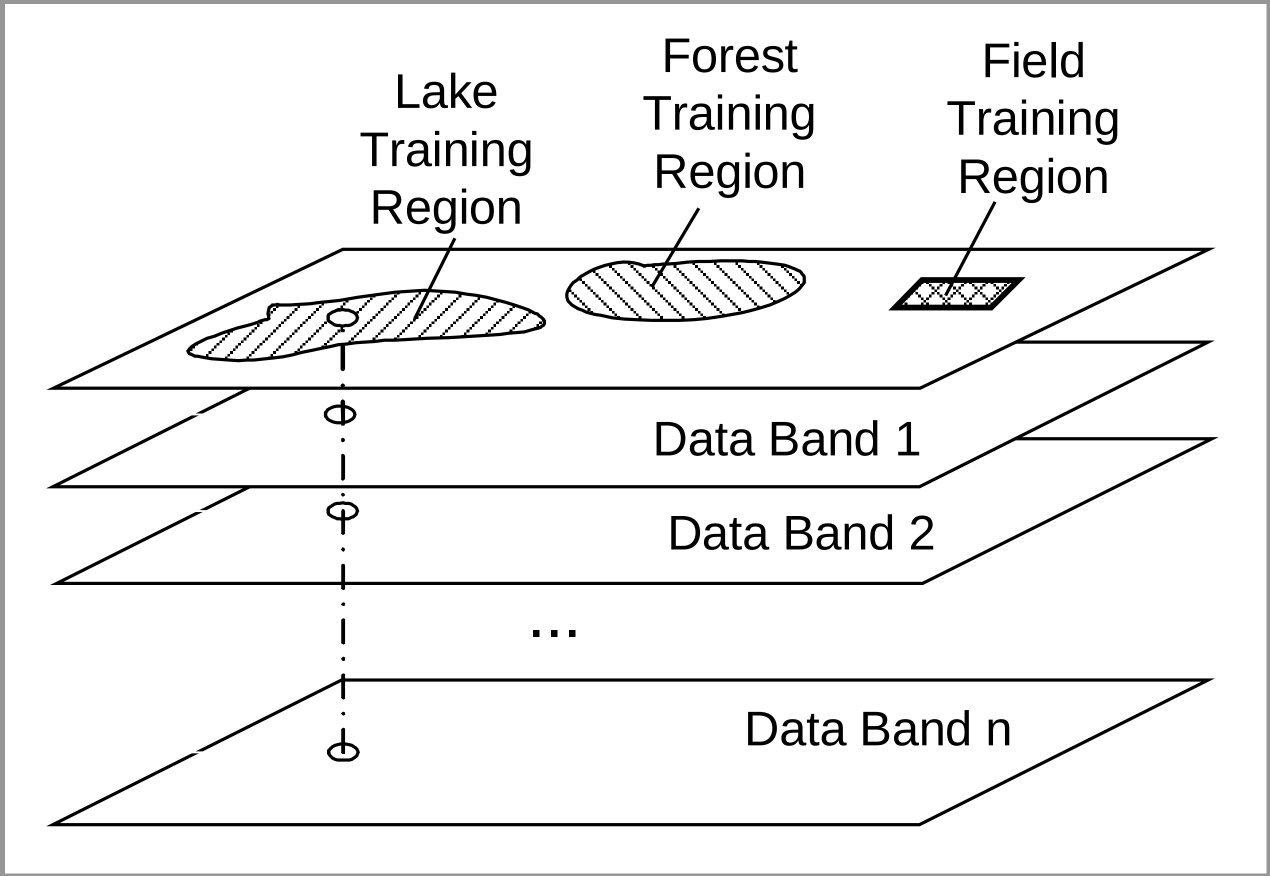
# Typical Object Detection Steps

1. Choose the target class(es).
2. Choose attributes to be sensed.
3. Acquire training data set (e.g. image).
4. Train the system.
5. Acquire a test data set.
6. Classify the test data set.
7. Locate blobs representing the class(es) of interest.

# Typical Object Detection Steps

1. Choose the target class(es).
2. Choose attributes to be sensed.
3. Acquire training data set (e.g. image).
4. Train the system
5. Acquire a test data set.
6. Classify the test data set.
7. Locate blobs representing the class(es) of interest.

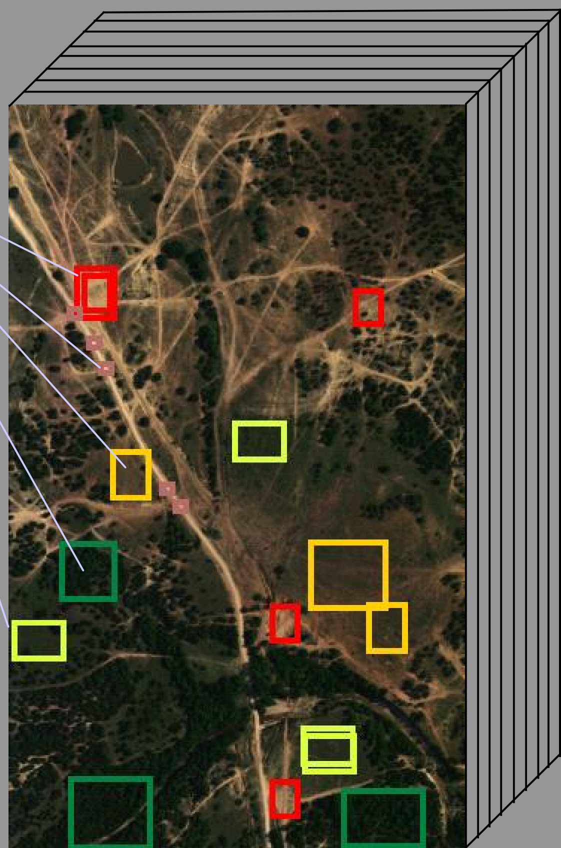
# Example: hyperspectral “cube”




Find optimum data subset for classification

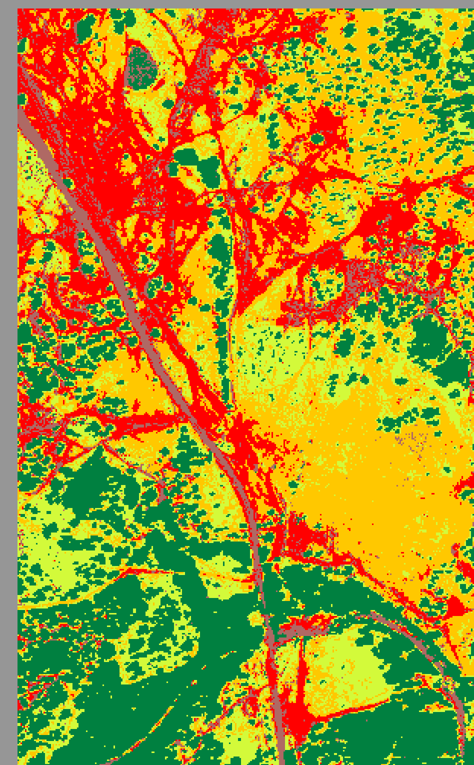
# Classification Objectives

Training regions



?

ASAP

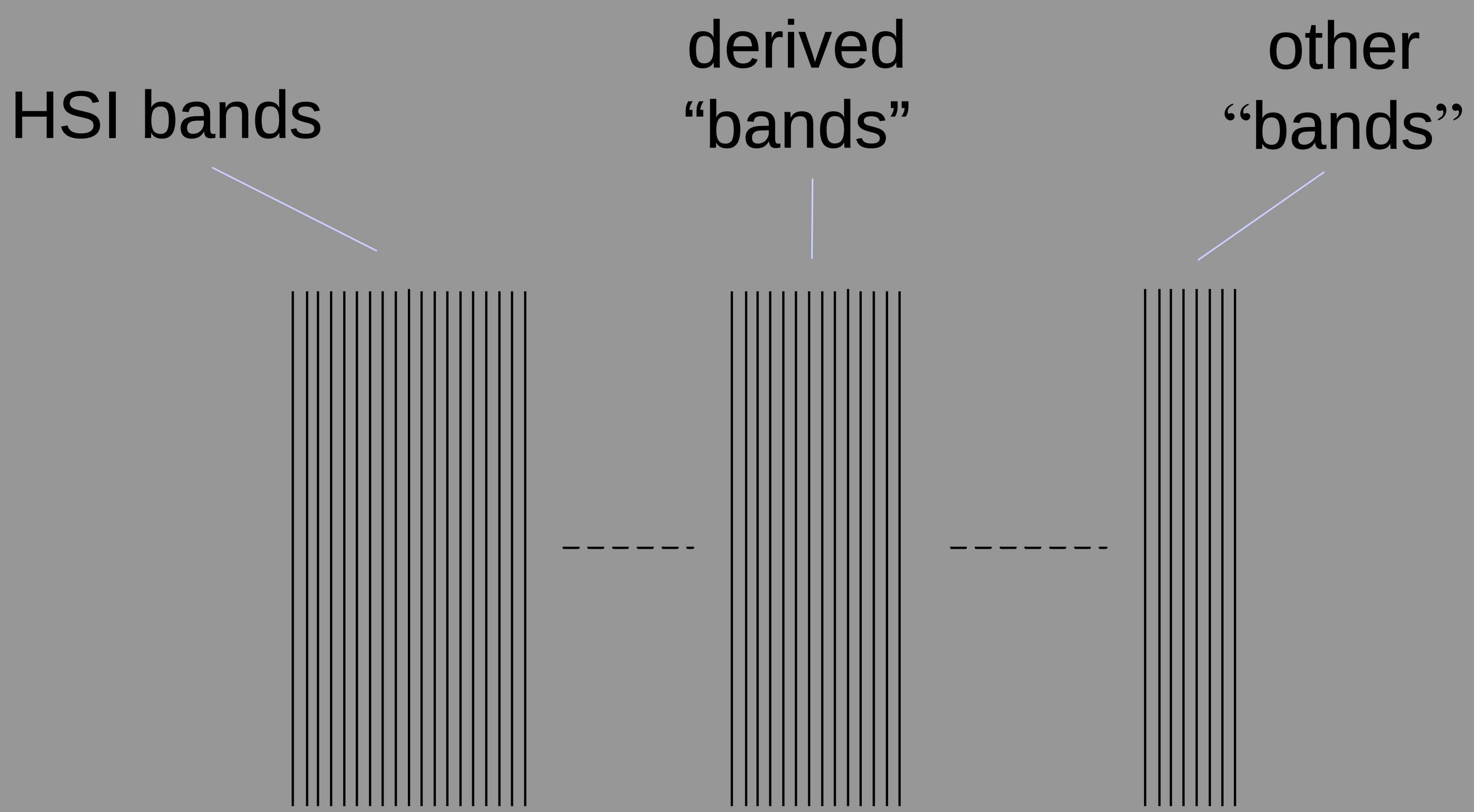


>1K bits/pixel

<5 bits/pixel

**WAY-2C**

# acquired bands and more



# Example: “Green Chips” Hyperspectral Cube

- 80 bands
- 12 bits A/D/pixel/band
- 16 total bits/pixel/band
- ~1K bits/pixel

*Courtesy Resonon, Inc.*

# “True Color” Image From Cube

*“Green Chips”*

| | Band | nm |
|---|---|---|
| R | 40 | 642 |
| G | 25 | 550 |
| B | 10 | 458 |

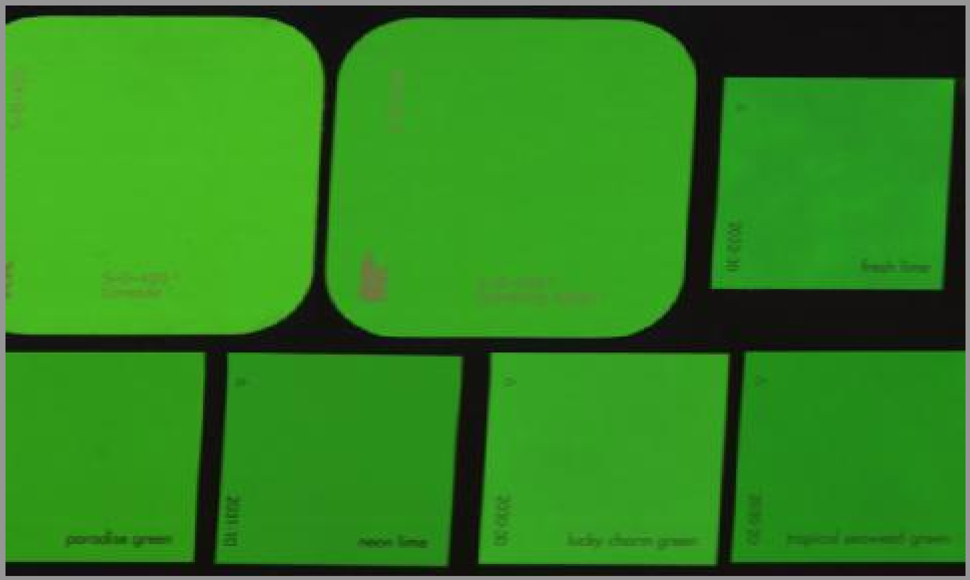


*Image Courtesy Resonon, Inc.*

## 8 Classes

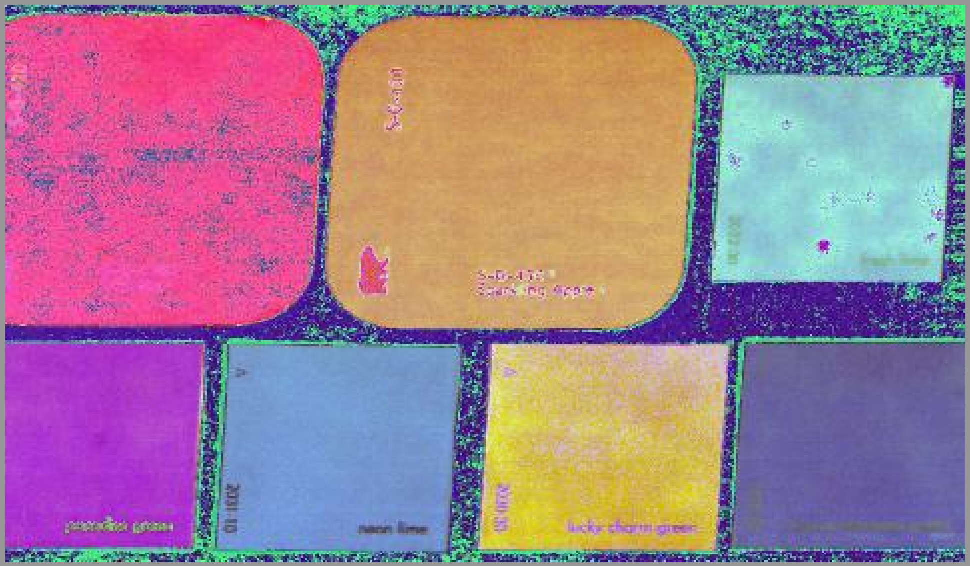
Required:

Band combination to differentiate reliably.

# “High Contrast” Image From Cube

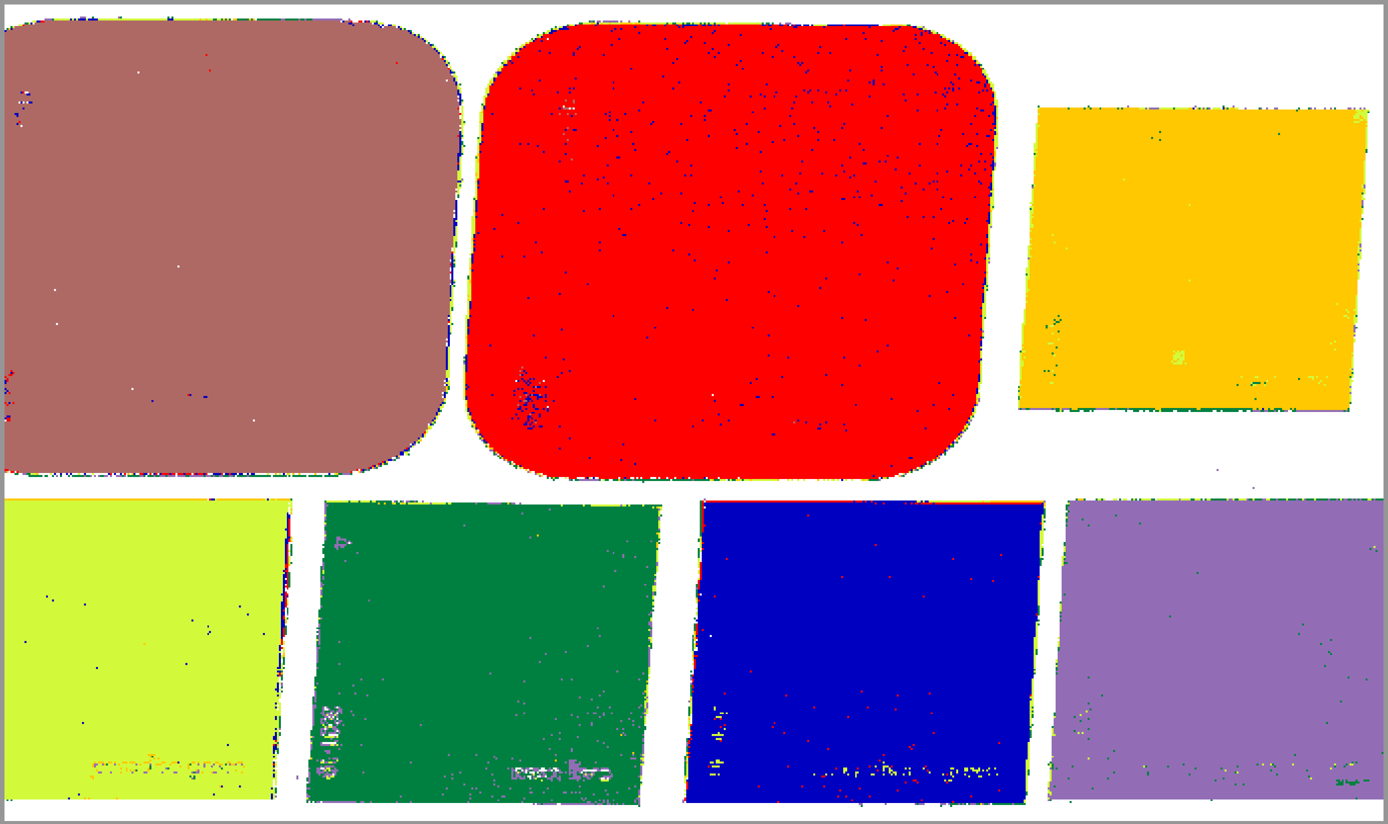
| | Band | nm |
|---|---|---|
| R | $60^{*}$ | 764 |
| G | $57^{*}$ | 745 |
| B | $17^{*}$ | 501 |

*some high order bits discarded

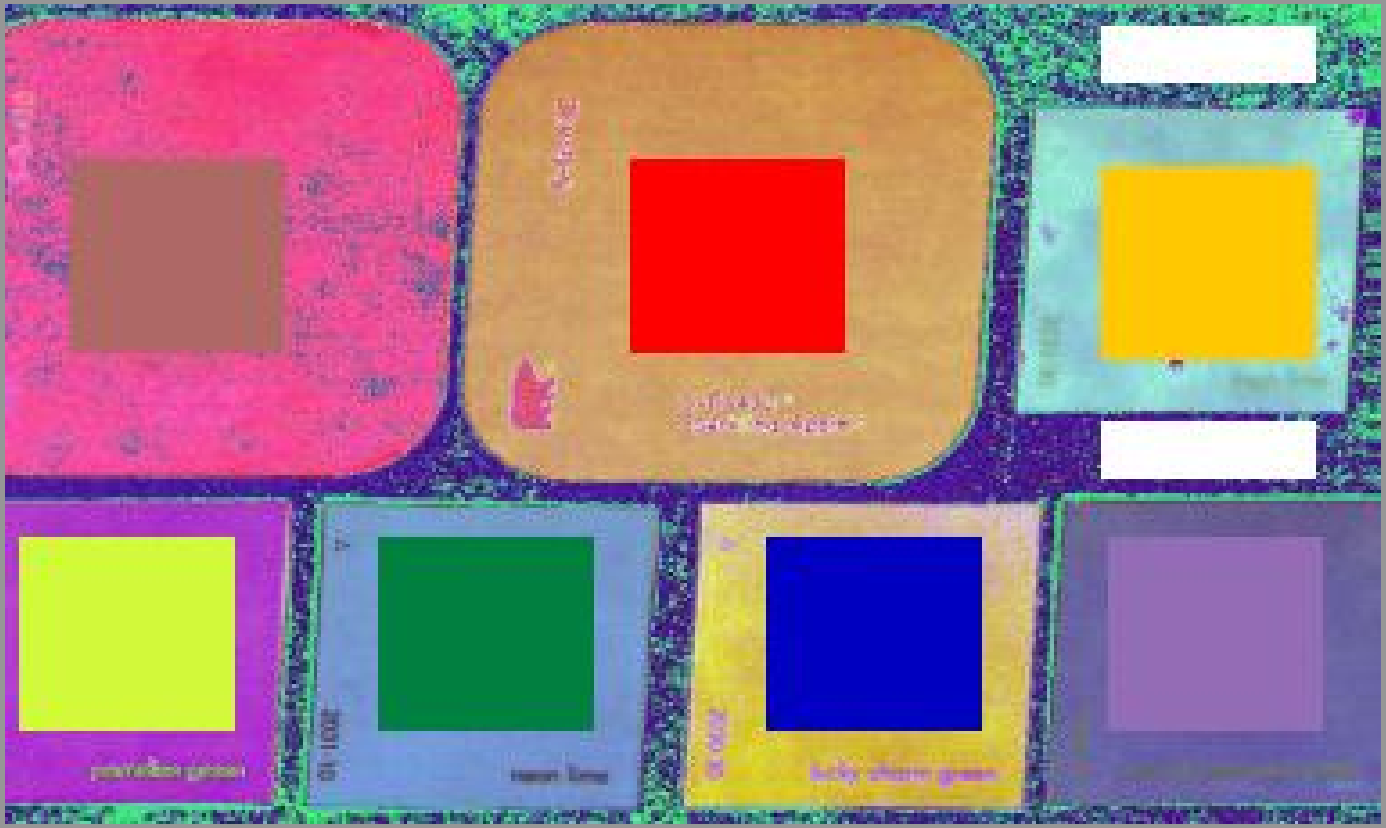


8 classes can be differentiated reliably.

# Pixel by pixel interpretation of high contrast “intermediate” image.

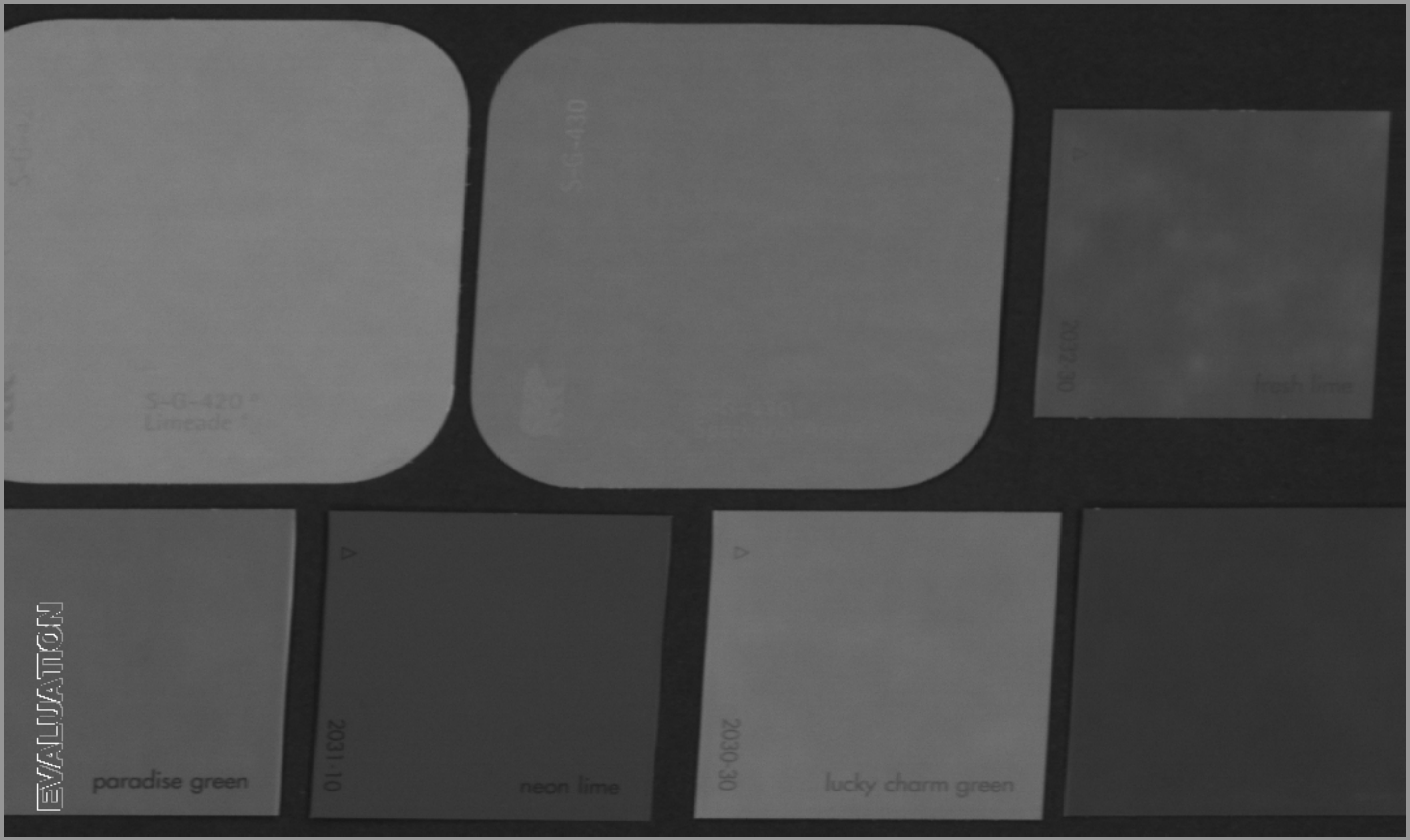


# Maximum likelihood classification based on ~30K pixel regions.



# Not all bands are equally relevant.

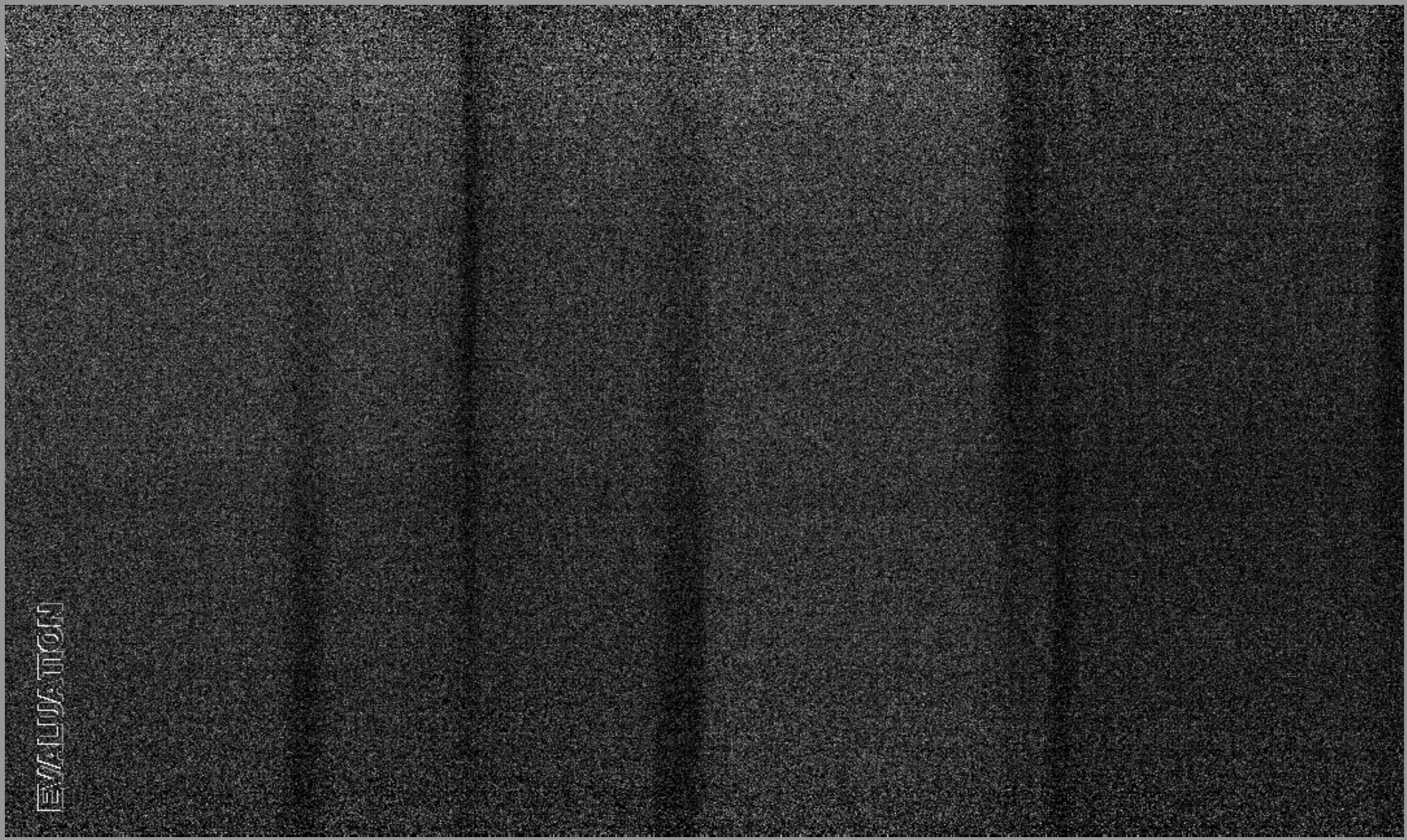
# Band 60 PSPA=67.88%





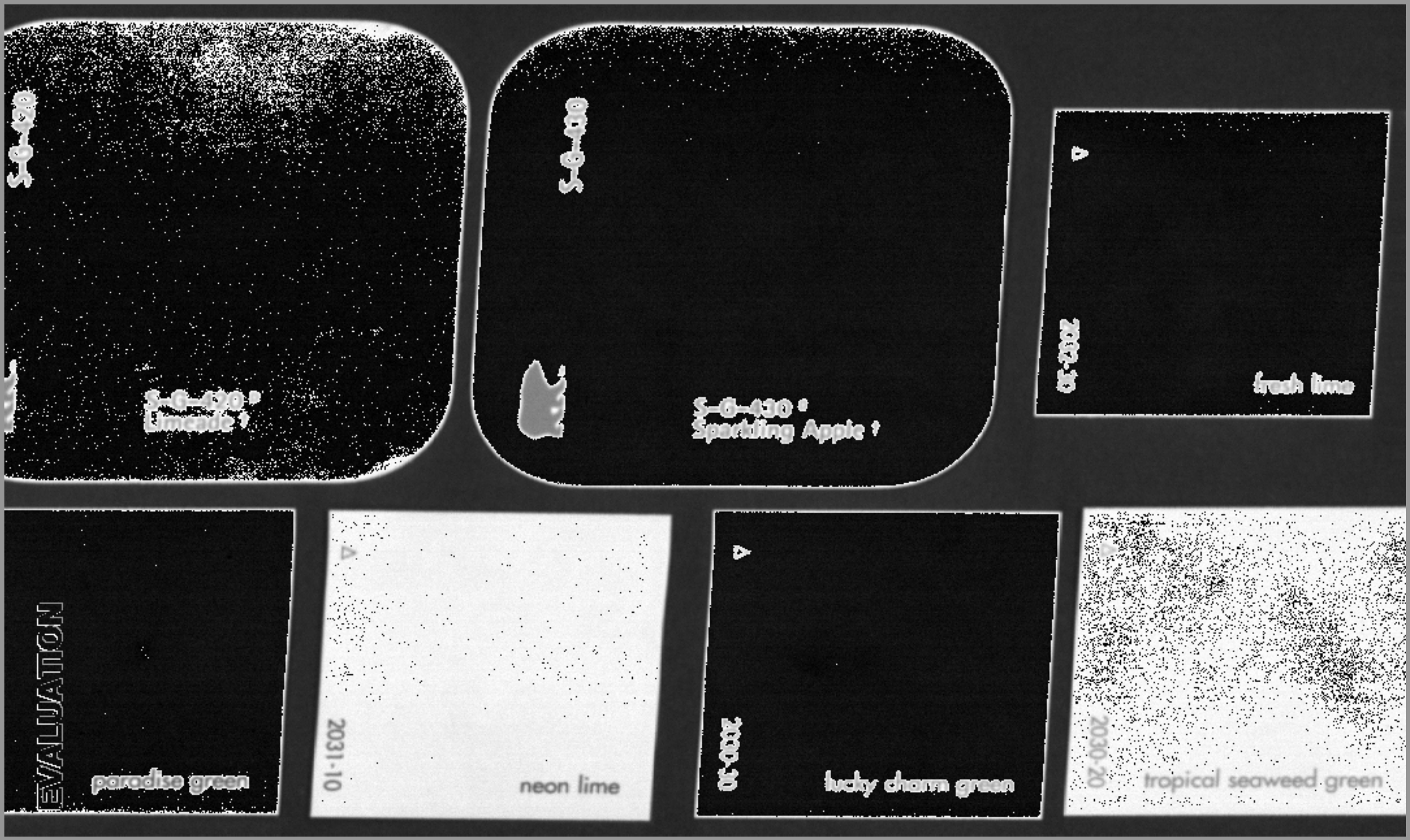

*Base Image Courtesy Resonon, Inc.*

# Band 00 PSPA=14.01%




*Base Image Courtesy Resonon, Inc.*

# Band 79 PSPA=27.19%



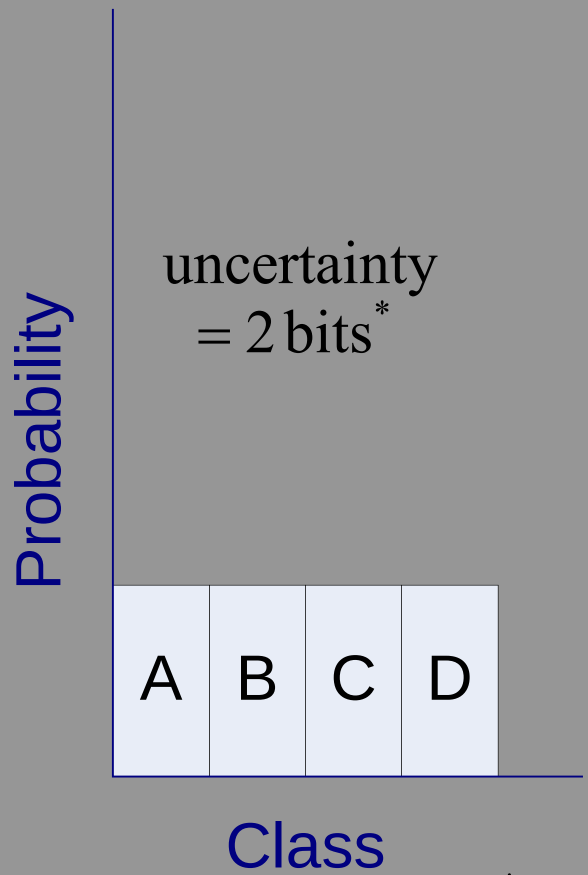

Base Image Courtesy Resonon, Inc.

# Entropy as a measure of uncertainty

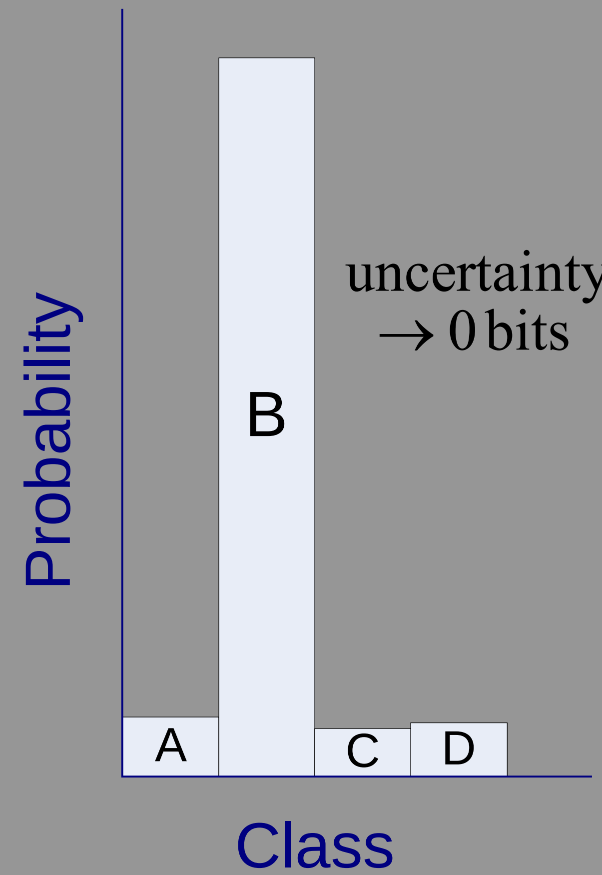
# “In mind’s eye”

## Classification→

Before




After

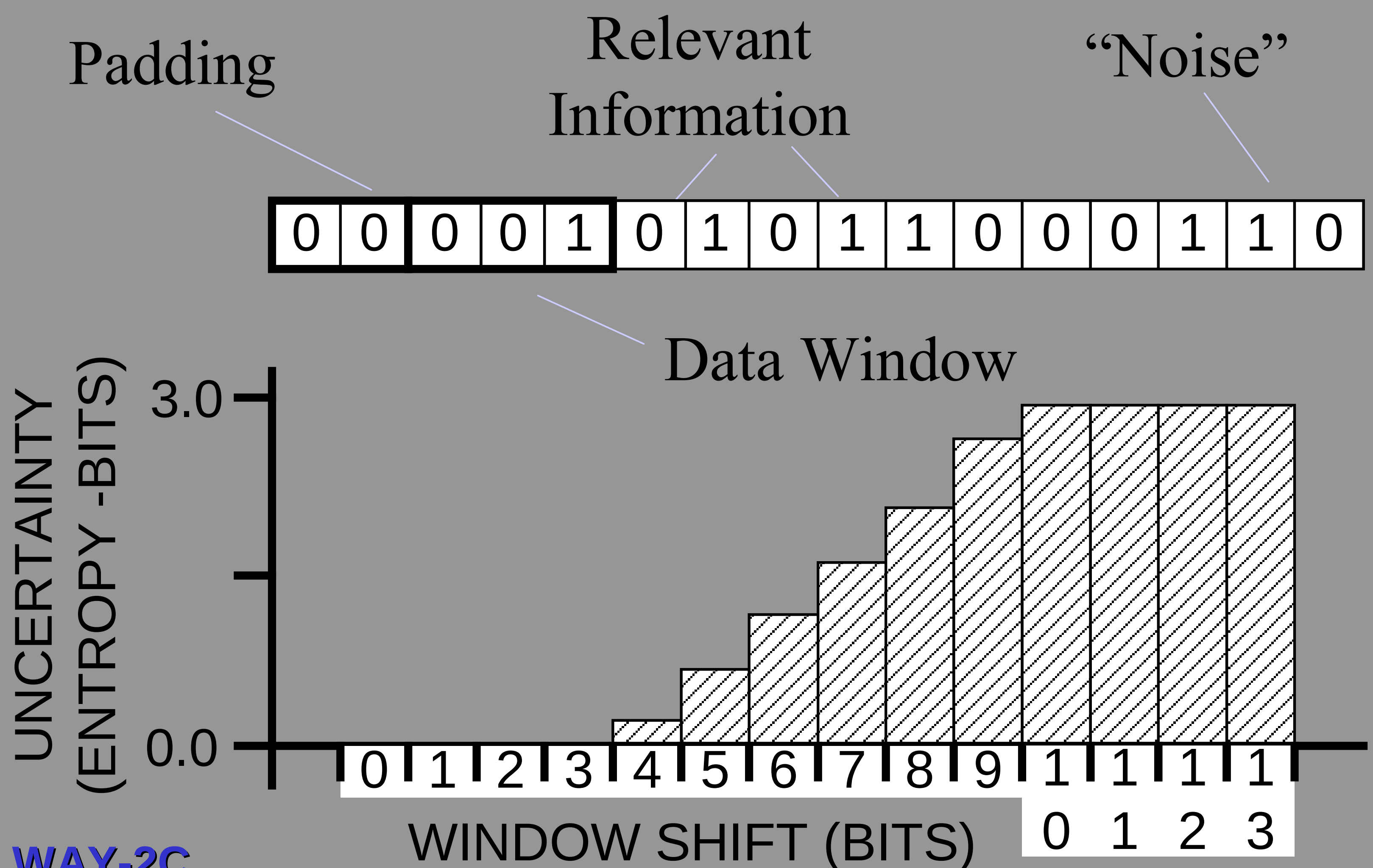



$$^{*}u = -\sum_{j}^{J} p_j \log_2(p_j)$$

**WAY-2C**

Not all positions within "band" are equally relevant.

# Most Relevant Information: Data Window Uncertainty



Objective: Find optimum combination of “bands” and data “windows” within those bands to differentiate the classes of interest.

# relevance

It can be shown that the relevant information in the data set $D$ for differentiating the class set $C$ (known as the mutual information between $C$ and $D$) is given by

$$\boxed{\mathit{relevance} = \mathrm{H}(C) + \mathrm{H}(D) - \mathrm{H}(C\&D)} \qquad (3)$$

where:

$\mathrm{H}(C)$ is the entropy of the class probability distribution,

$\mathrm{H}(D)$ is the entropy of the data distribution and

$\mathrm{H}(C\&D)$ is the entropy of the joint distribution of $C\&D$.

# uncertainty

It can also be shown that information required to completely specify the class once the data is known, must be given by

$$\boxed{\mathit{uncertainty} = \mathrm{H}(\boldsymbol{C\&D}) - \mathrm{H}(\boldsymbol{D})} \qquad (6)$$

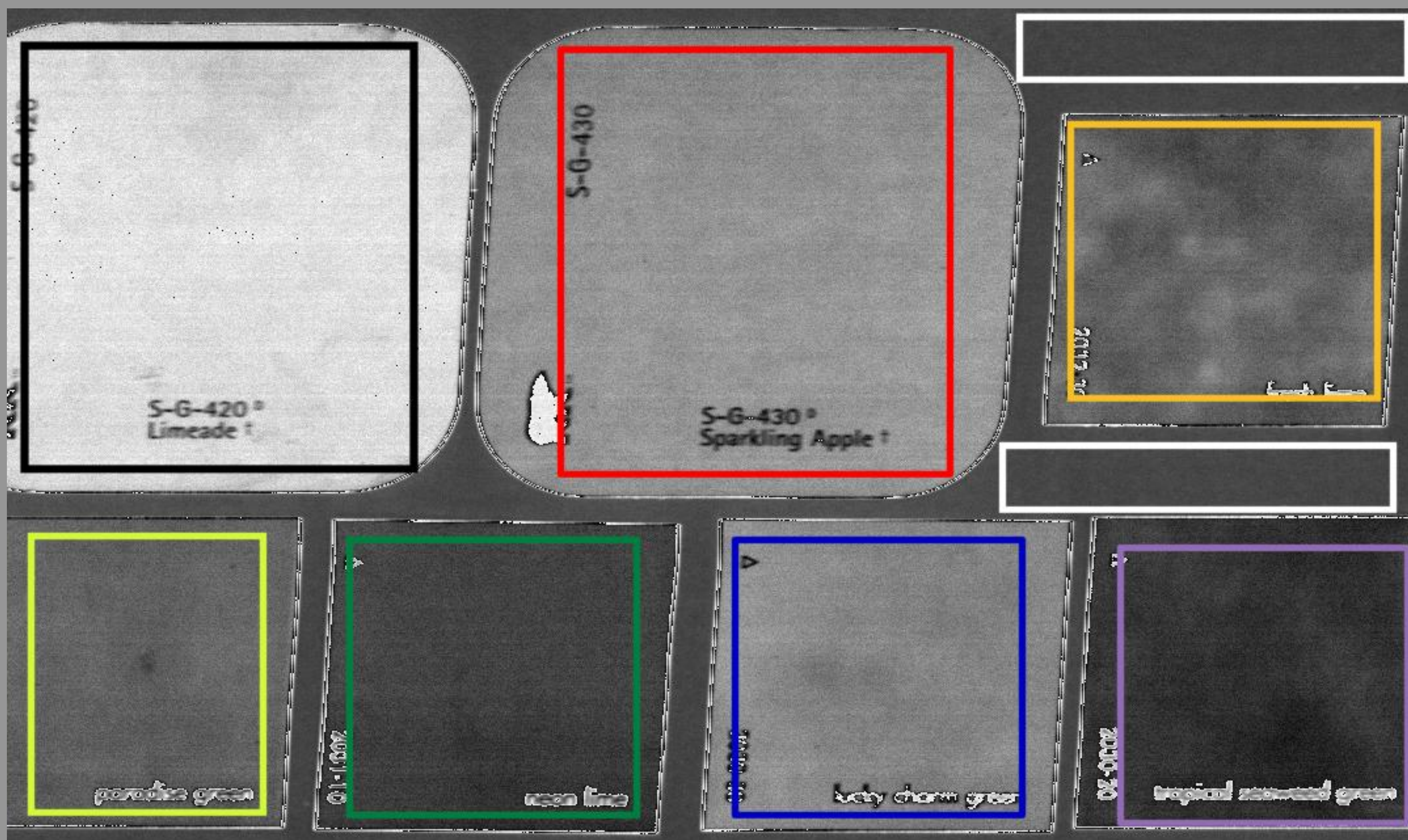
$= 0$ when $\boldsymbol{C}$ and $\boldsymbol{D}$ are perfectly correlated and

$= \mathrm{H}(\boldsymbol{C})$ for no correlation between $\boldsymbol{C}$ and $\boldsymbol{D}$.

Maximizing relevance minimizes uncertainty.

# Finding Most Relevant Data to Construct Intermediate Image for Classification

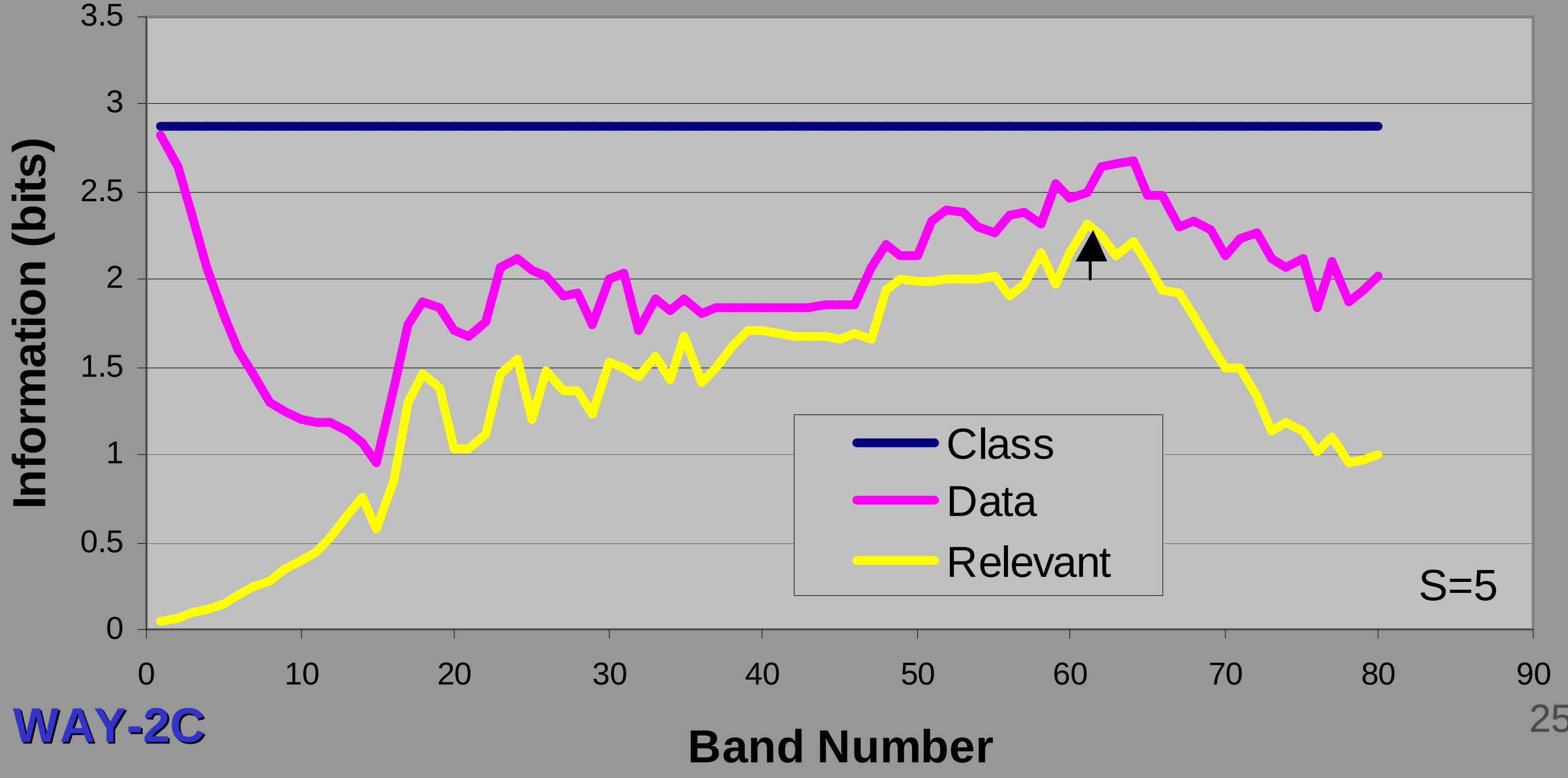
# Training Regions






*Base Image Courtesy Resonon, Inc.*

# Most relevant window: measured single band information contribution.

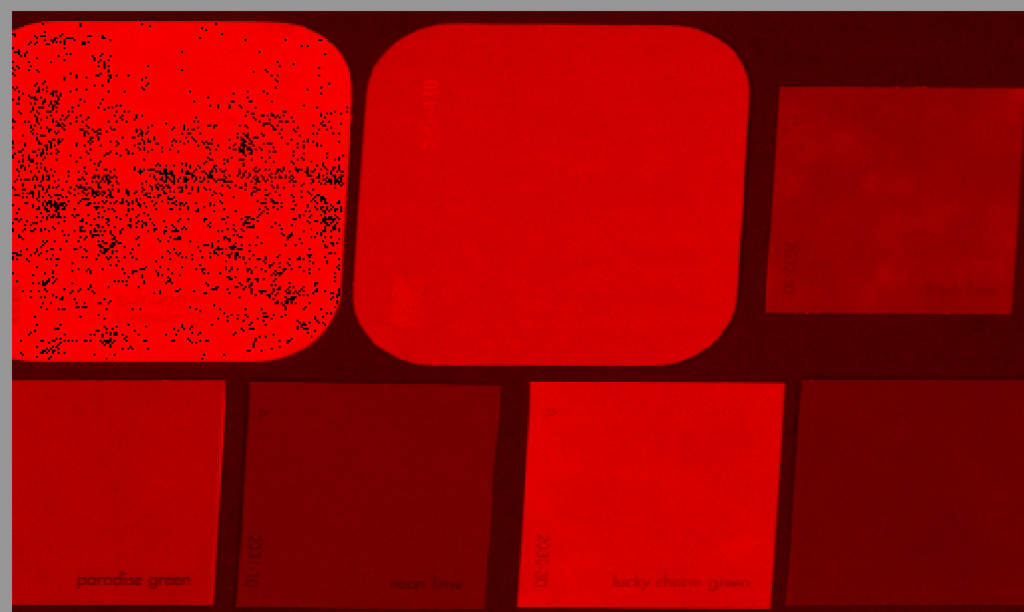


# Image constructed from single most relevant band and window only.

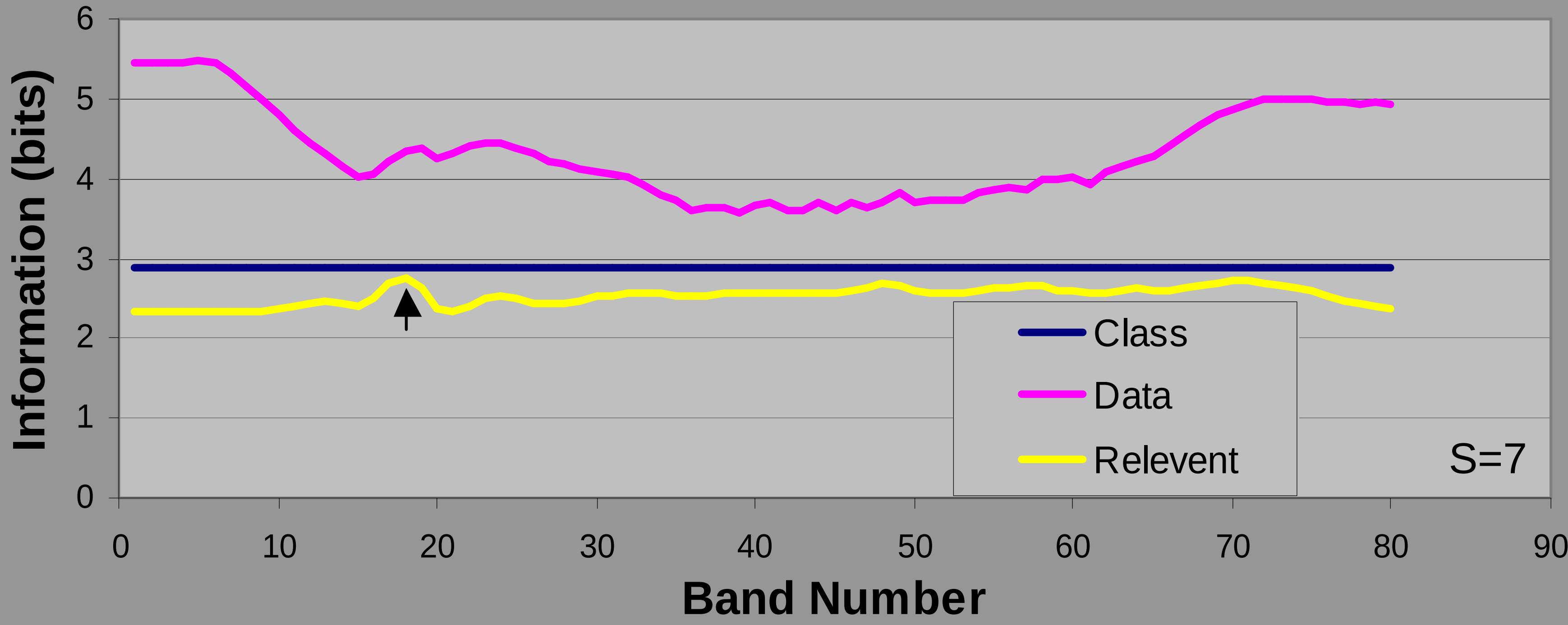
| Plane | Band | nm |
|---|---|---|
| R | $60^*$ | 764 |

rel = 2.318 bits

PSPA = 67.88%



# Relevance: (band & window) pair information contribution (band 60 window fixed)



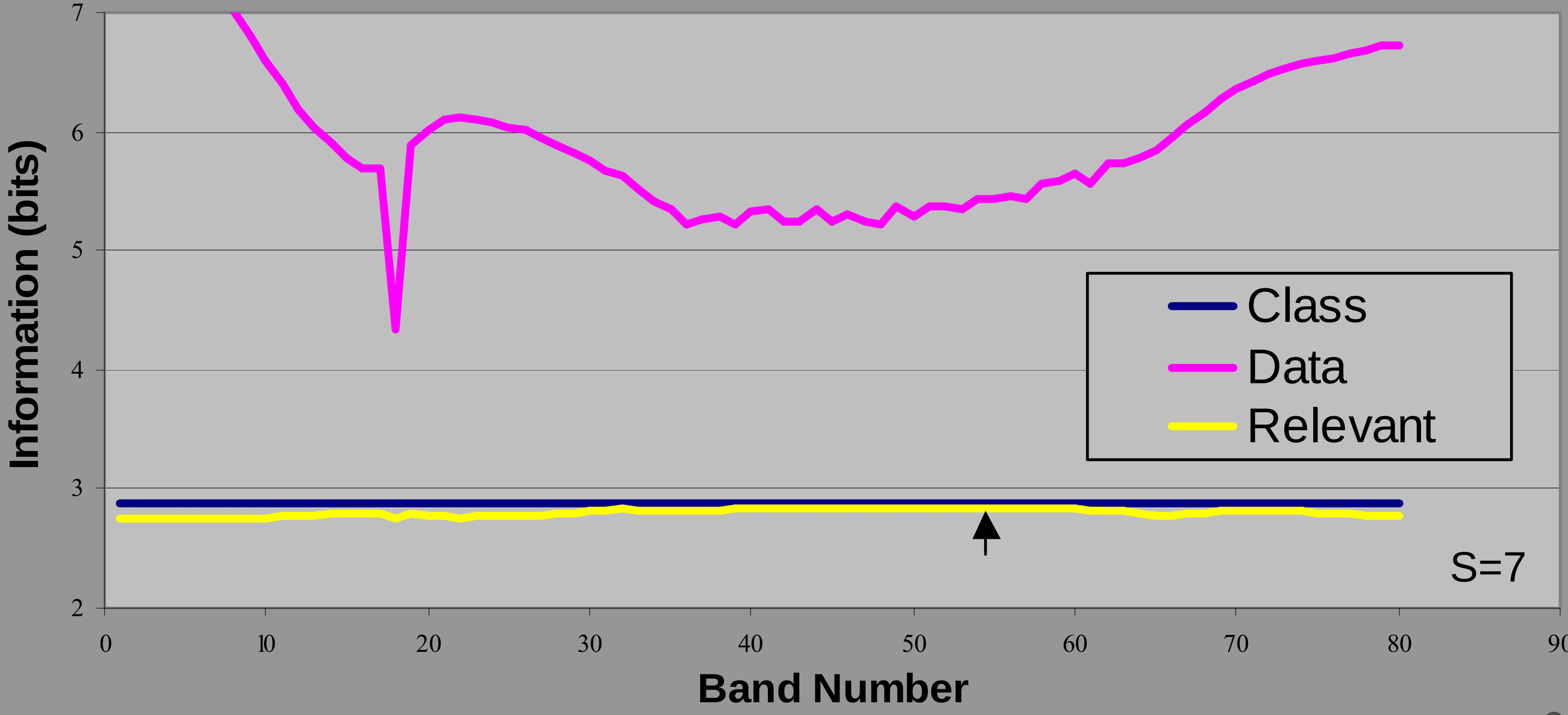

WAY-2C

# Image constructed from most relevant band pair.

| Plane | Band | nm |
|---|---|---|
| R | 60* | 764 |
| B | 17* | 501 |

rel = 2.742 bits
PSPA = 91.07%

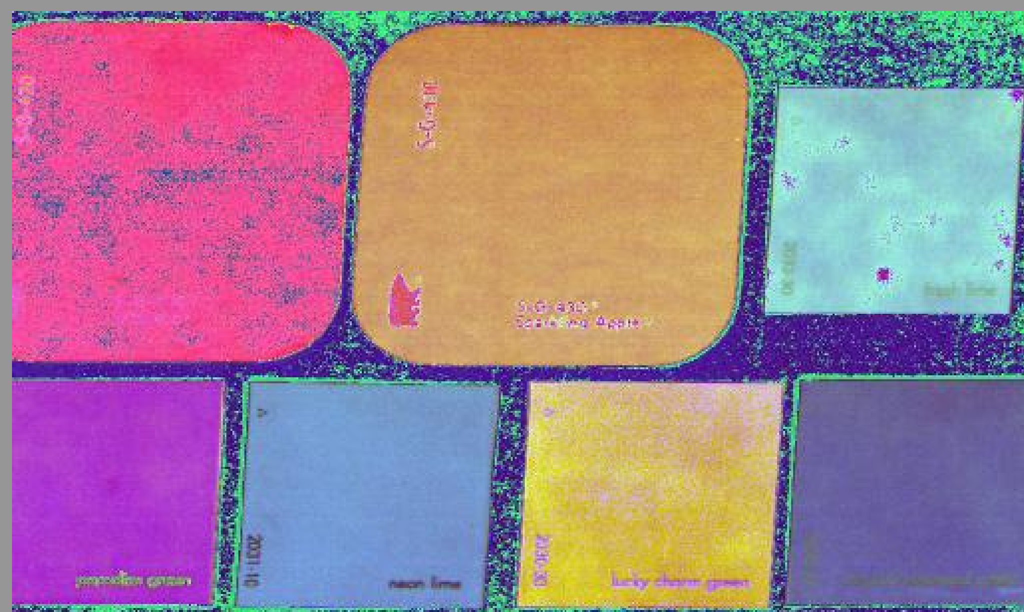
# Relevance: (band & window) triplet information contribution (bands 60,17 windows fixed)




**WAY-2C**

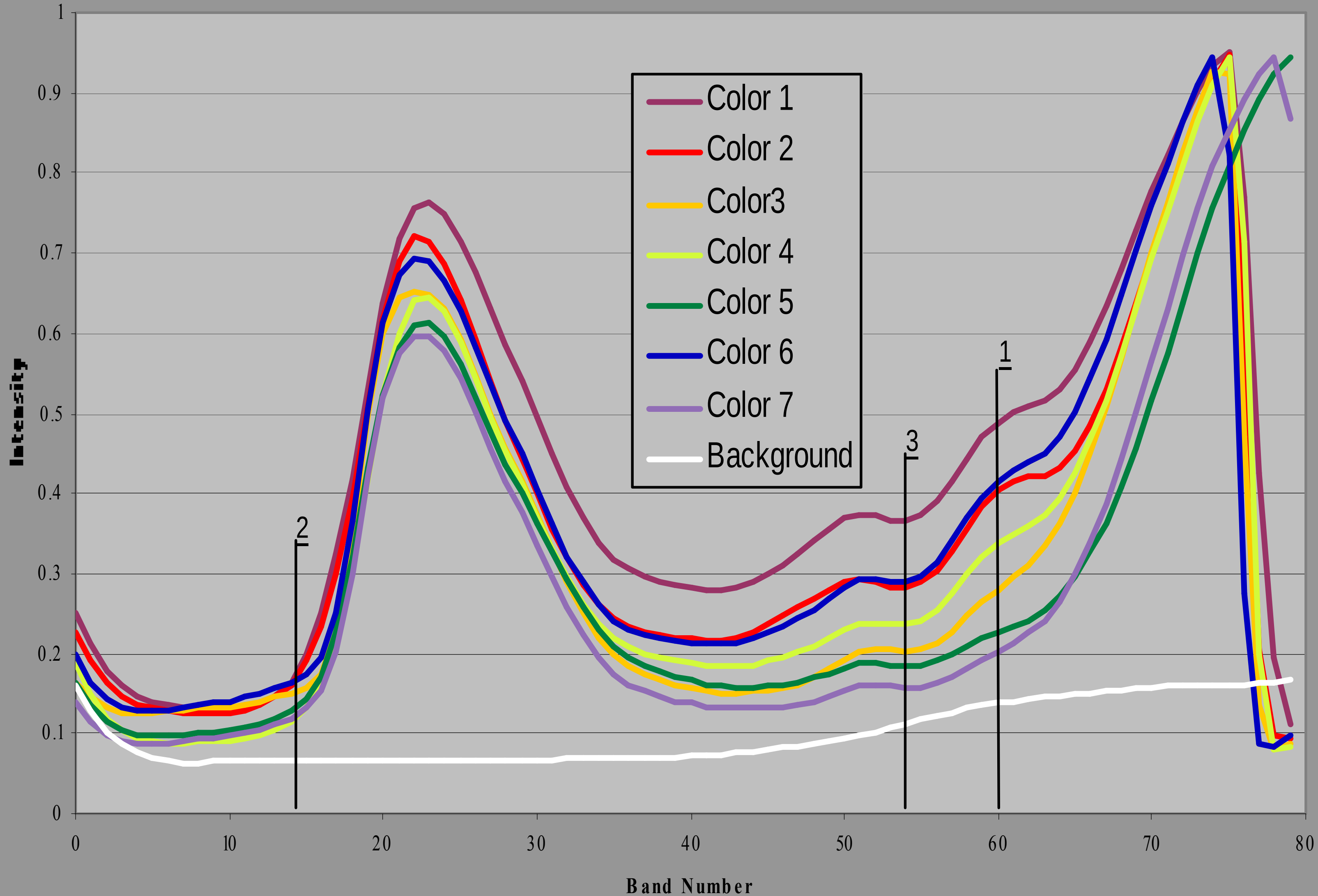
# Intermediate image from most relevant (band & window) triplet.

| Plane | Band | nm |
|---|---|---|
| R | 60* | 764 |
| G | 57* | 745 |
| B | 17* | 501 |



rel = 2.844 bits
PSPA = 97.74%

# Automatically Selected Bands



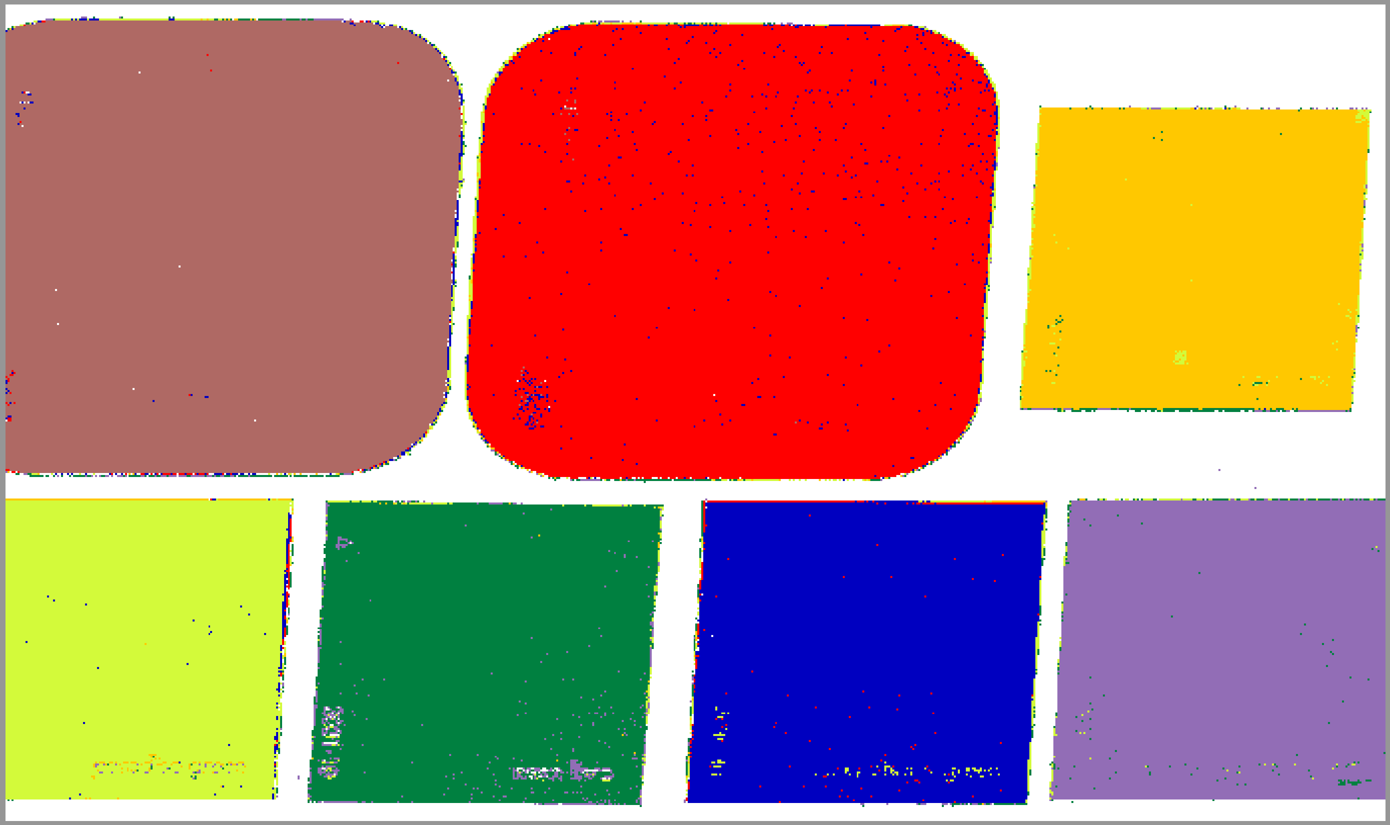



WAY-2C

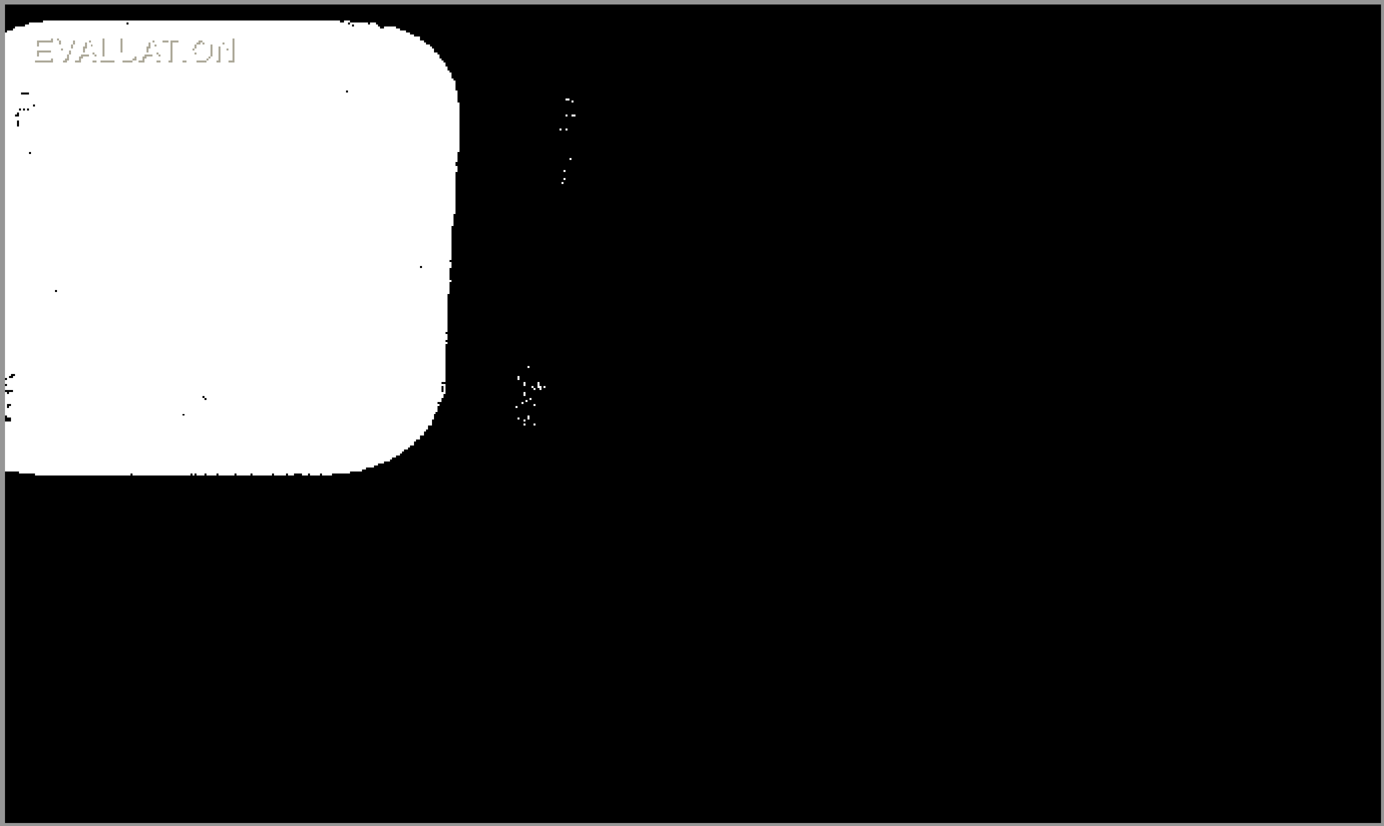
# Now, using only optimum data subset:

1. Construct an “intermediate” image.
2. Train using the same class training regions as for locating that subset.
3. Perform maximum likelihood interpretations on this and similar images based only on the optimum data subset.

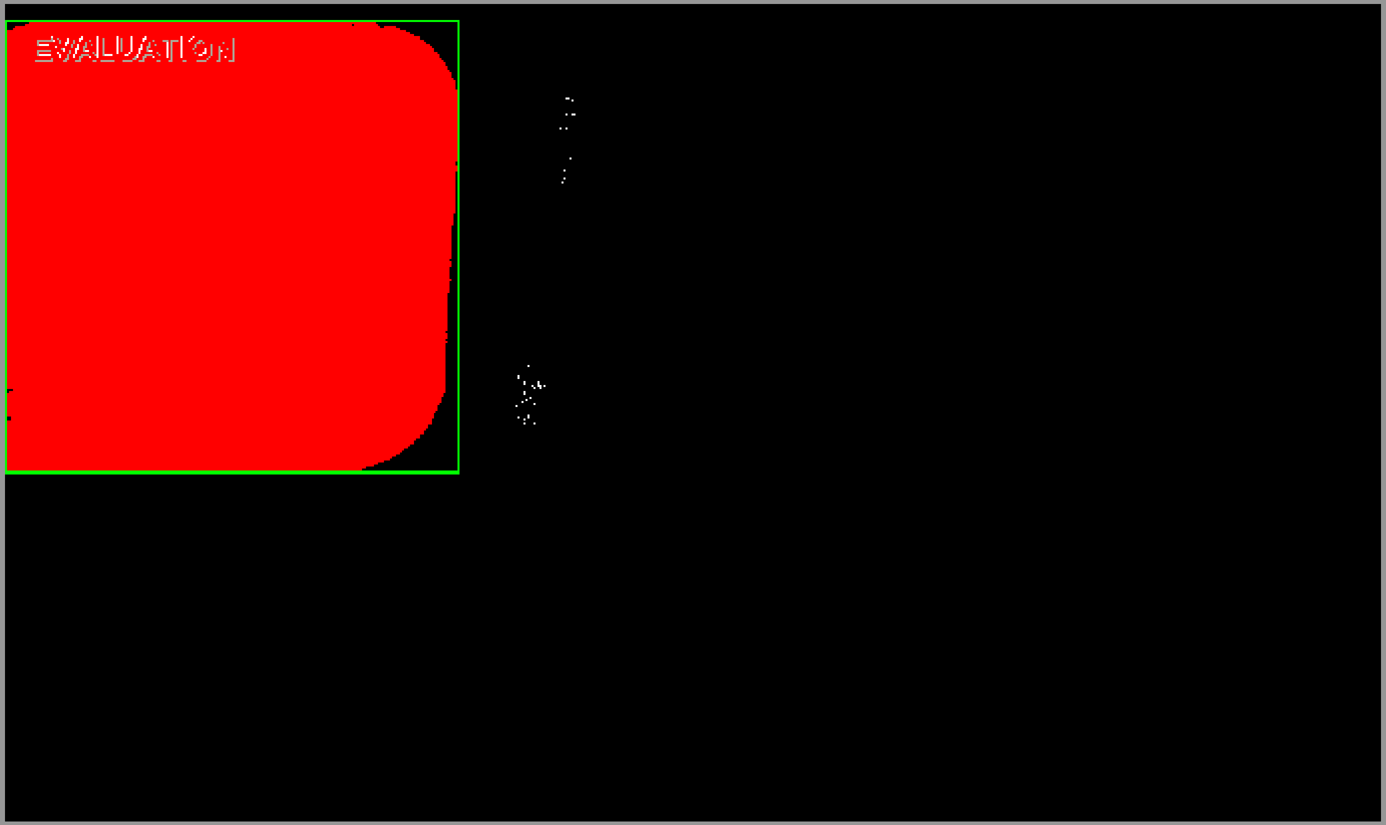
# Pixel by pixel interpretation of most relevant band triplet image.



# Pixel by pixel class # 1 extraction

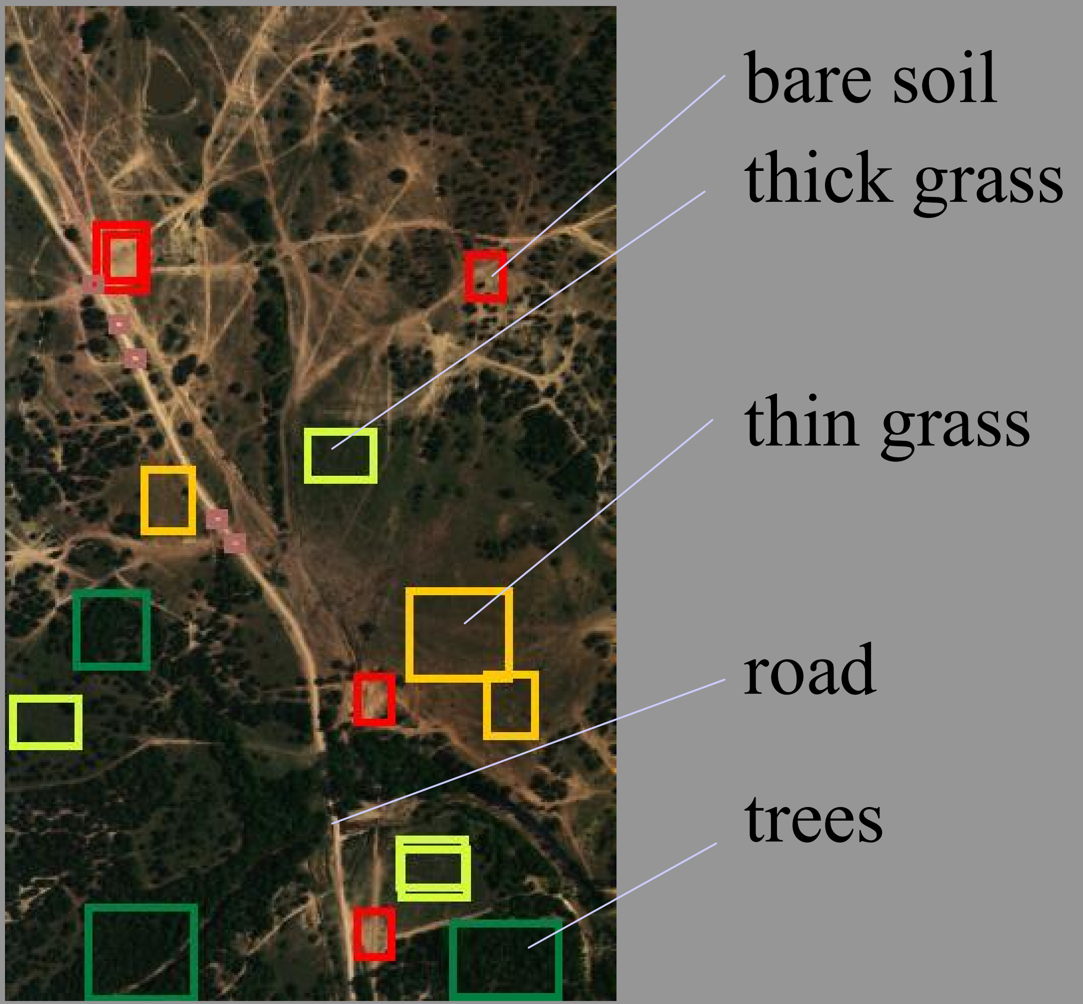


# Blob location and processing.

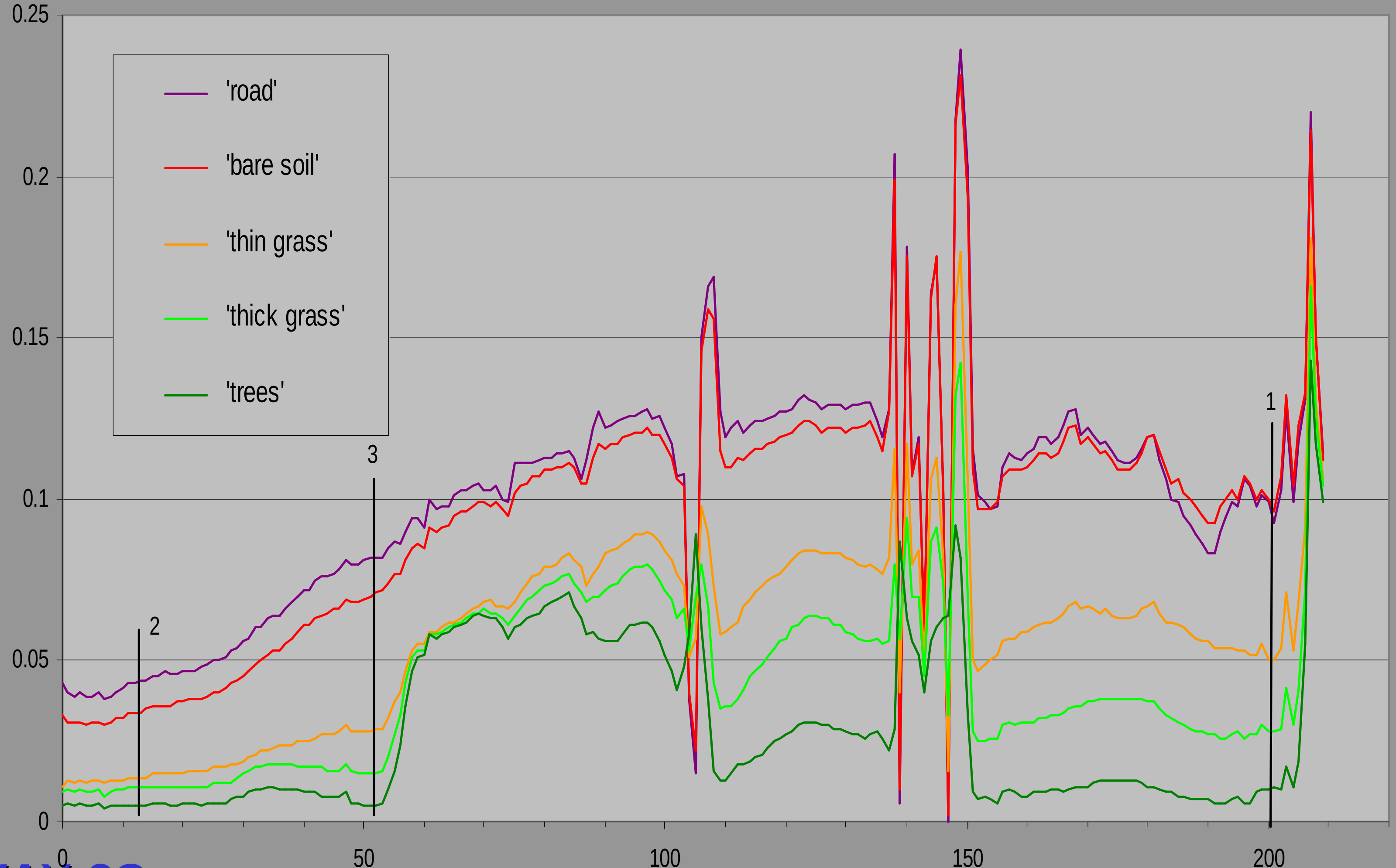


# “True Color” TERRAIN Image

HYDICE
bands 49,35,15



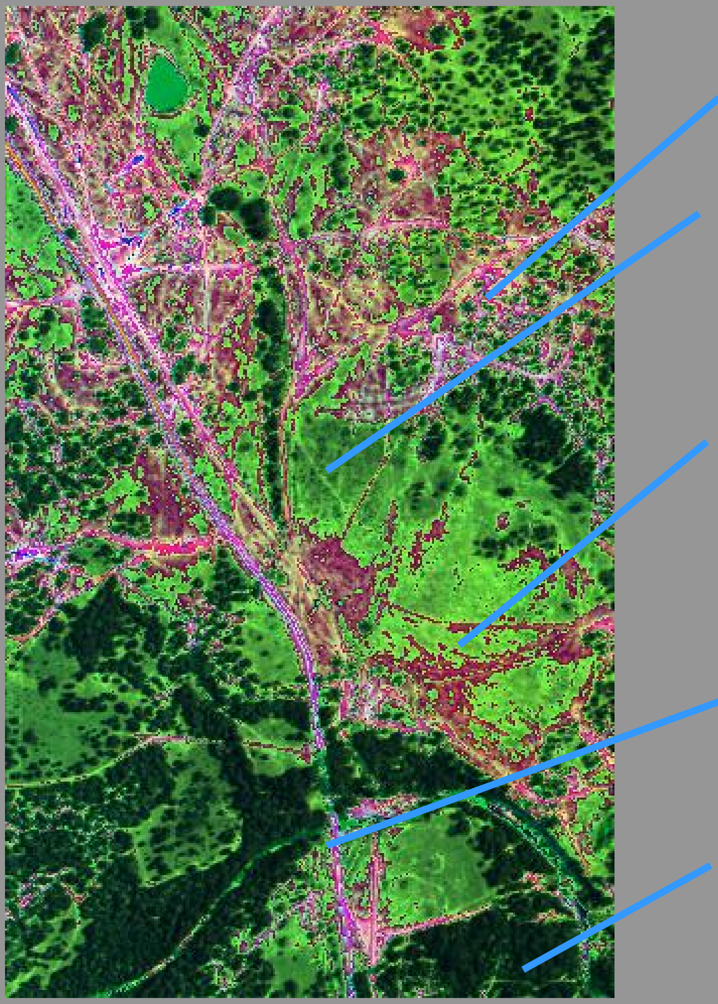
# Terrain: Mean Spectra and Selected Bands



# “Most Relevant” TERRAIN Image

HYDICE
bands 199,52,17

> 98% data reduction from original hyperspectral cube.

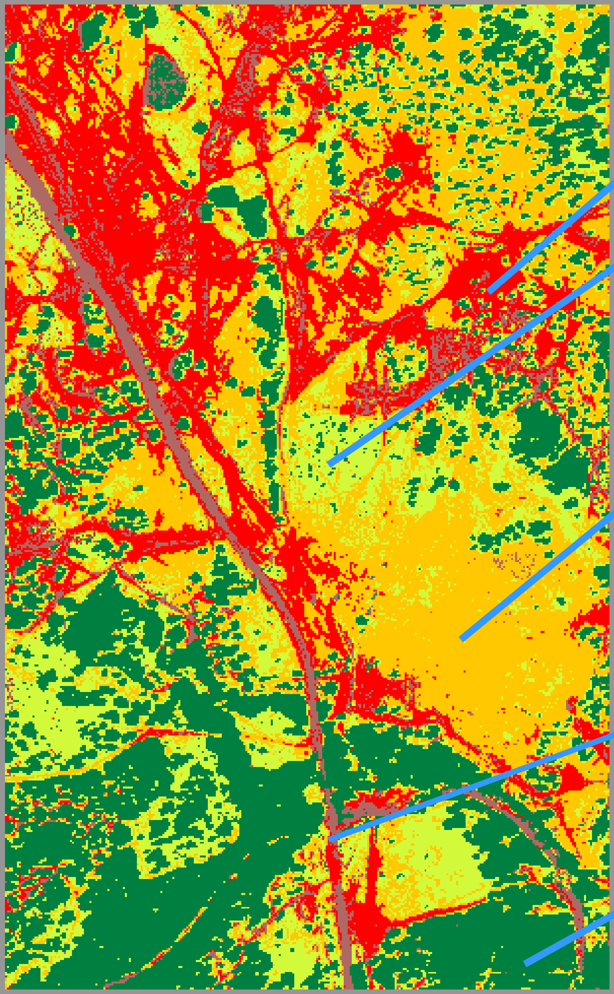



WAY-2C

# Interpreted TERRAIN Image

Based on
HYDICE
bands 199,52,17

uncertainty = 0.48 bits
PSPA = 71.8%





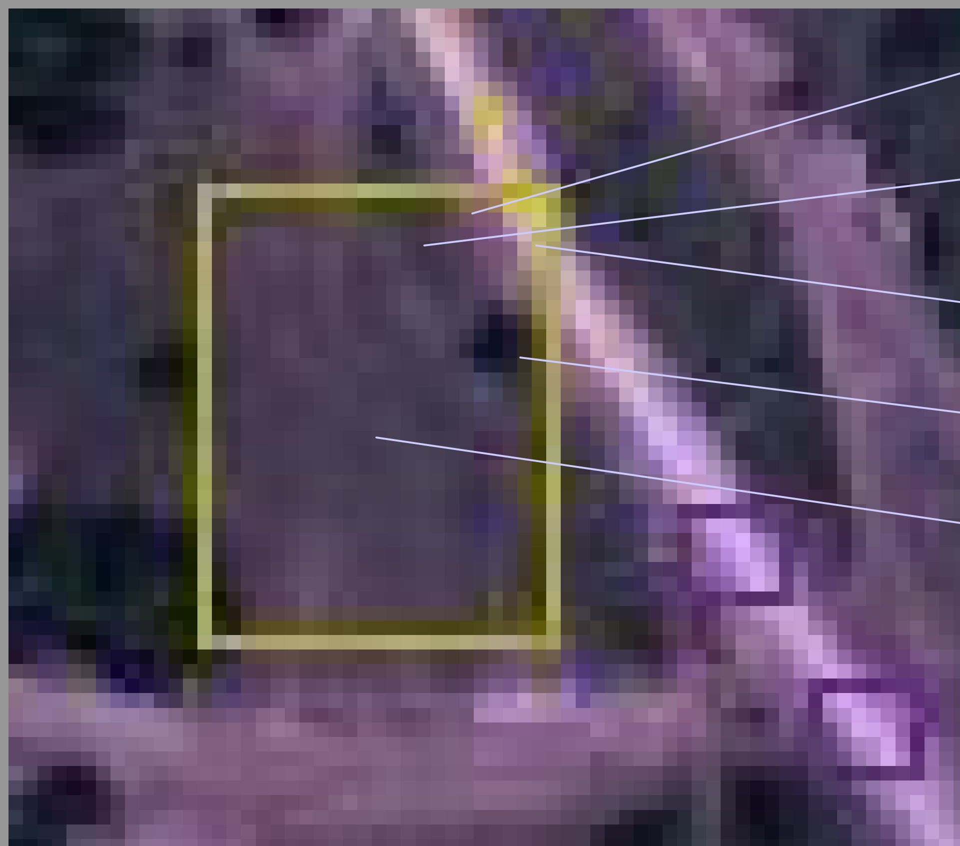
WAY-2C

# Predicted Agreement

We define the *PSPA*, or predicted single pixel agreement (with training) by:

$$\boxed{\log(PSPA) = \mathrm{H}(D) - \mathrm{H}(C\&D)} \qquad (7)$$

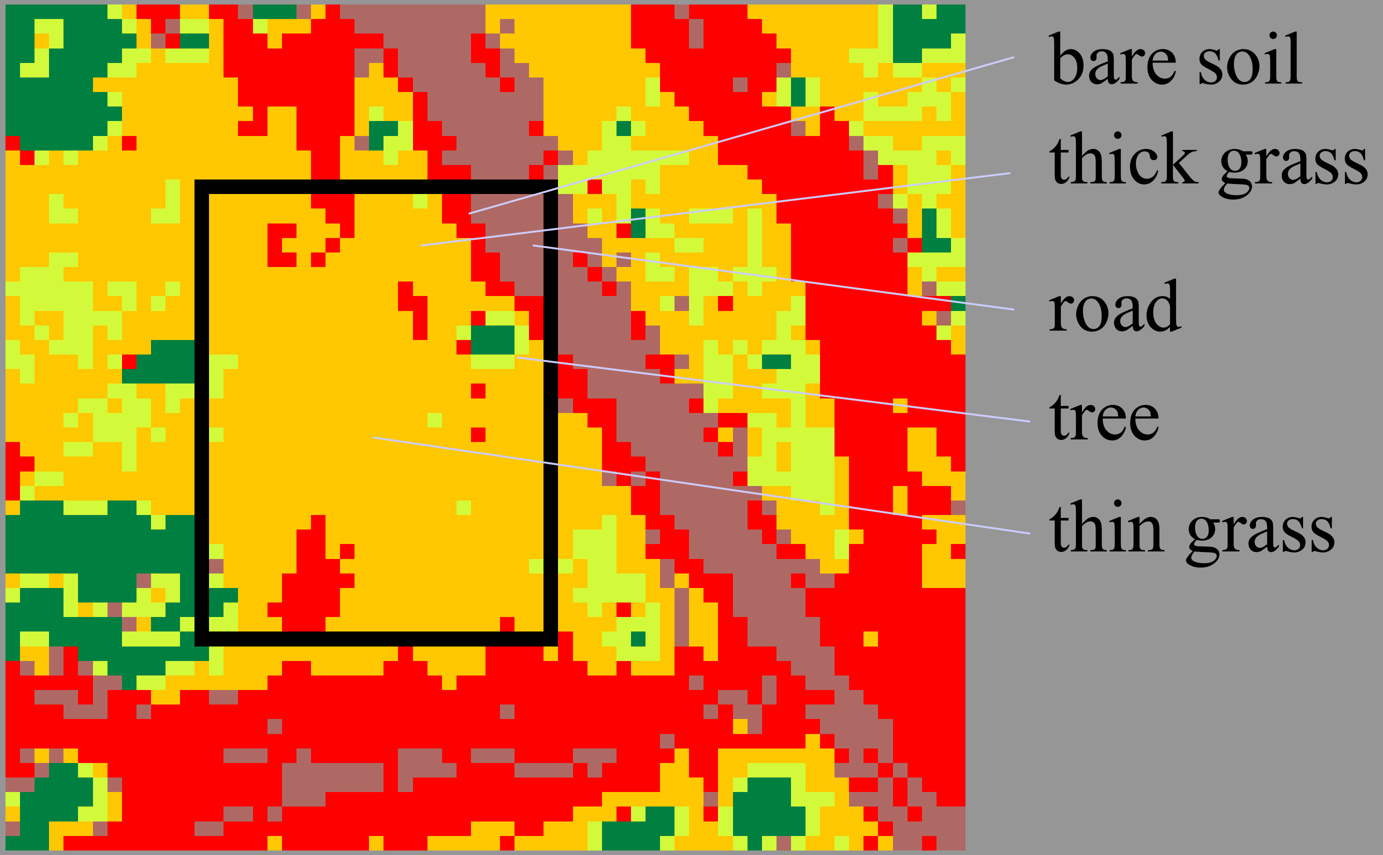
# “Thin Grass” Training Region Detail




Assigned training regions may not be “pure”, leading to system accuracy underestimation.

WAY-2C

# “Thin Grass” Training Region Classification




System may correctly classify but assumes it’s wrong, thus underestimating its own accuracy.

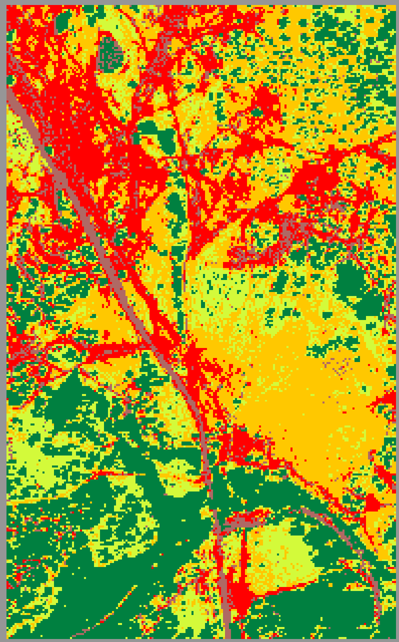
WAY-2C

# Locating the Most Relevant Identification Information in Spectral Cubes and Similar Large Data Sets

## End

Technology Available
US Patent 8,918,347

Robert K. McConnell
WAY-2C
rkm@way2c.com



**WAY-2C**