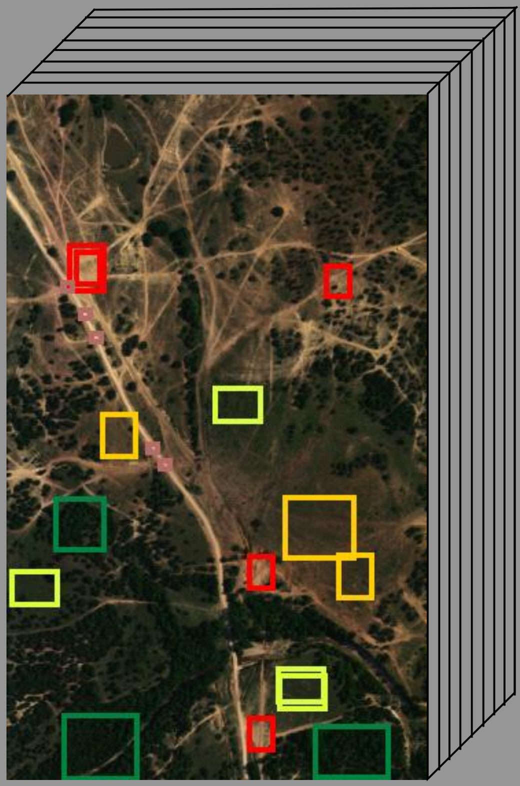
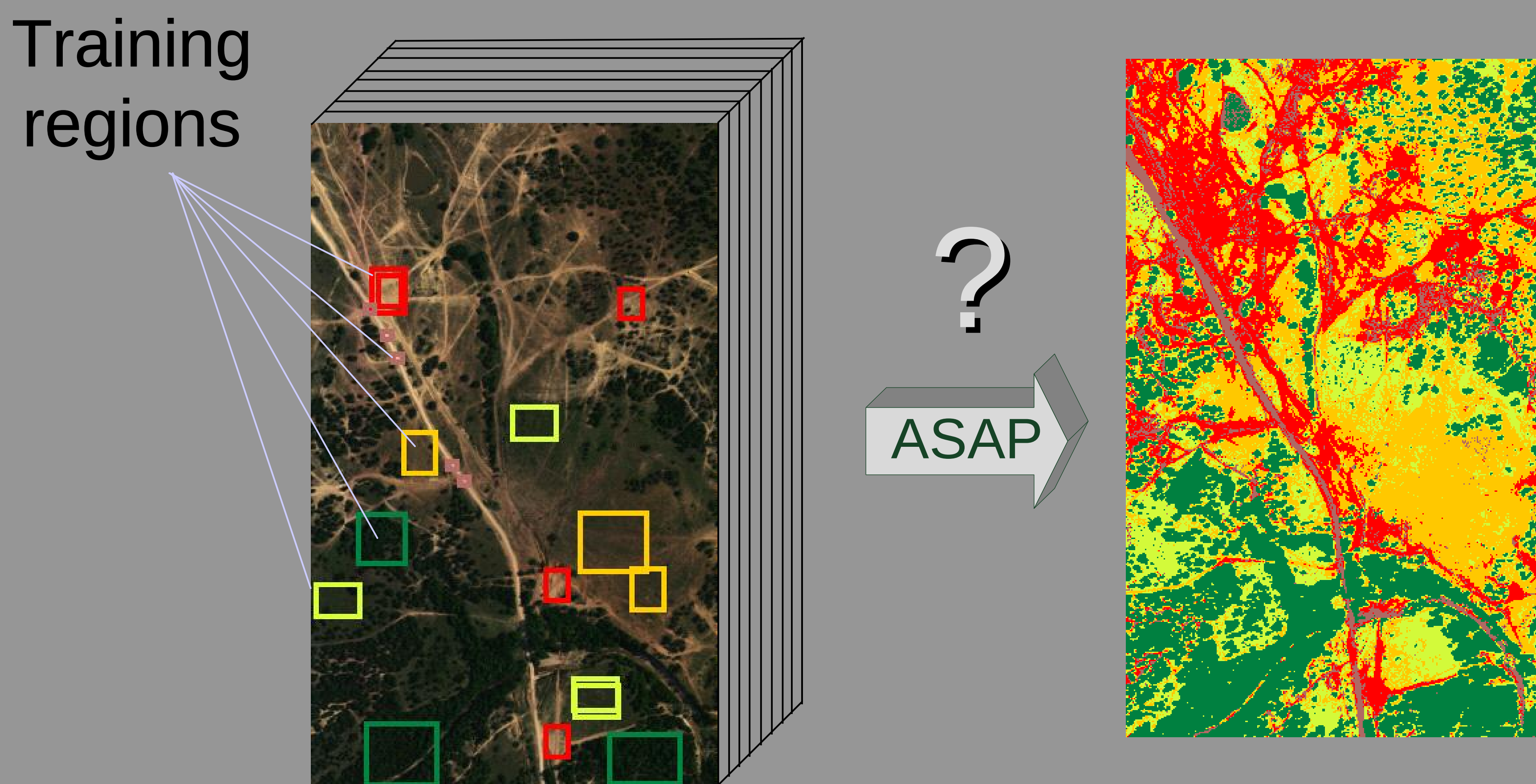
A New Method for Supervised Optimum Multispectral and Hyperspectral Data Subset Selection



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

WAY-2C

Classification Objectives



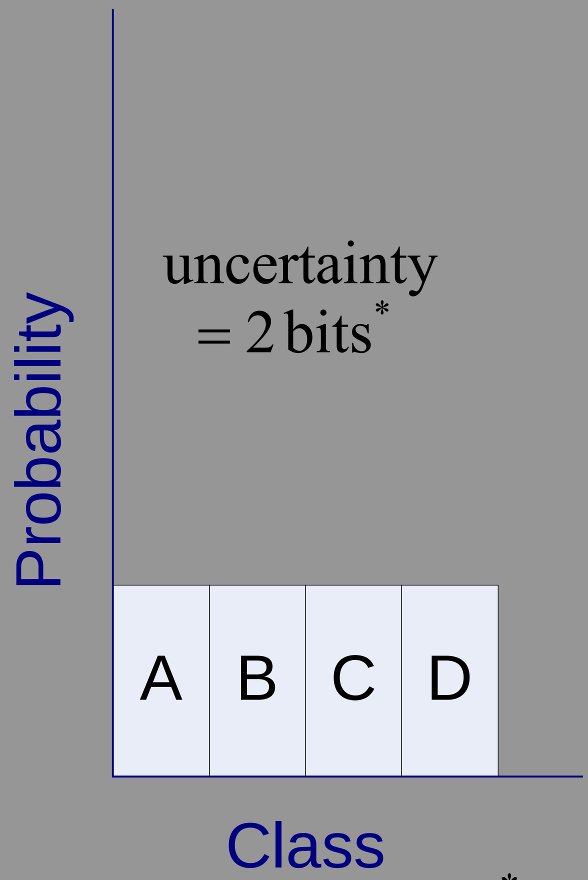
>1K bits/pixel

<5 bits/pixel

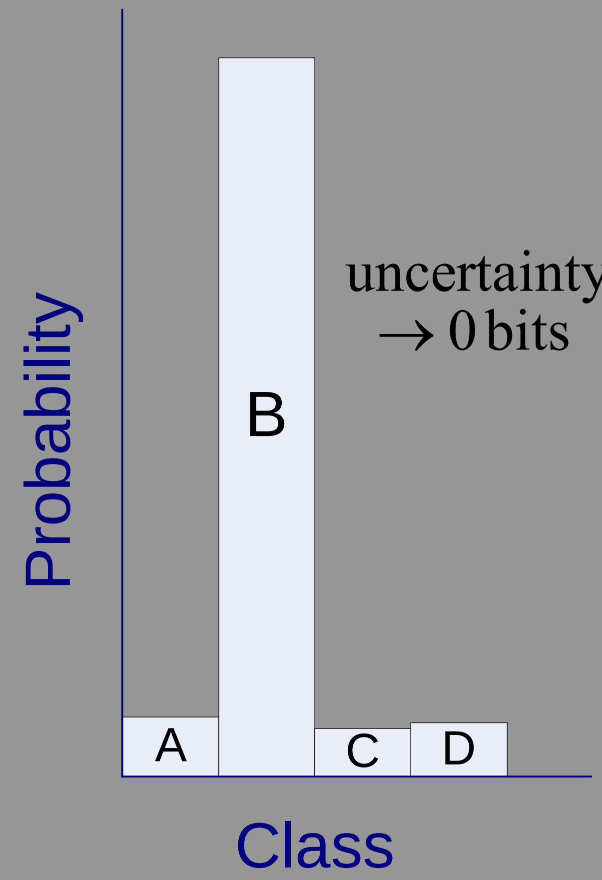
"In mind's eye"

Classification→

Before



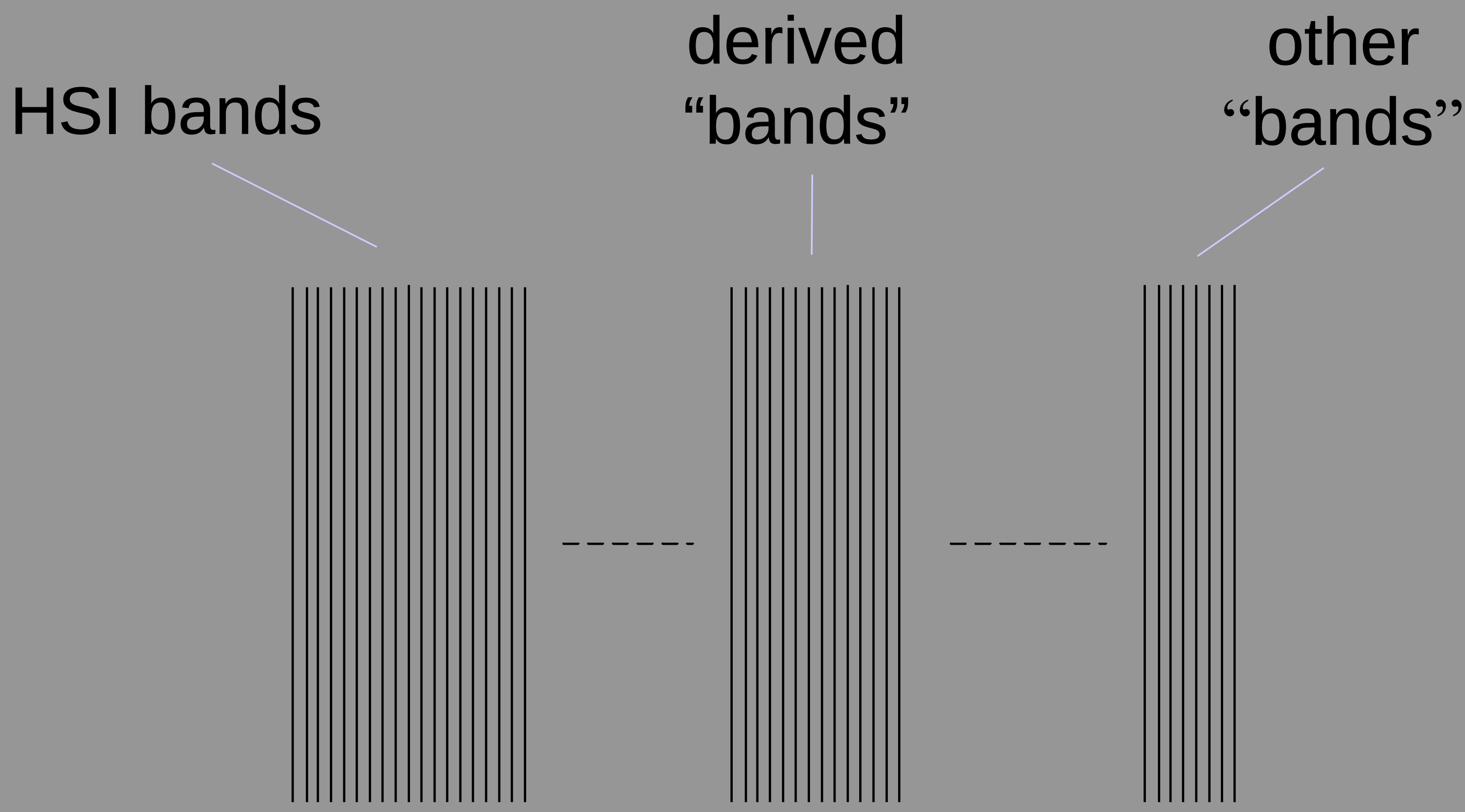
After



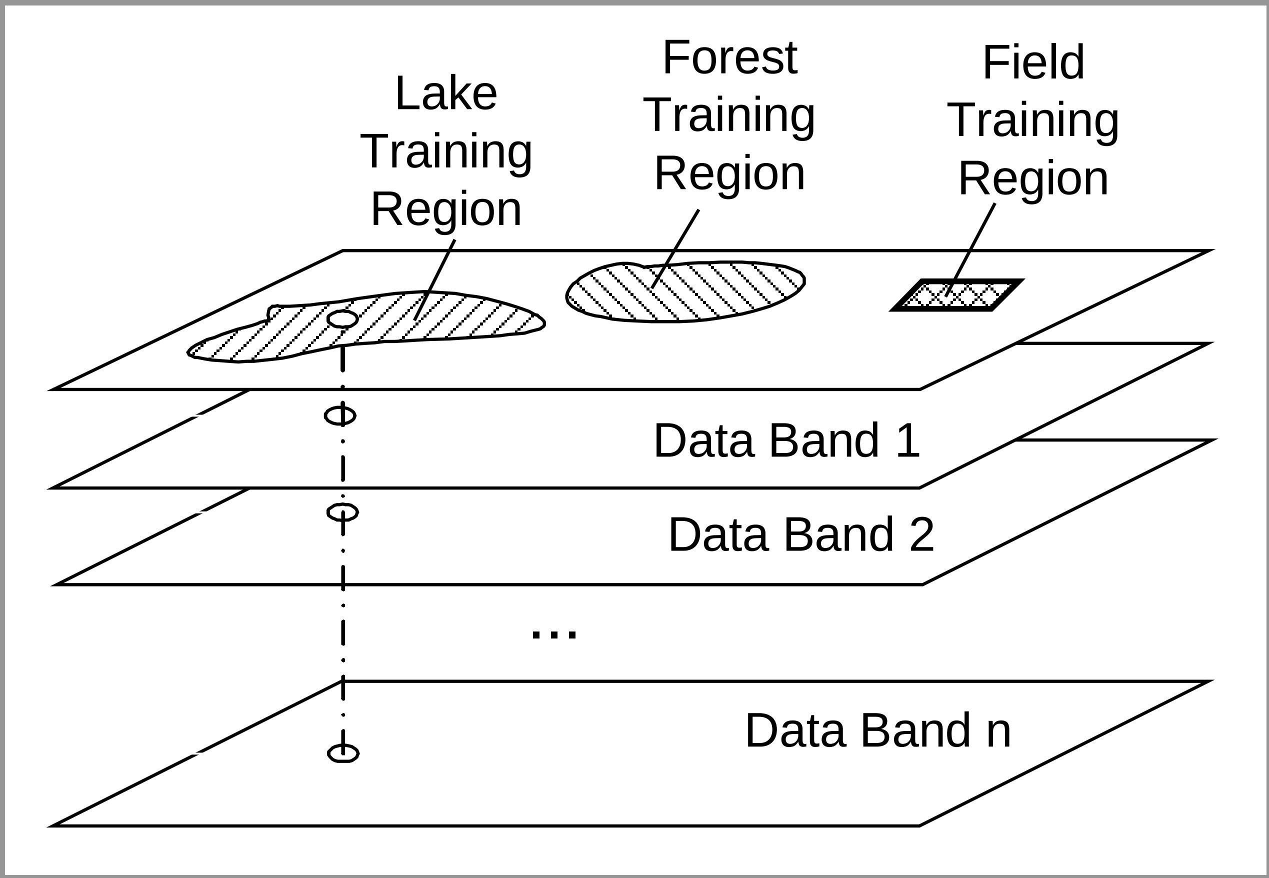
$$^*u = -\sum_{j}^{J} p_j \log_2(p_j)$$

WAY-2C

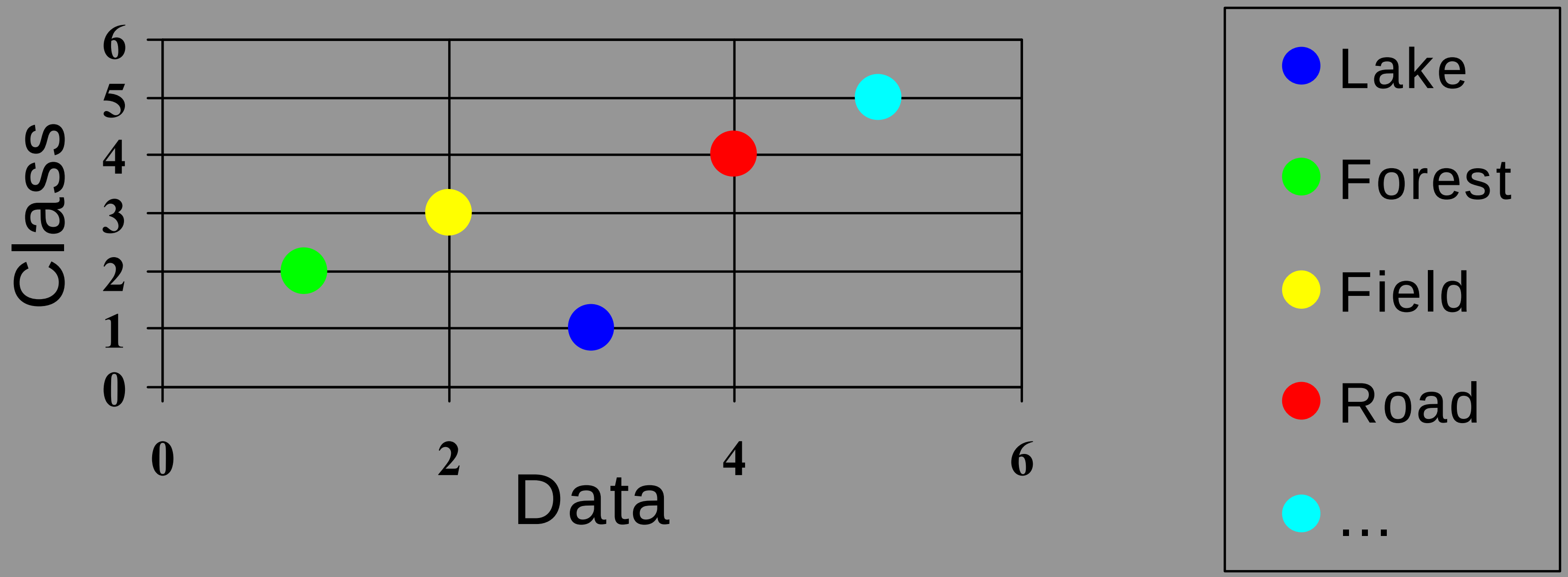
acquired bands and more



Given a training cube ...



... ideally find data subset that correlates one-to-one with class



WAY-2C

Outline

Using a simple example explore one way to:

- Find a most relevant data subset upon which to base a classification.
- Create an intermediate image using that most relevant subset.
- Perform that classification.

Finding a most relevant data set

treated strictly as a black box operation to automatically select out all data but that most relevant for differentiating the specified classes of interest.

Data may be discarded because the information it contains is either irrelevant or redundant.

A simple example: optimum band set for green “ paint chip” differentiation

"True Color" Image From Cube

	Band	nm
R	40	642
G	25	550
B	10	458

"Green Chips"

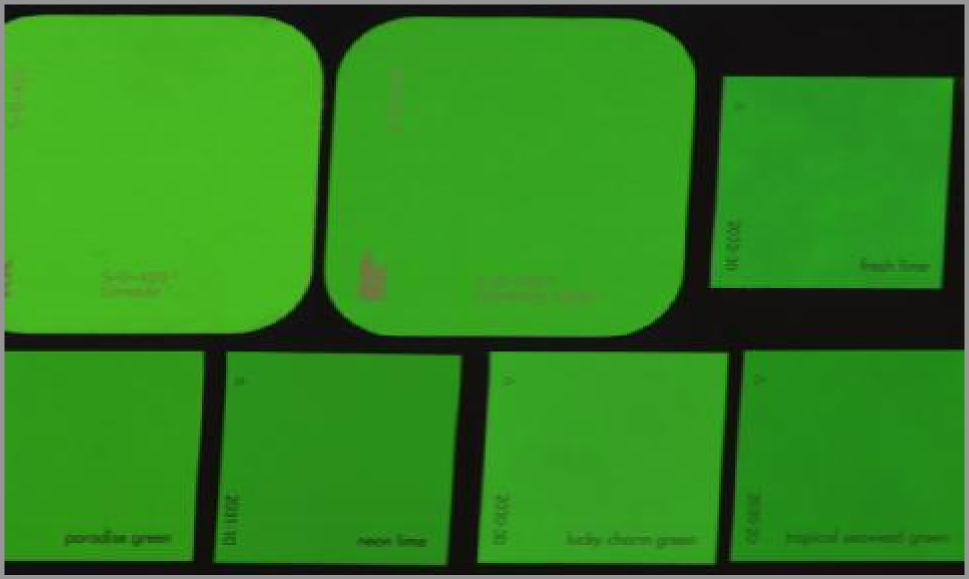


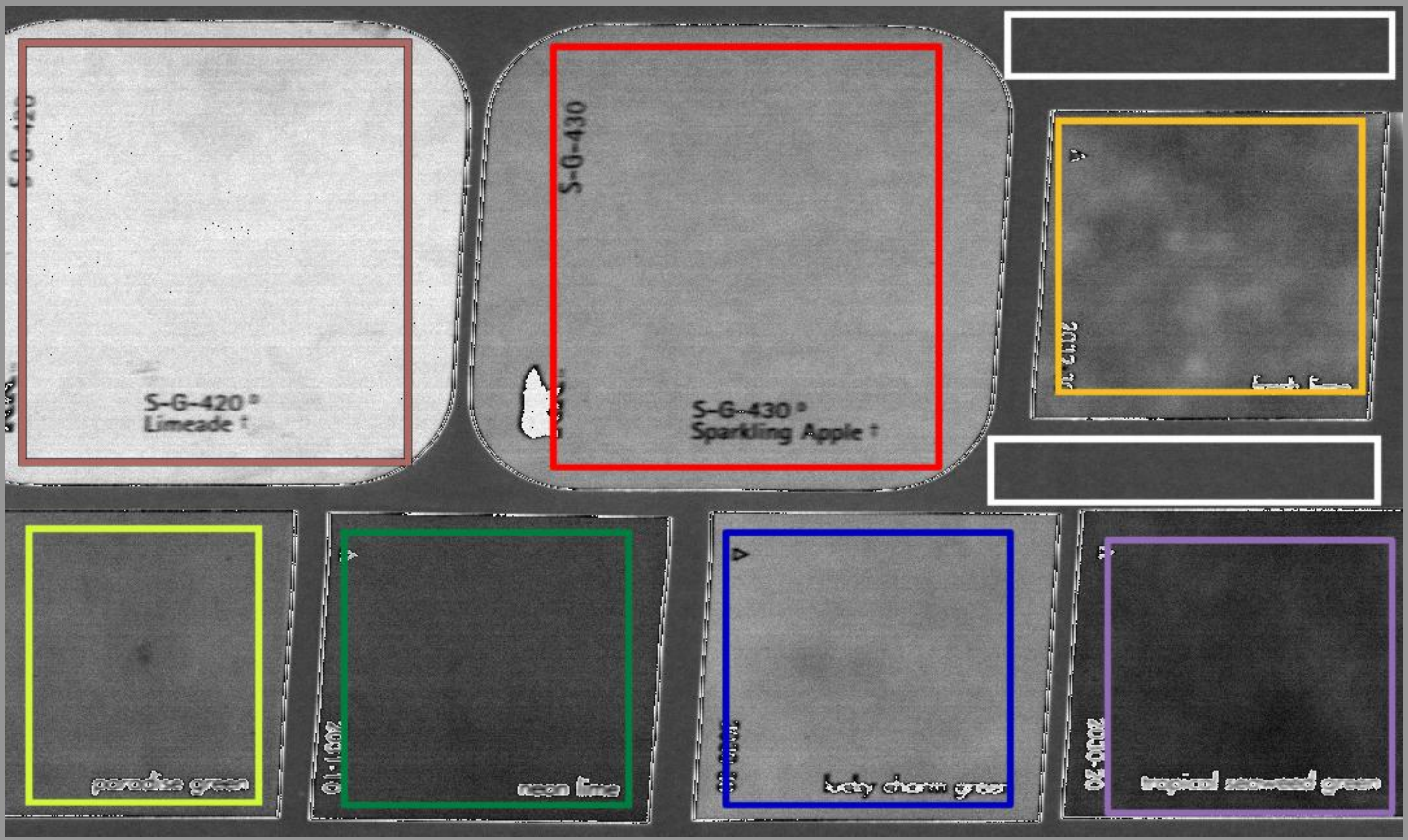
Image Courtesy Resonon, Inc.

8 Classes

Required:

Band combination to differentiate reliably.

"Green Chip" Training Regions



WAY-2C

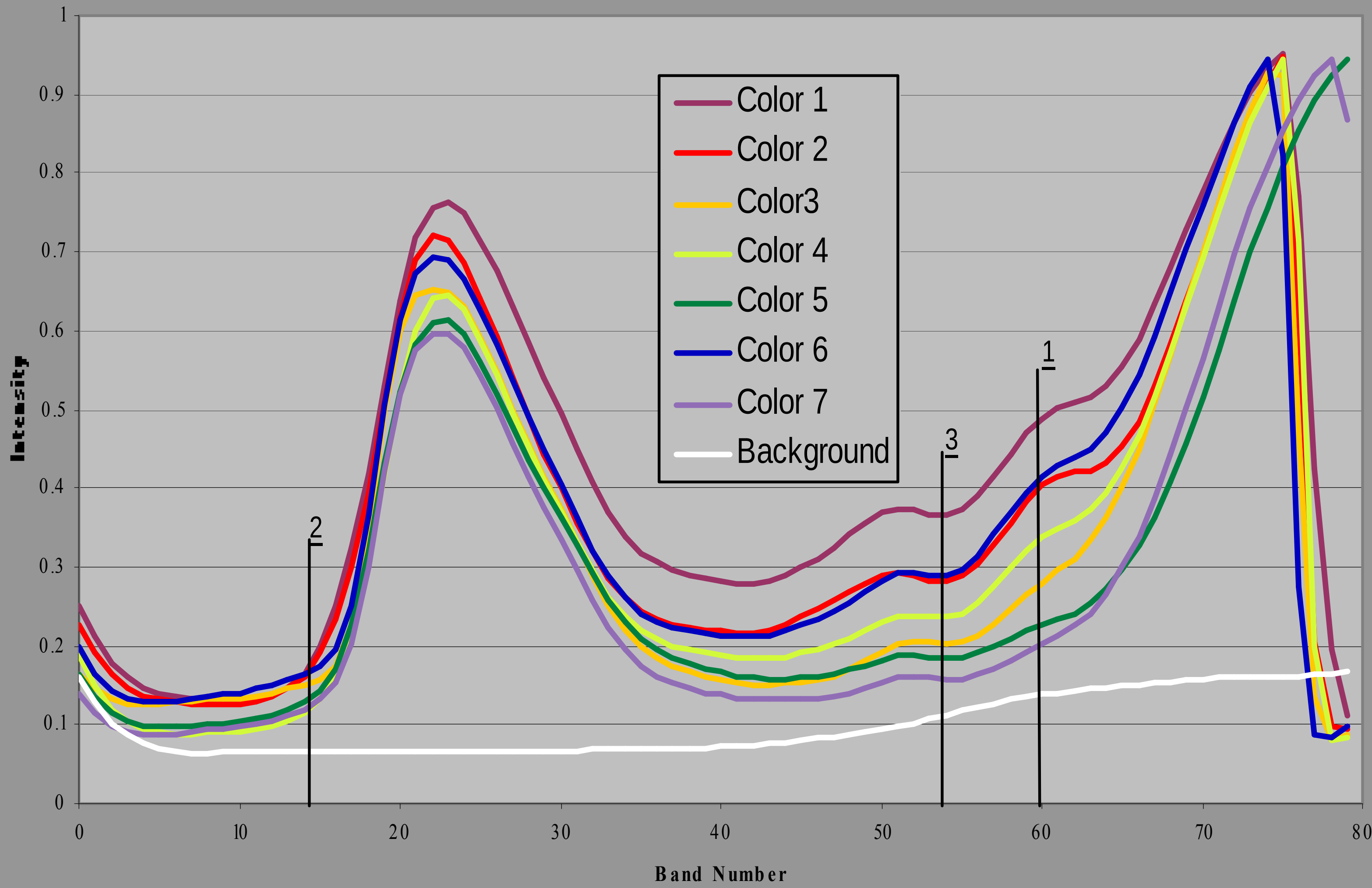
Base Image Courtesy Resonon, Inc.

Data Processing for Image Interpretation

Traditional Approach:

- Compare spectrum of each pixel with mean spectra from classes of interest to find best fit.

Traditional: Compare Pixel to Mean Observed Spectra



Data Processing for Image Interpretation

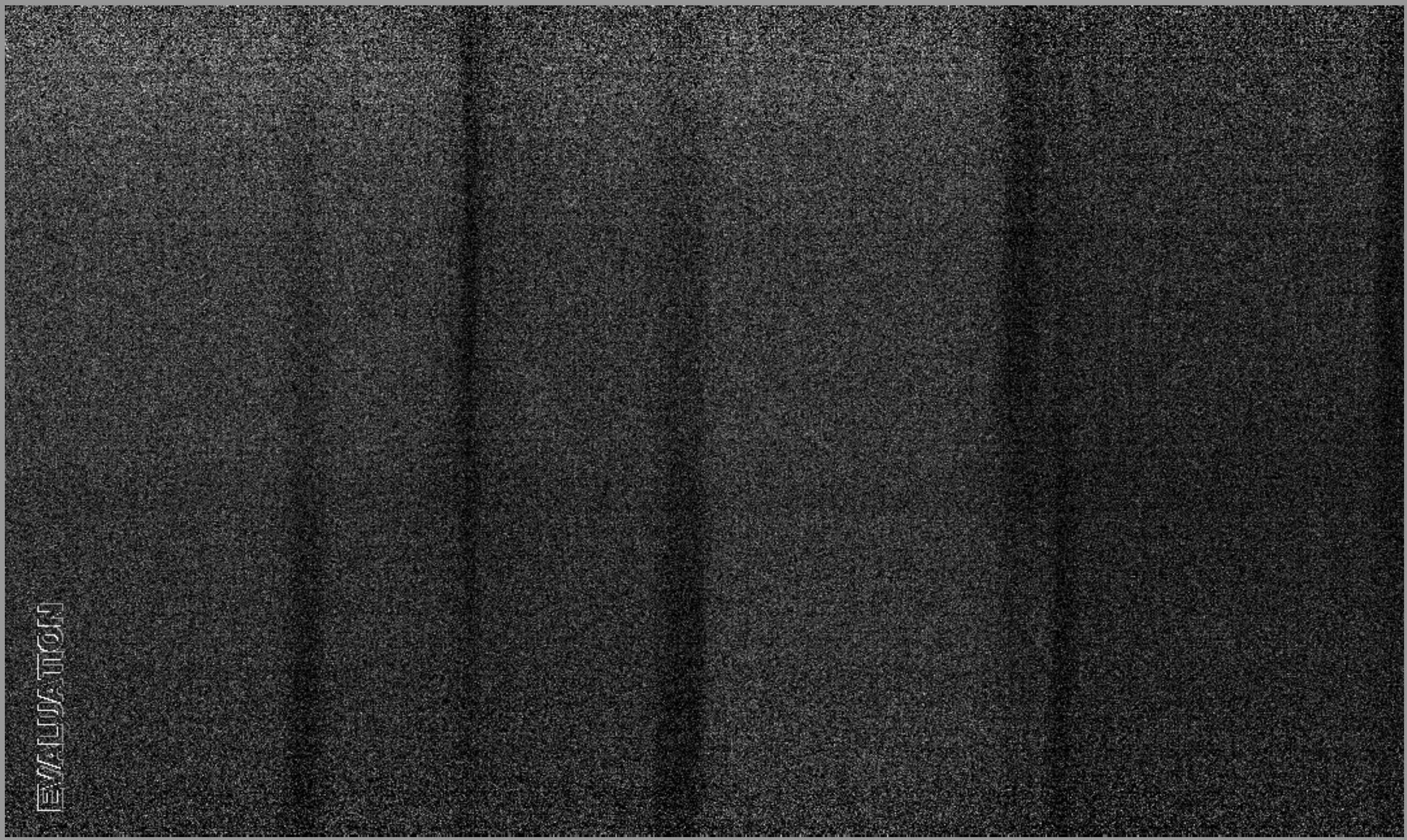
Traditional Approach:

- Compare spectrum of each pixel with spectra from known classes to find best fit.

Problem:

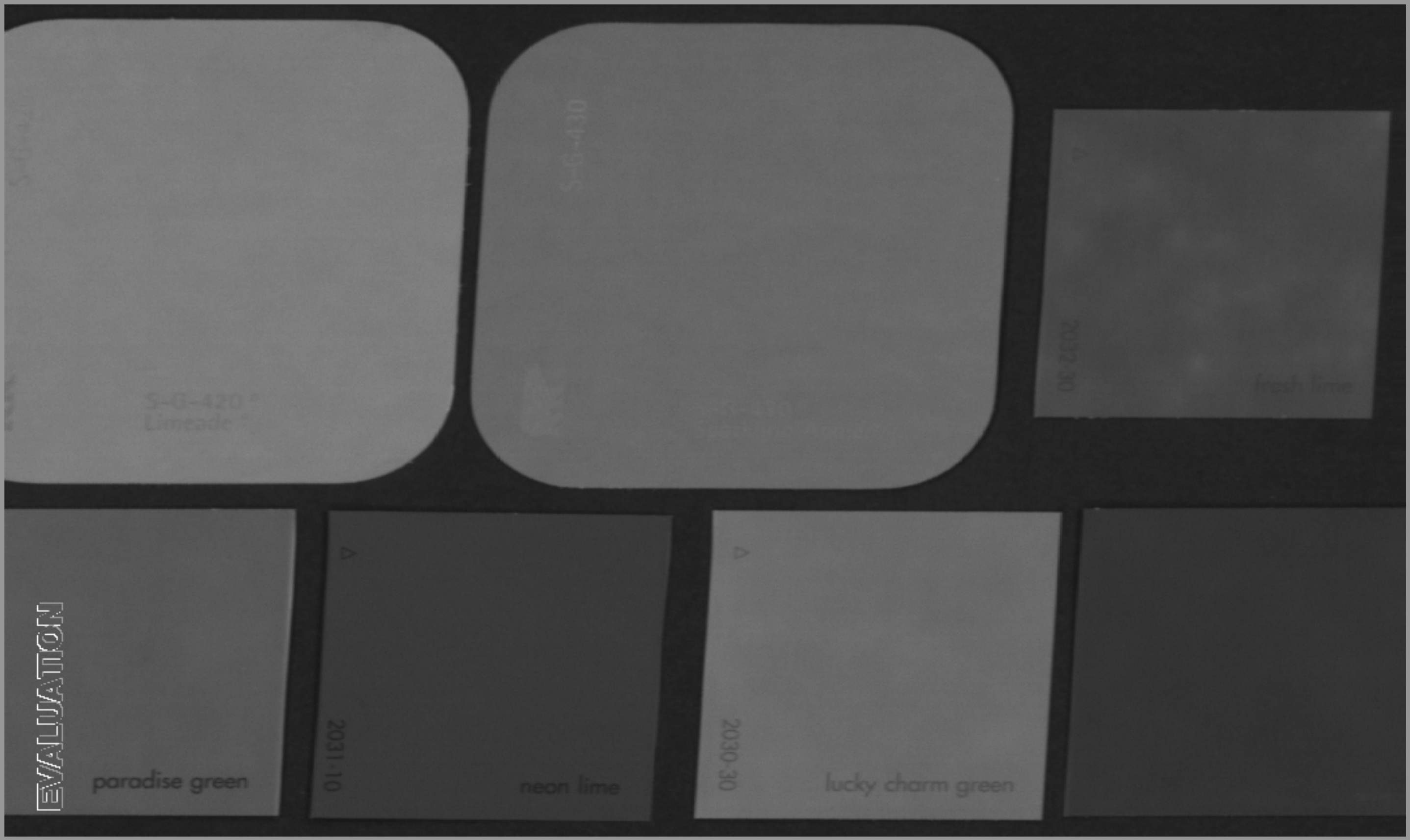
- Not all data bands are equally relevant for differentiation of specified classes.

Band 00 PSPA=14.01%



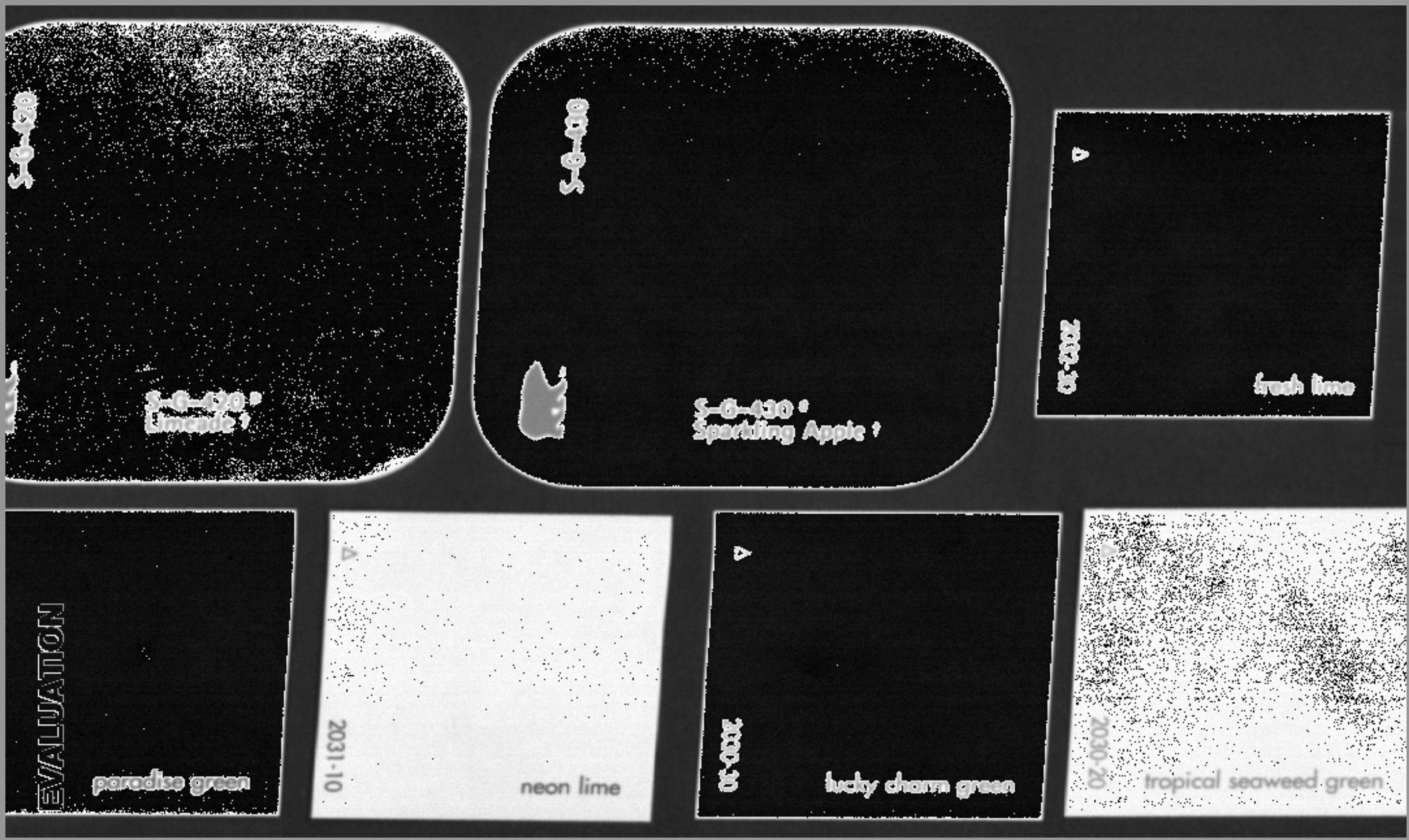
Base Image Courtesy Resonon, Inc.

Band 60 PSPA=67.88%



WAY-2C *Base Image Courtesy Resonon, Inc.*

Band 79 PSPA=27.19%



WAY-2C *Base Image Courtesy Resonon, Inc.*

Data Processing for Image Interpretation

Traditional Approach:

- Compare spectrum of each pixel with spectra from known classes to find best fit.

A newer (and evolving) approach:

- Select combination of only most relevant data to compare with similar data from known classes.

Processing Steps

Automated methods to:

1. Scan an entire training cube to find the most relevant bands and window positions within those bands upon which to base a classification.
2. Construct intermediate image based on the selected information.
3. Build reference “template” for each class based on intermediate image probability distributions.
4. Use reference templates to interpret current and other intermediate images.

Information Theory

Using histograms, allows us to treat uncertainty much like a physical quantity such as mass or charge.

Relevant information reduces uncertainty.

Data Entropy (uncertainty)

Consider any subset D of the available data as a collection of symbols, one for each location of interest in the training regions:

$$D(d_1, d_2, ..., d_k, ...) \quad (1)$$

the average amount of information necessary to specify the data value at a location is the data entropy

$$\mathrm{H}(D) = -\sum_k \mathrm{p}(d_k)\log_b(\mathrm{p}(d_k)) \quad (2)$$

where: $\mathrm{p}(d_k)$ is the frequency of occurrence of value d_k.

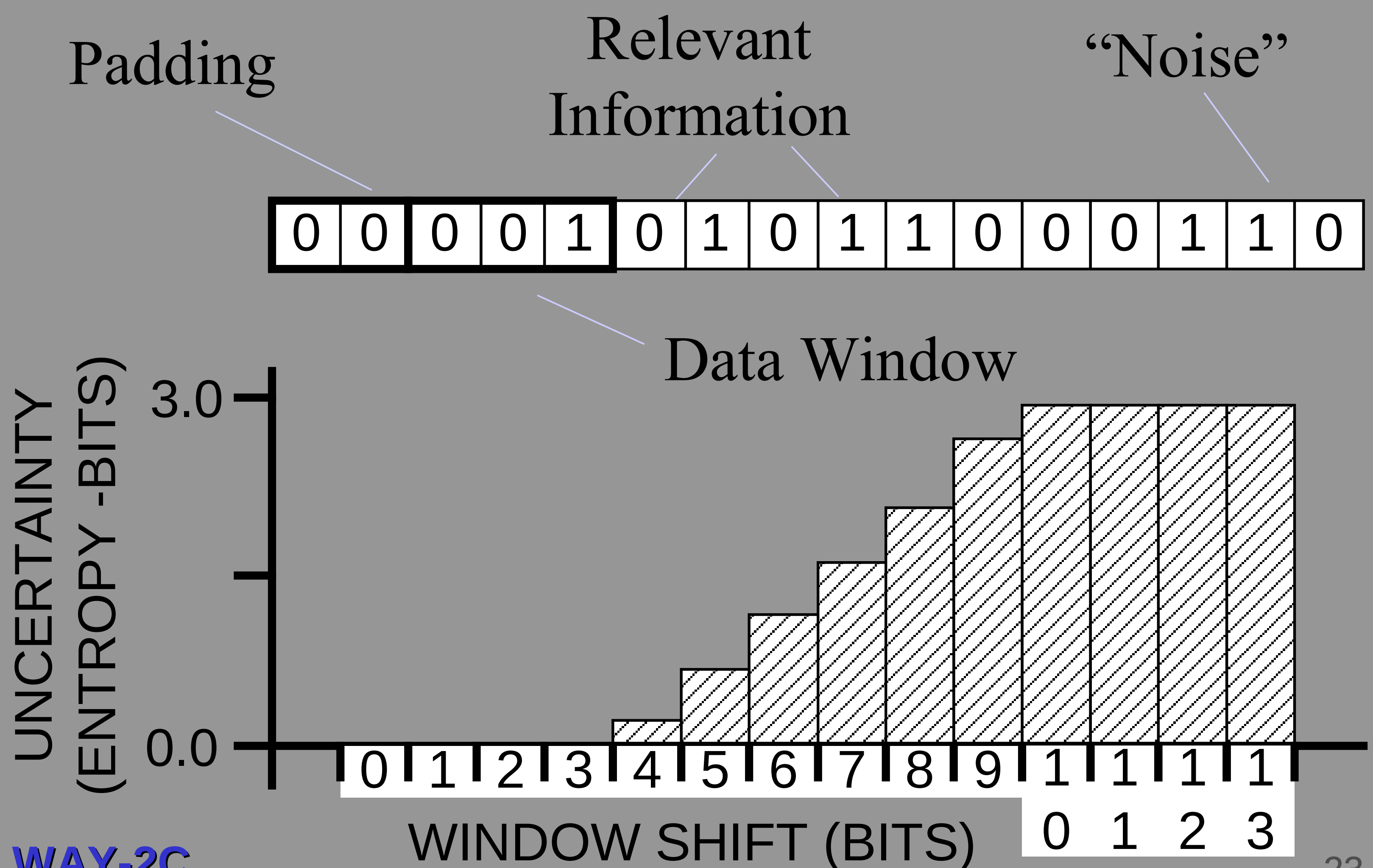
Data Entropy (uncertainty)

If the base of the logarithms b is 2 then

$$\mathrm{H}(D) = -\sum_k \mathrm{p}(d_k)\log_2(\mathrm{p}(d_k))$$

has units of bits.

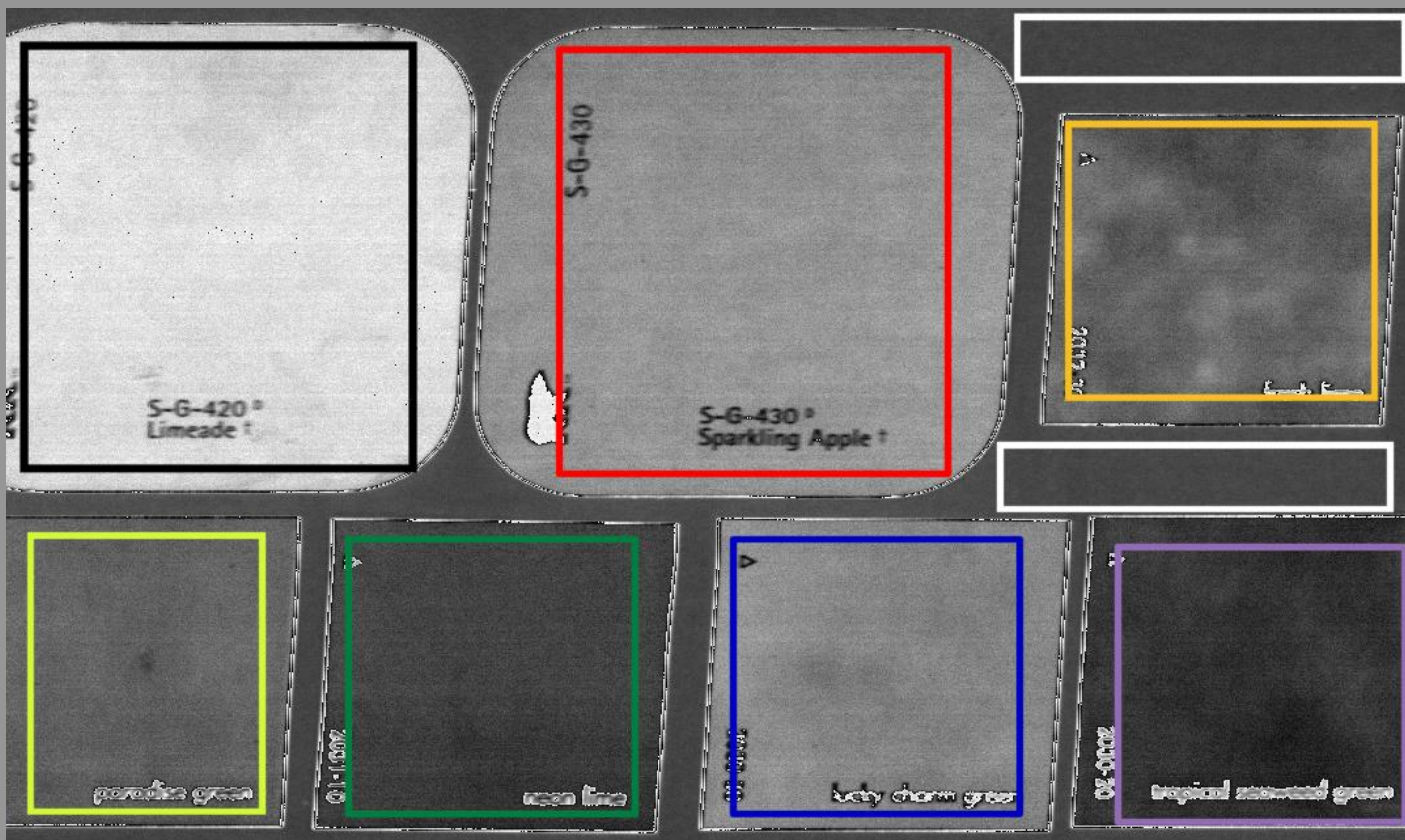
Most Relevant Information: Data Window Uncertainty



WAY-2C

Measuring Relevance

Measuring Relevance (1)



WAY-2C

Base Image Courtesy Resonon, Inc.

Measuring Relevance (2)

Consider a set of training regions in an image, where both class and data values are known.

For the set of classes

$$C(c_1, c_2, ..., c_j, ...)$$

and the set of data

$$D(d_1, d_2, ..., d_k, ...)$$

Measuring Relevance (3)

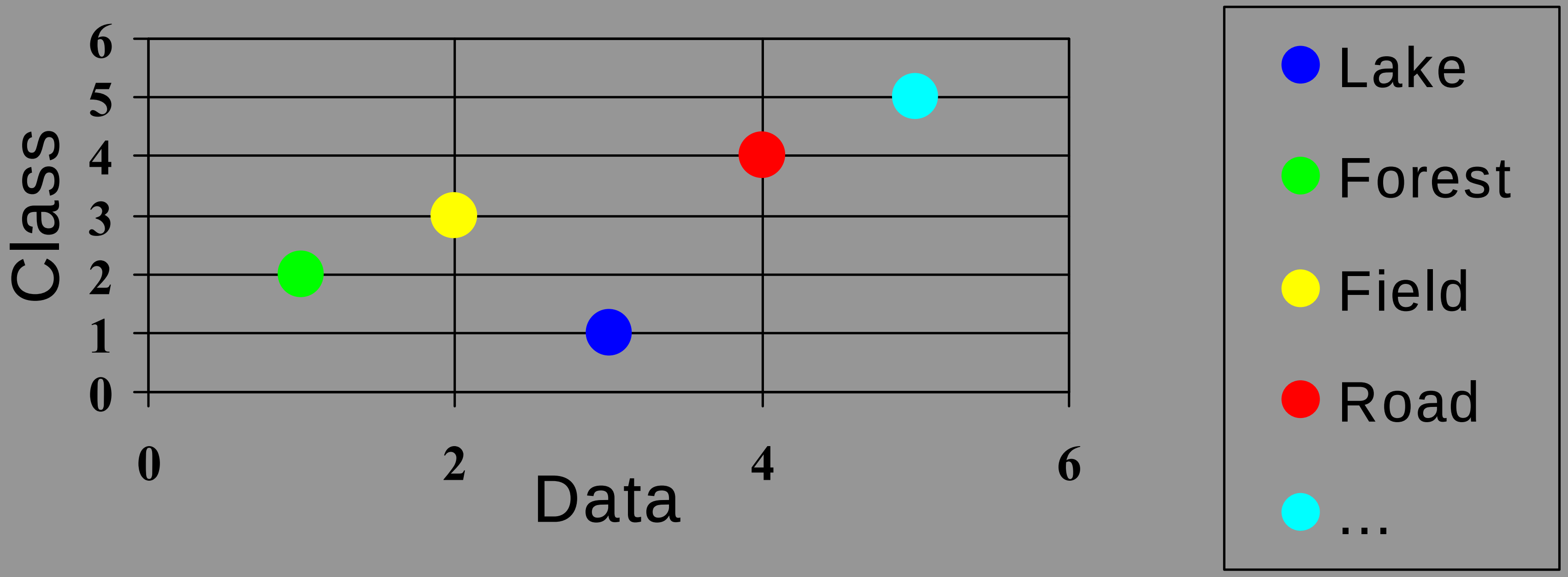
It can be shown that the relevant information in *D*, also known as the mutual information, is given by

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

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 $C\&D$.

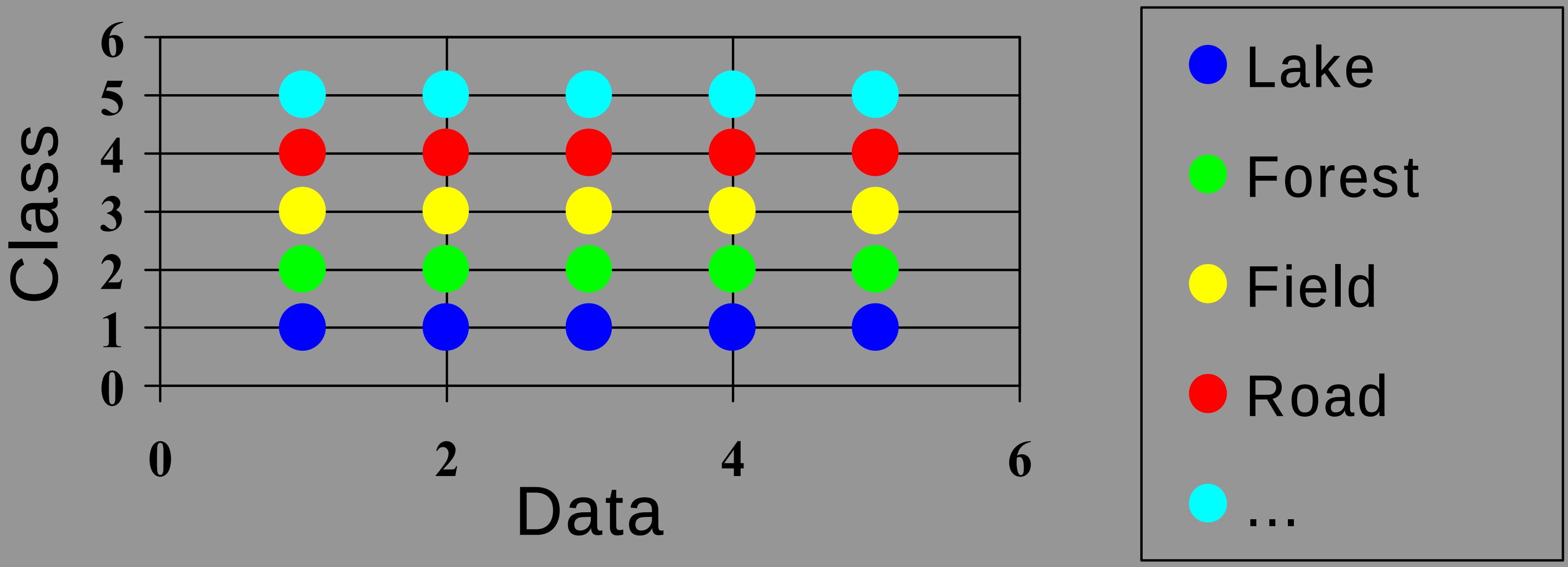
If one-to-one correlation between class and data...



$$\boxed{H(C\&D) = H(C) = H(D)} \qquad (4)$$

WAY-2C

If no correlation between class and data …



$$\boxed{H(C\,\&\,D) = H(C) + H(D)} \qquad (5)$$

WAY-2C

uncertainty

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 $\boldsymbol{D}$.

Predicted Agreement

We define the *PSPA*, or predicted single pixel agreement by:

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

Finding Most Relevant Data and Constructing Intermediate Image

Most relevant window: measured single band information contribution.

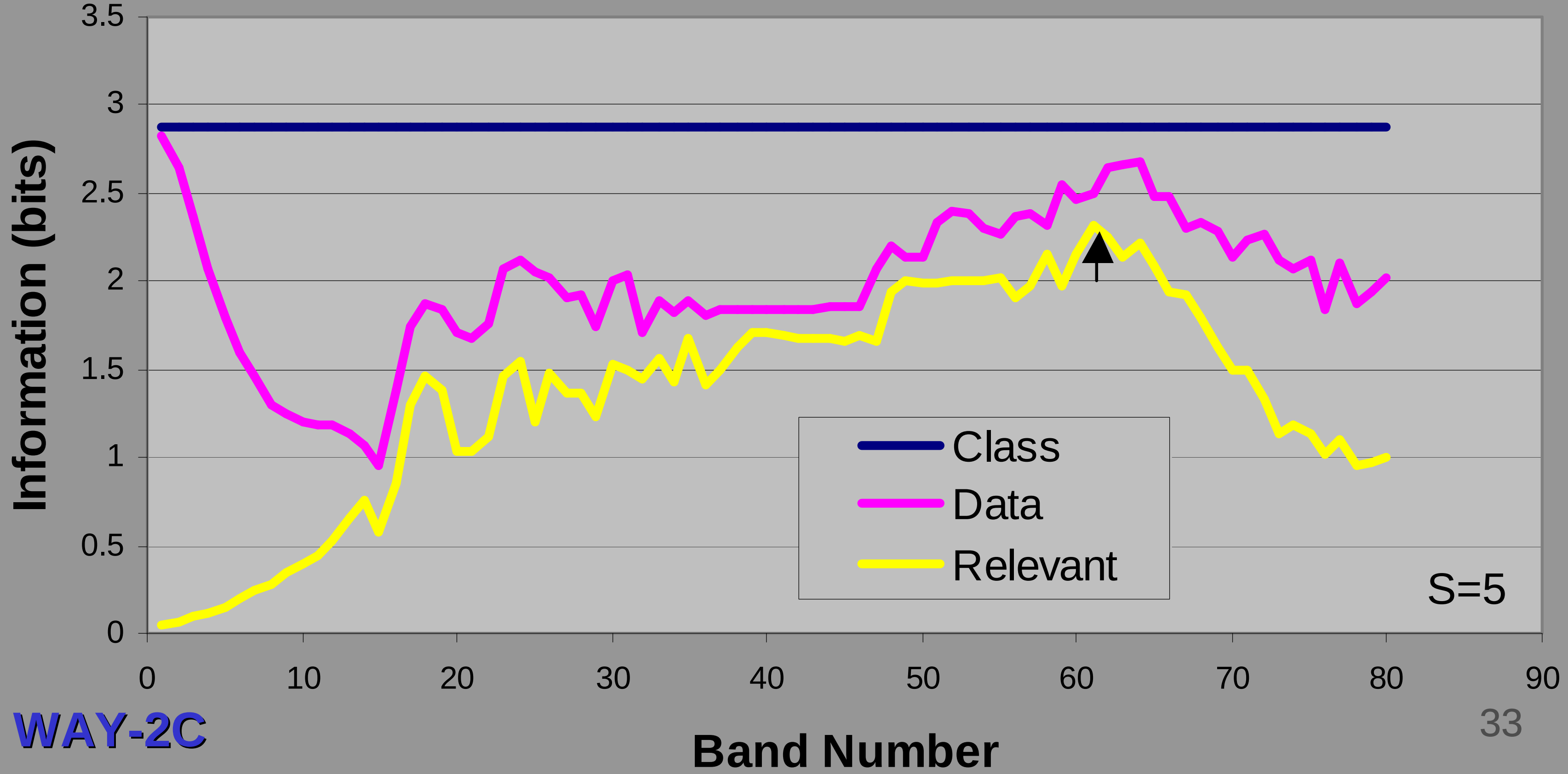
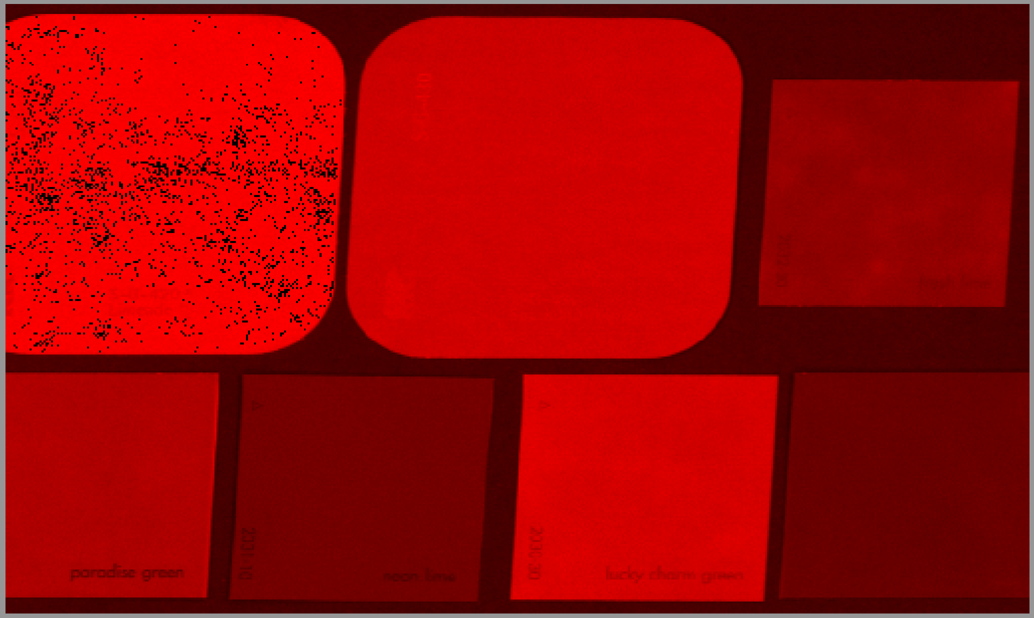


Image constructed from single most relevant band and window only.

Plane	Band	nm
R	60	764

rel = 2.318 bits

PSPA = 67.88%



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

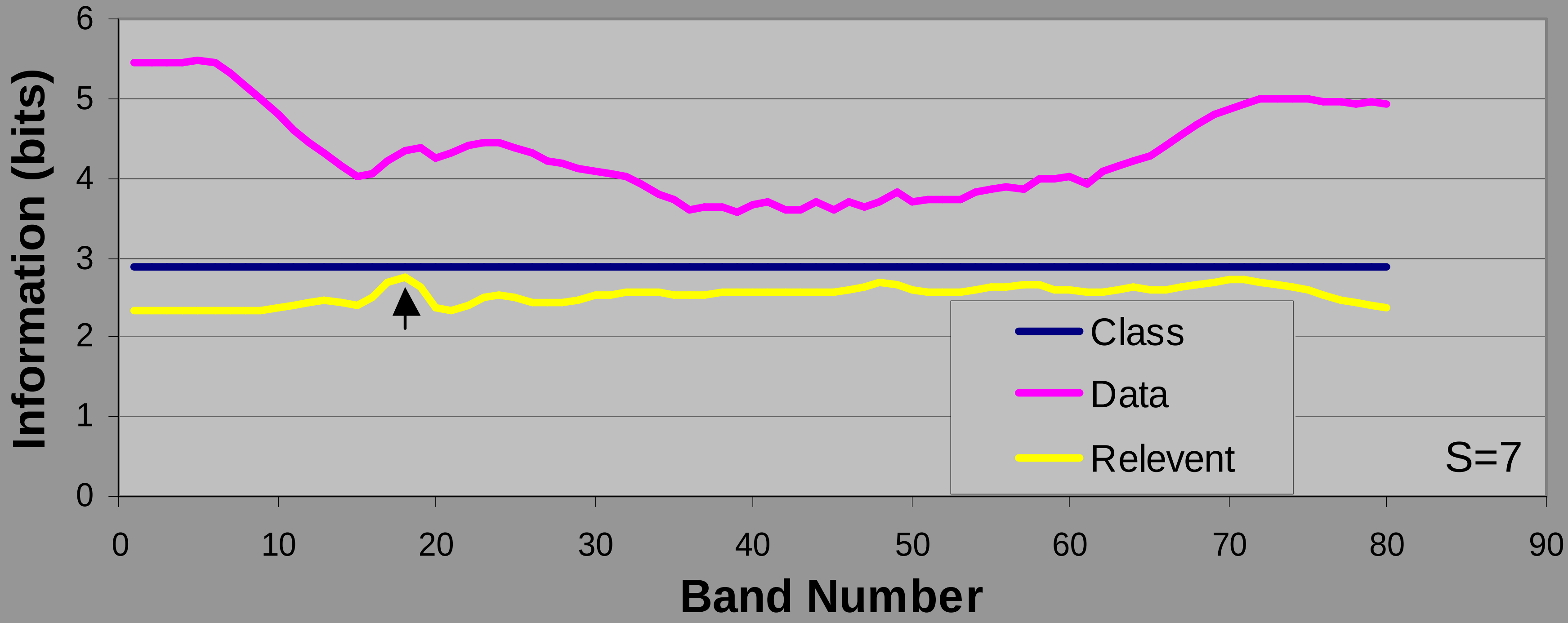
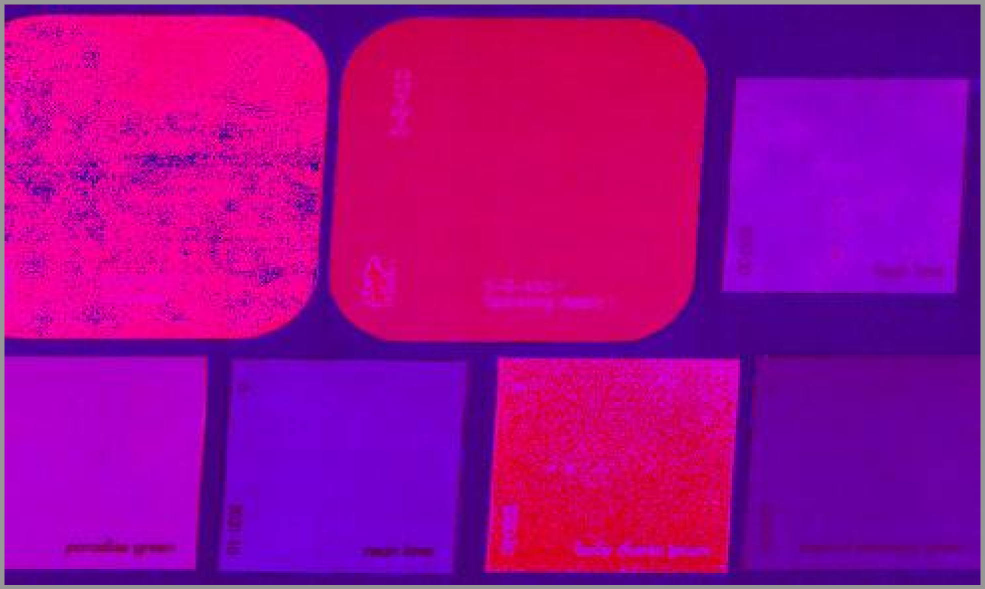


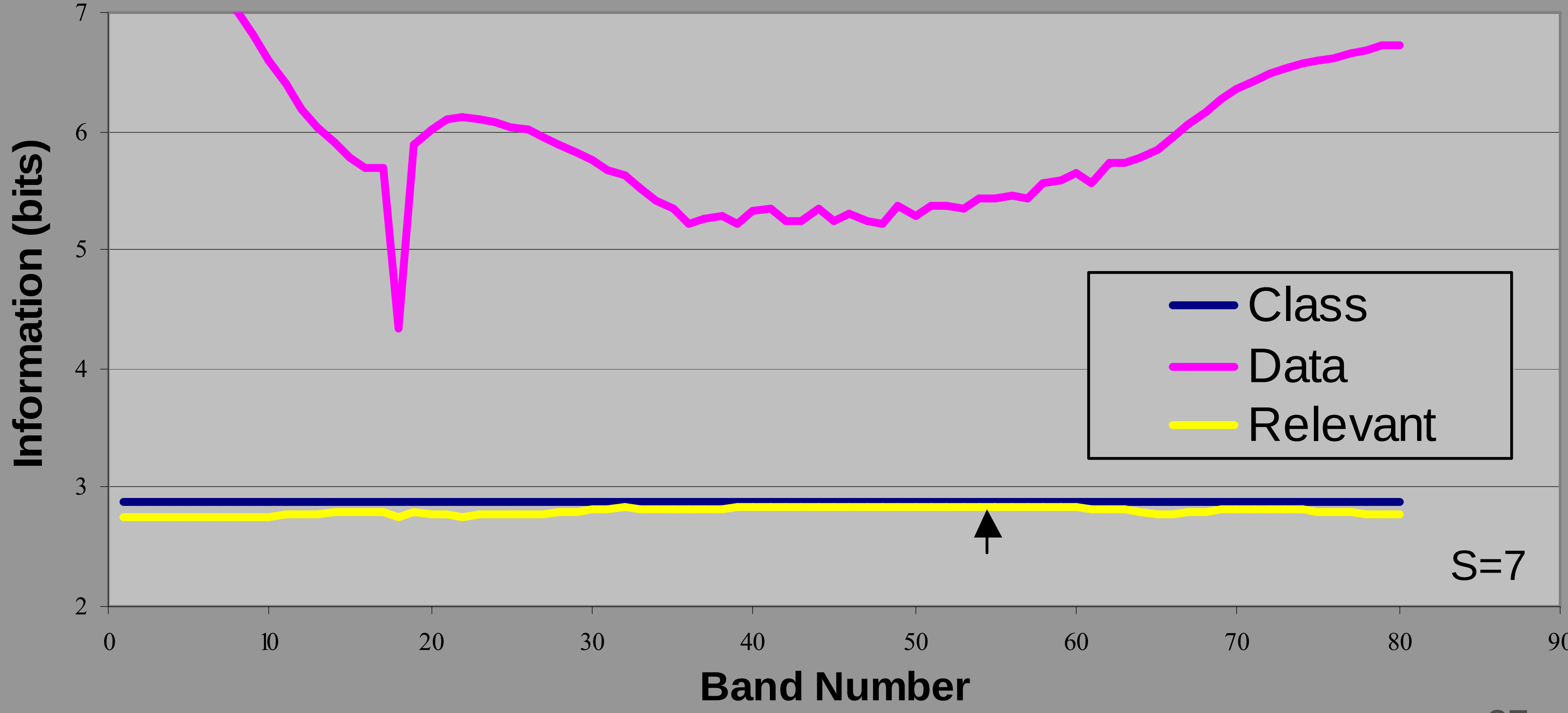
Image constructed from most relevant band pair.

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)

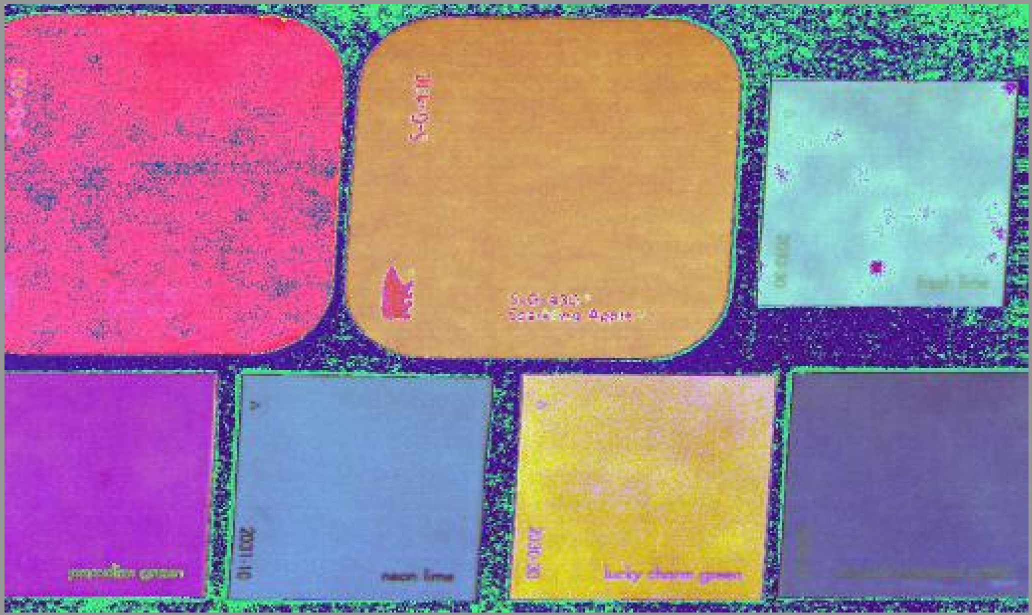


WAY-2C

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

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

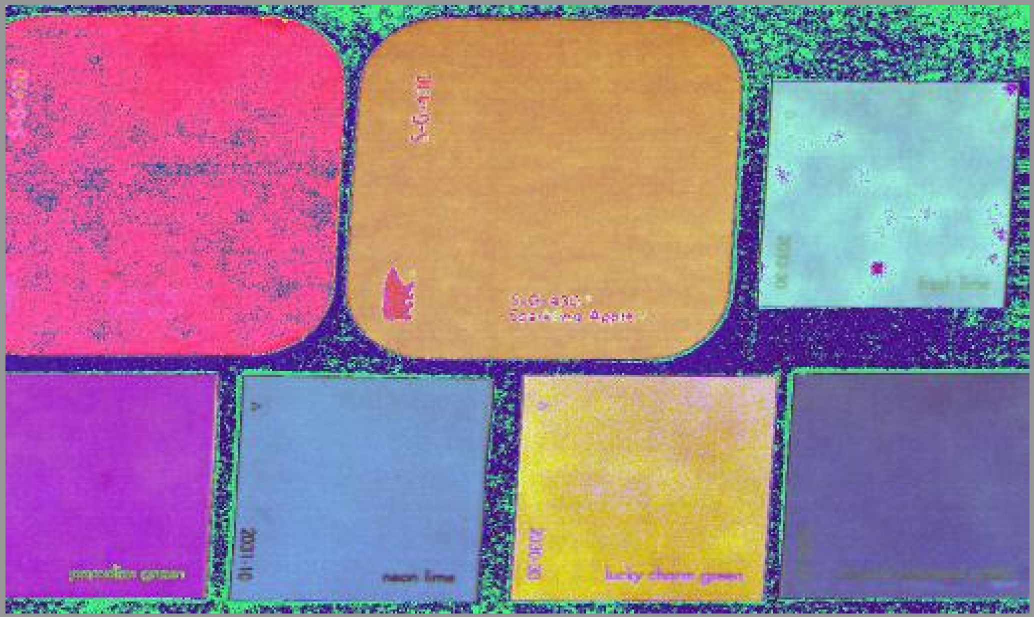
rel = 2.844 bits
PSPA = 97.74%



>95% band count reduction from original hyperspectral cube.

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

rel = 2.844 bits
PSPA = 97.74%



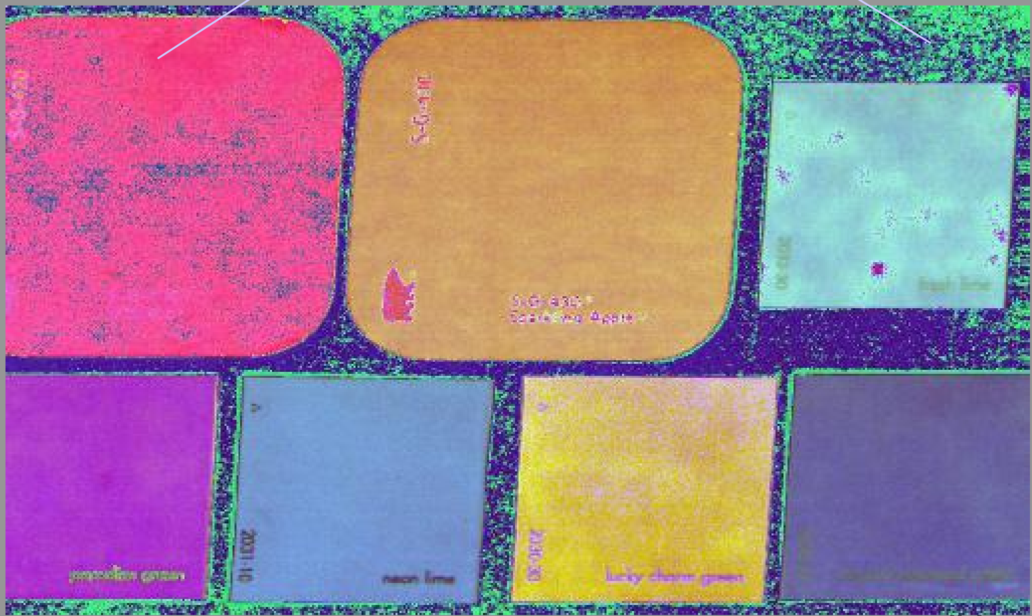
Training System To Recognize Classes of Interest

Building
Reference “Templates”

Maximum Relevance Intermediate Image

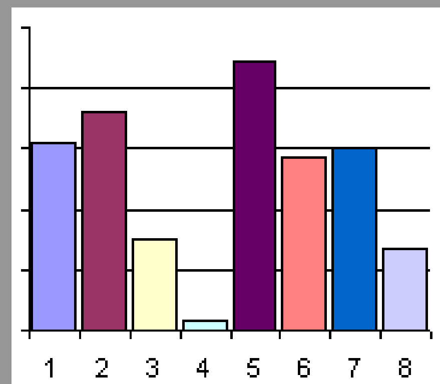
Multimodal color distributions
(Least squares class fitting won’t work)

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



Determining most likely source

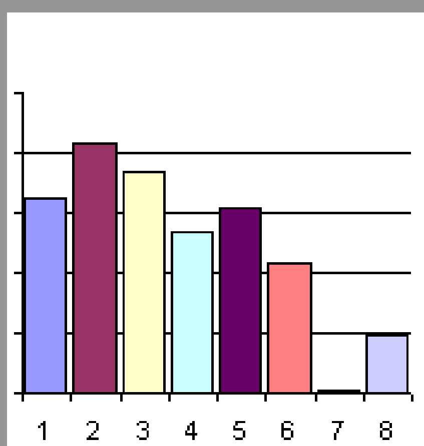
T



…

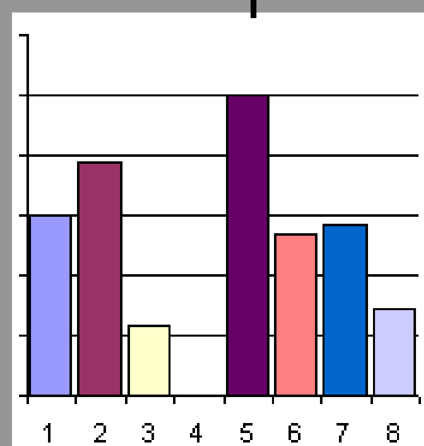
Most likely source NOT necessarily best T and R correlation!

? **?** **?**



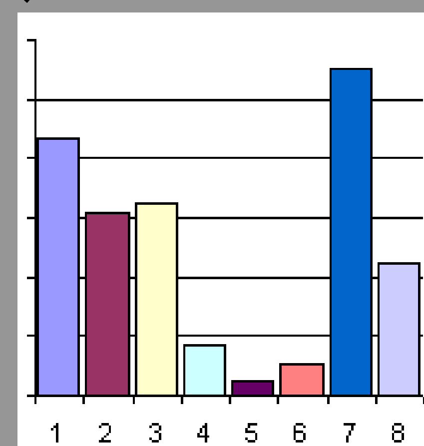
…

R(A)



…

R(B)



…

R(C)

…

WAY-2C

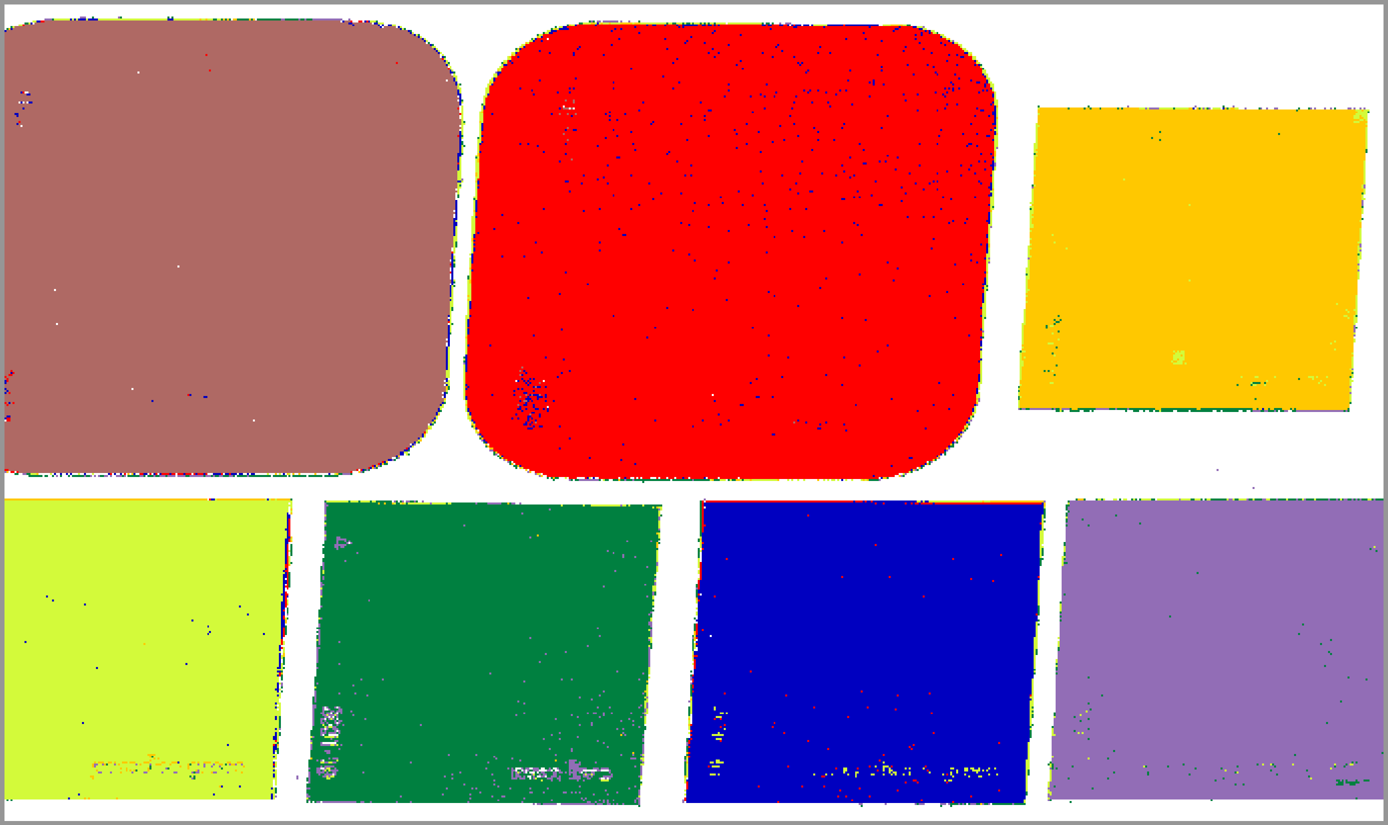
Relative source probability (class pair odds) assuming A and B equally likely

$$\frac{P(A|T)}{P(B|T)} = b^{N_T \sum_i^I p_{iT}(\log_b(p_{iA}) - \log_b(p_{iB}))}$$

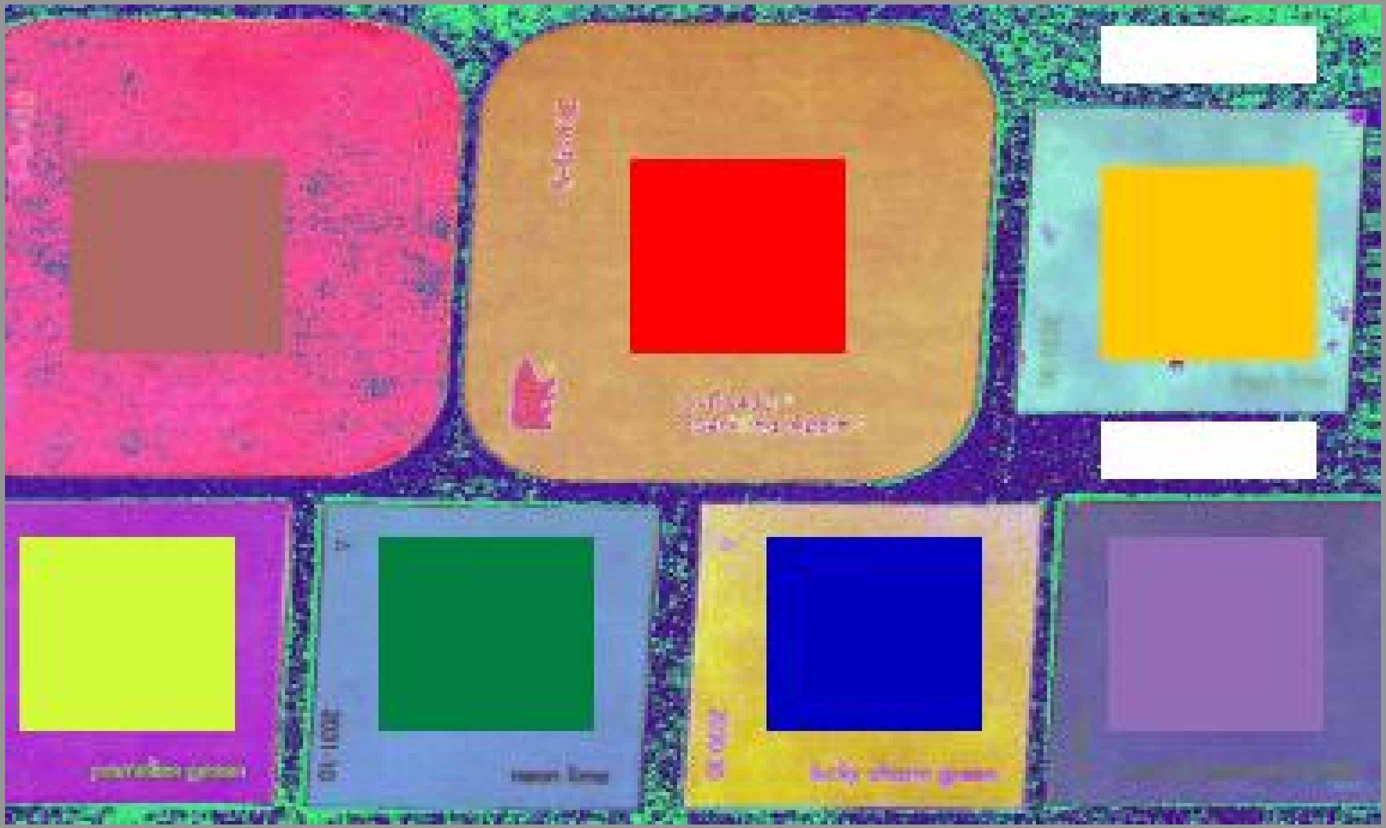
When the probabilities of all colors are known for each of a pair of possible classes, the likelihood of the correct class being determined increases exponentially with number of independent samples (e.g. pixels) in the histogram for the test region.

Interpreting Intermediate Image(s)

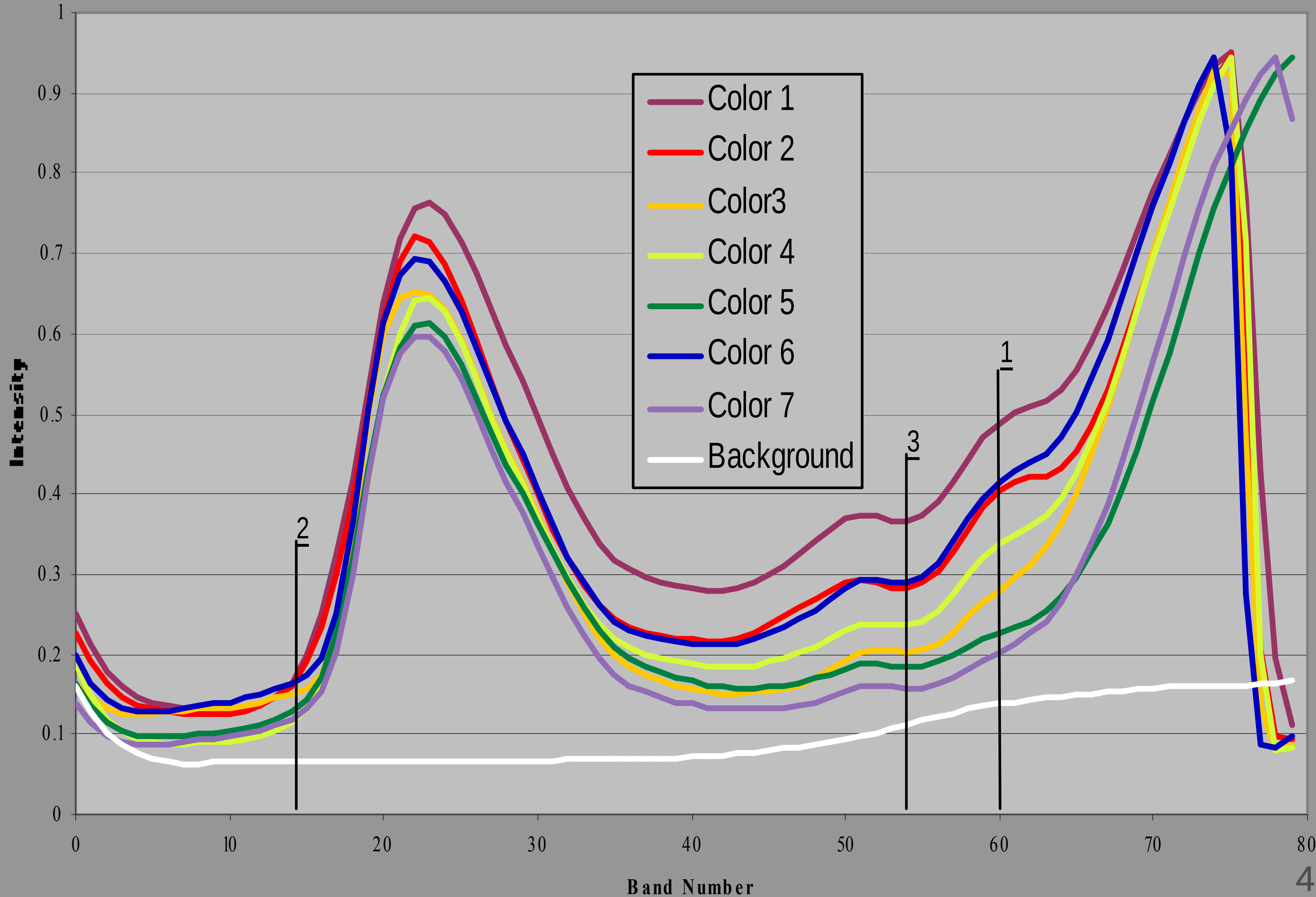
Pixel by pixel interpretation of most relevant band triplet image.



Classification of most relevant band triplet image (~30K pixel blocks).

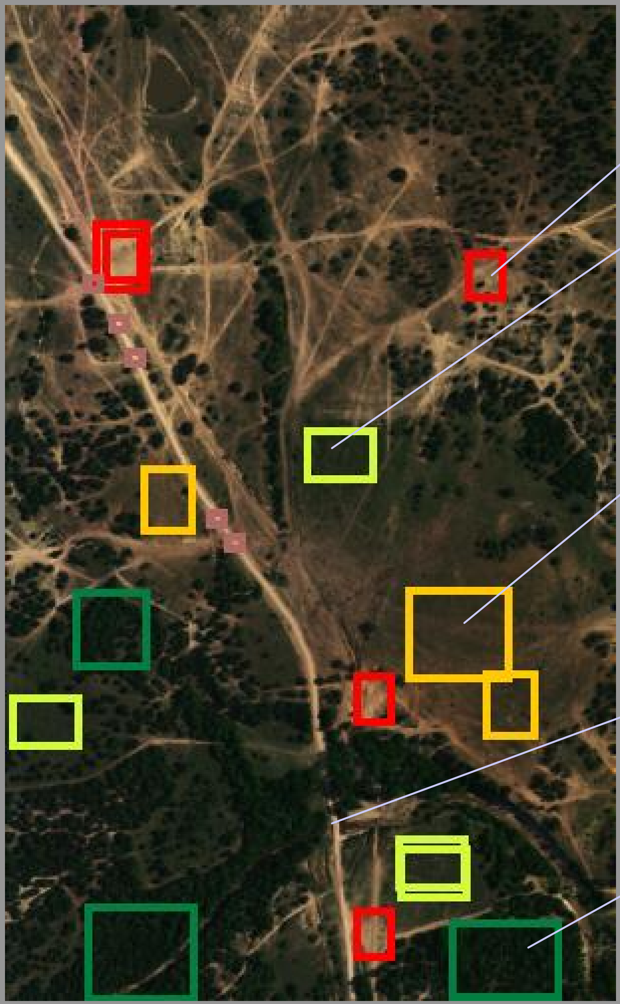


Automatically Selected Bands



"True Color" TERRAIN Image

HYDICE
bands 49,35,15



bare soil

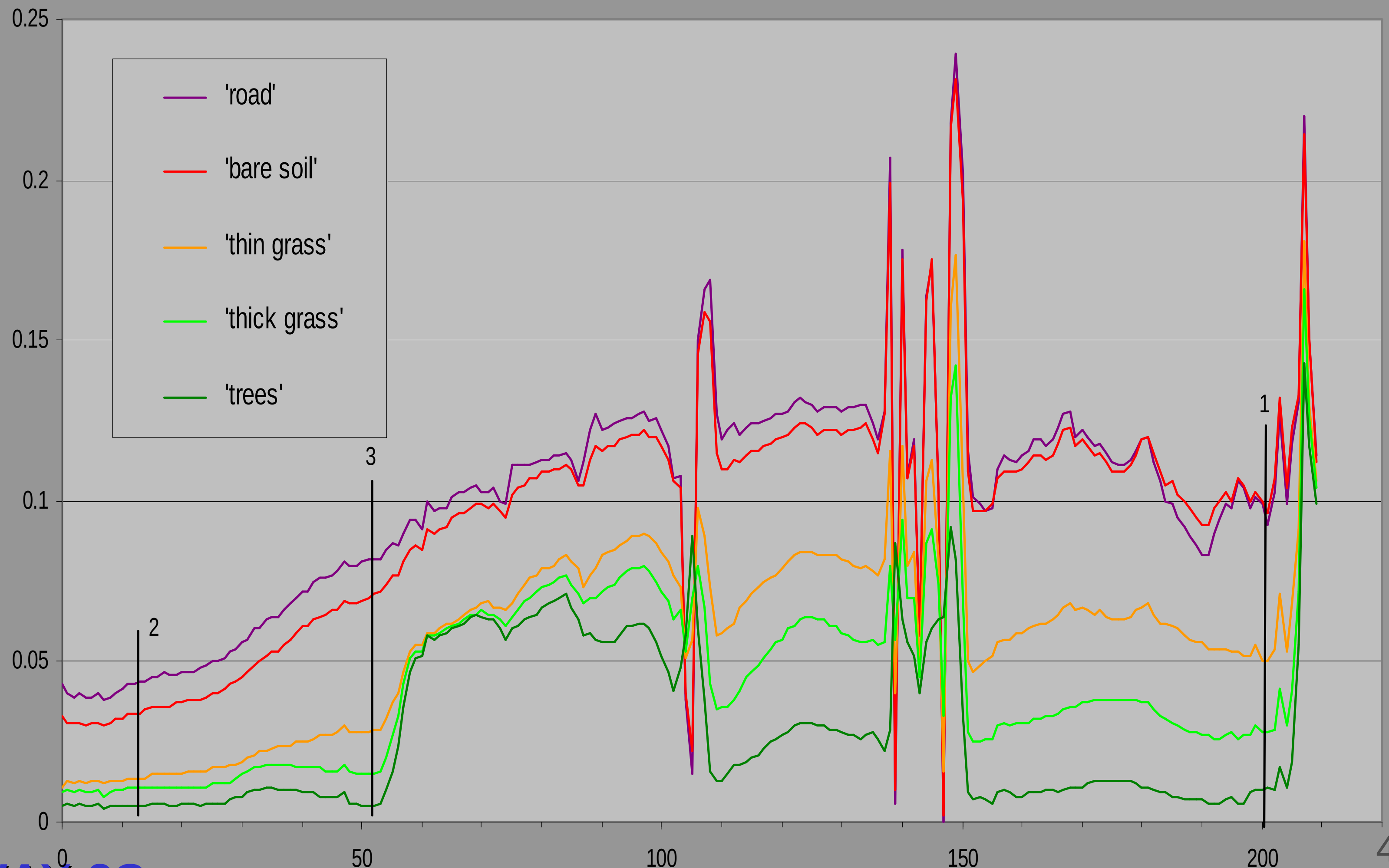
thick grass

thin grass

road

trees

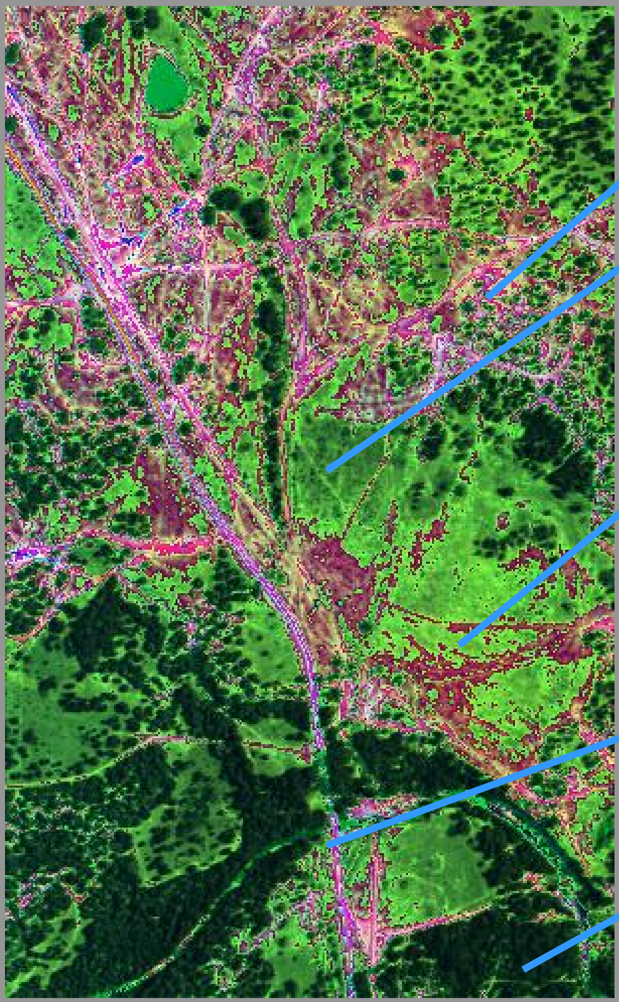
Terrain: Mean Spectra and Selected Bands



"Most Relevant" TERRAIN Image

HYDICE
bands 199,52,17

> 98% data reduction from original hyperspectral cube.

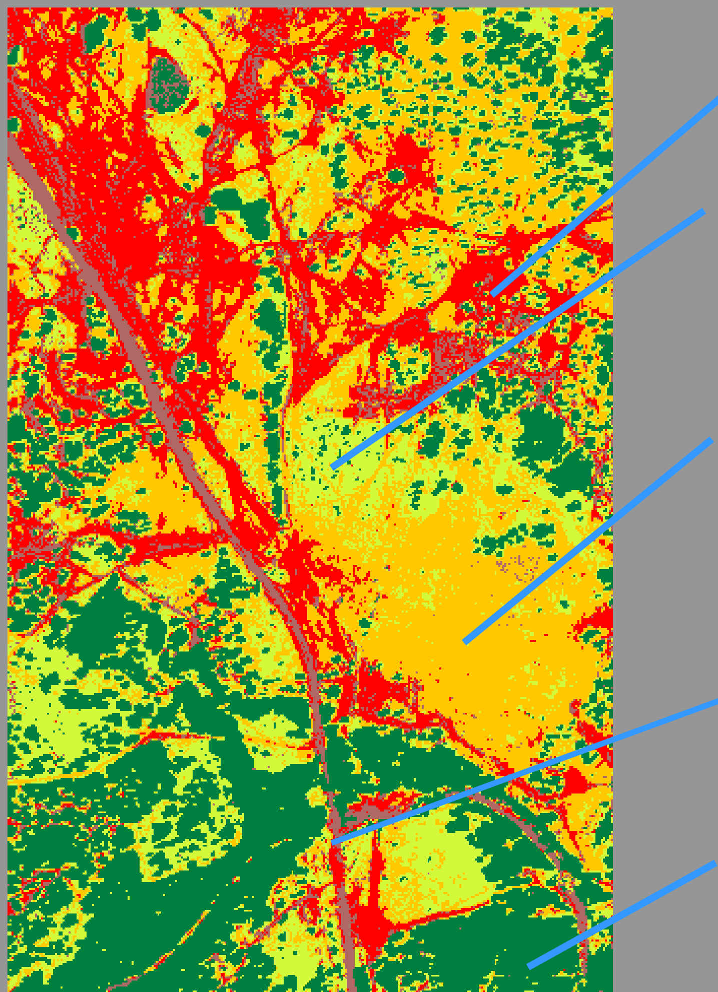


WAY-2C

Interpreted TERRAIN Image

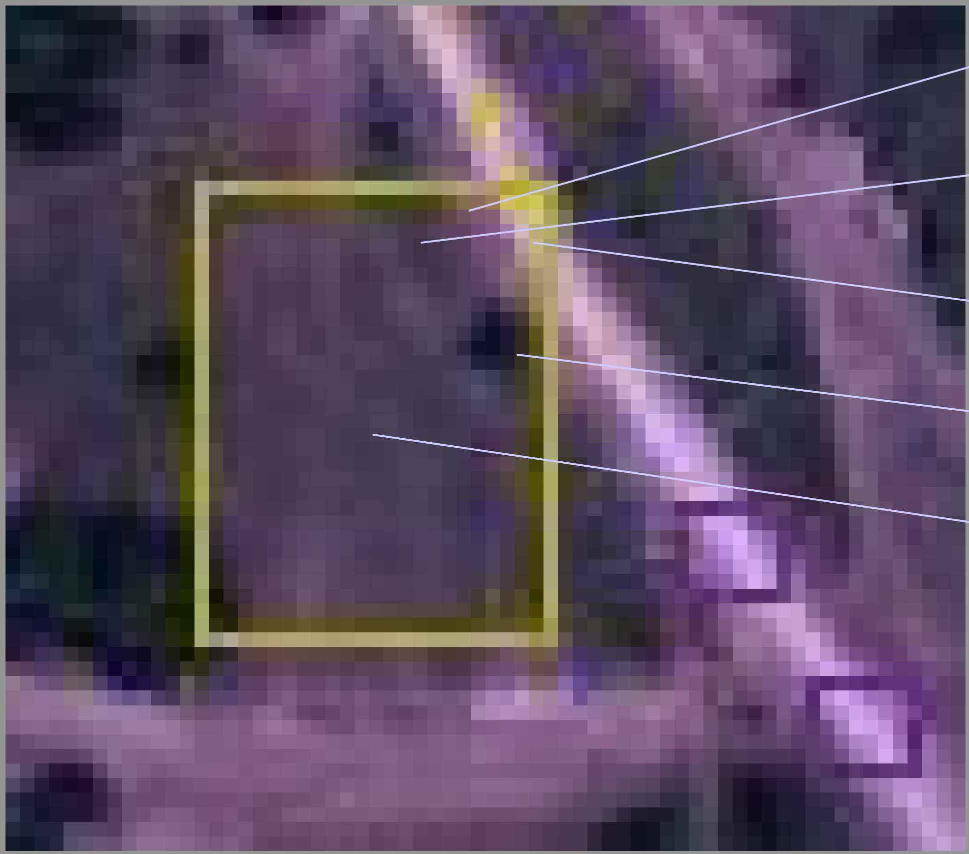
Based on
HYDICE
bands 199,52,17

uncertainty = 0.48 bits
PSPA = 71.8%



WAY-2C

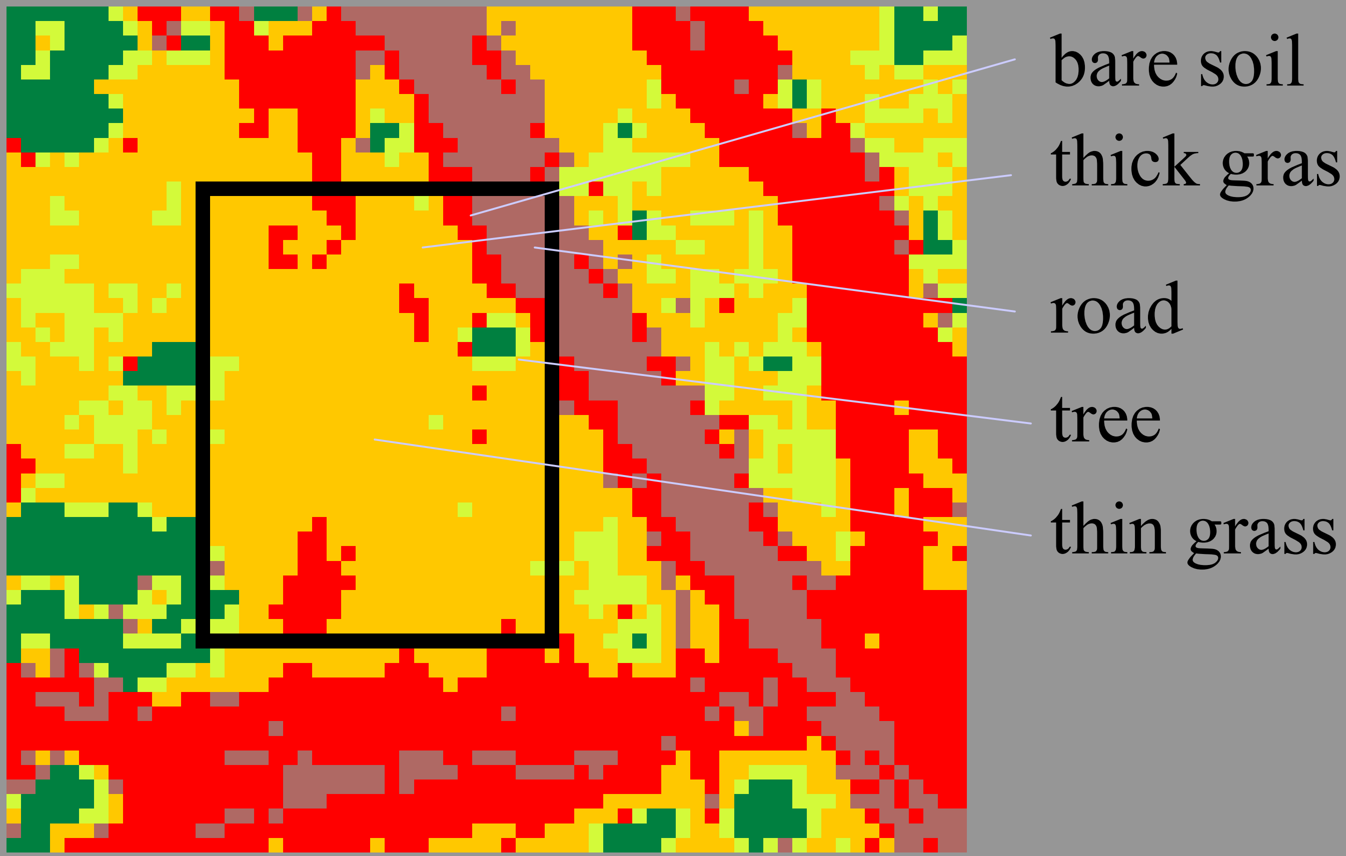
“Thin Grass” Training Region Detail



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

Challenge: Differentiating Apples From Leaves.

Simulated Orchard

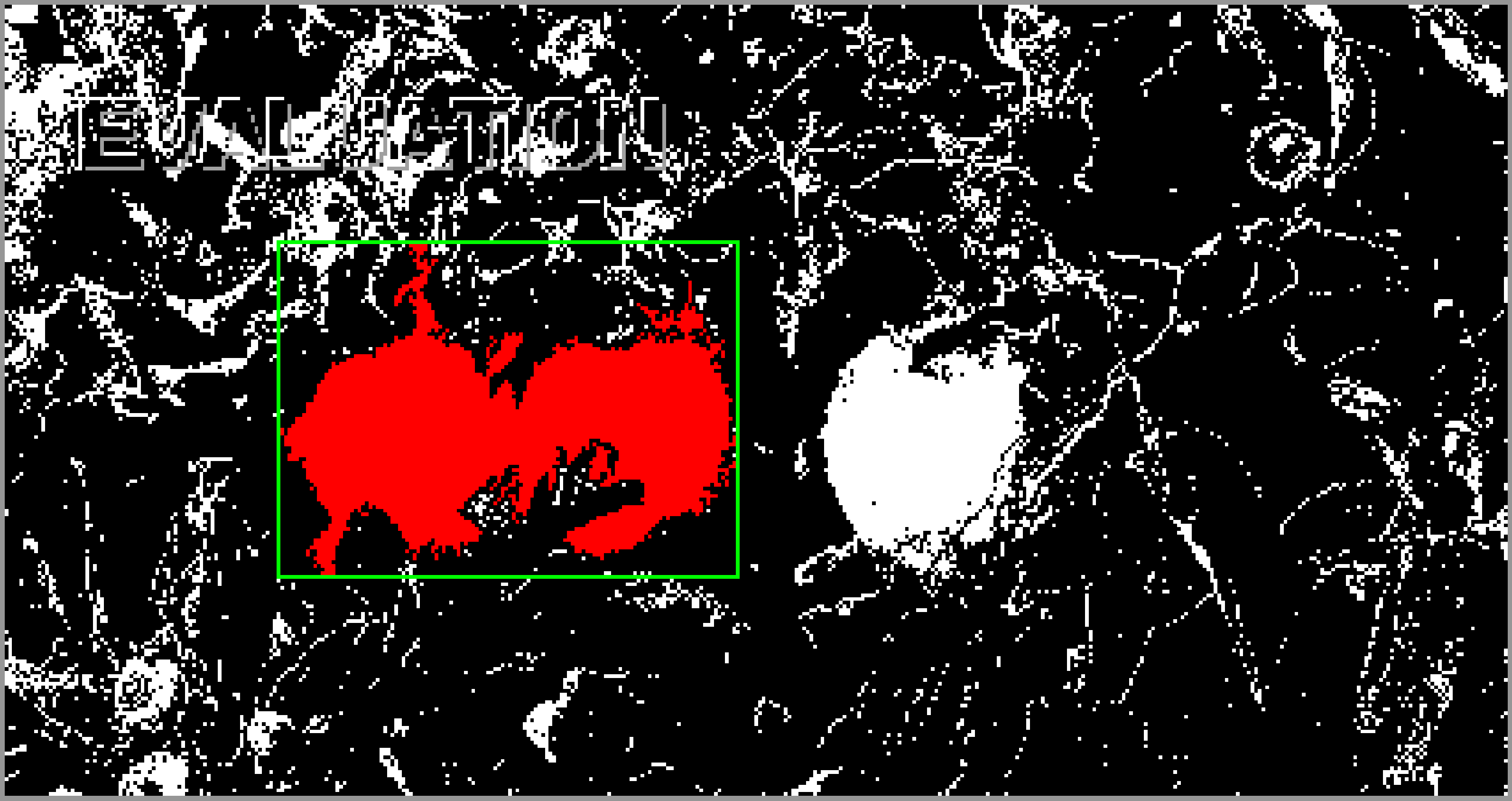
Courtesy Jürgen Hillmann, XIMEA GmbH

Training Regions:
(apple and background)

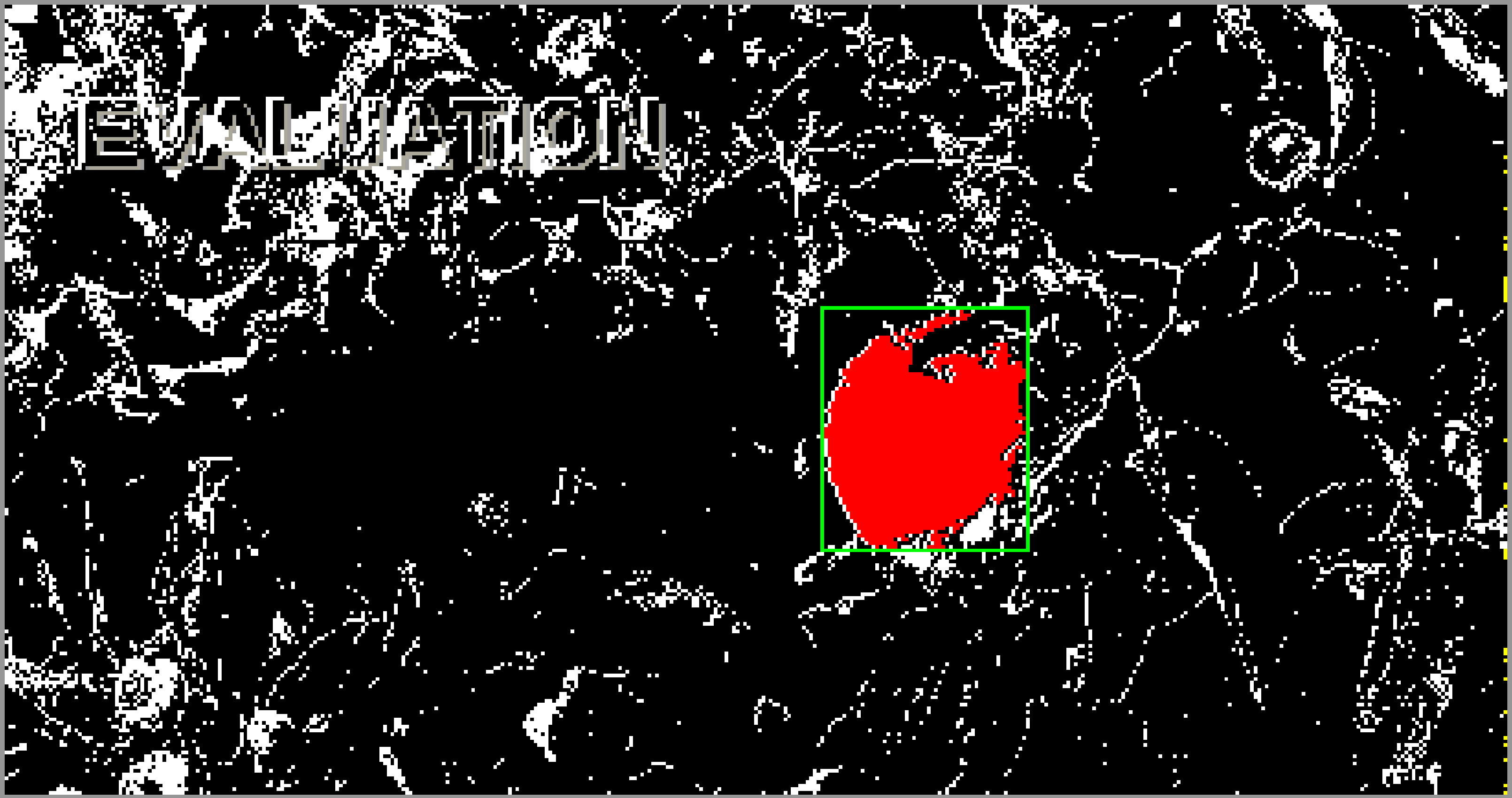
Intermediate Image

Pixel by Pixel Interpretation

Largest Target Apple Blob Bounding Box for Robot Picker



2nd Largest Target Apple Blob Bounding Box for Robot Picker



Method Summary

- Quickly locates most relevant data subset.
- Can dramatically reduce transmission storage and computation requirements.
- Eliminates need for manual data editing.
- Provides numerical estimate of accuracy.
- Speeds data interpretation.
- Requires minimal user skills and training

Successfully tested on:

Hyperspectral:

- ENVI: BIP,BIL,BSQ
- AVIRIS
- HYDICE
- HYPERION

Multispectral from:

- DigitalGlobe, Inc.

Other:

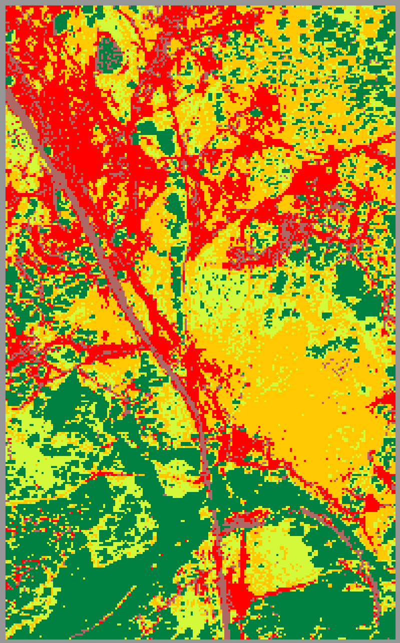
- Accelerometer array time series.
- Oil & gas well log data

Opens opportunities for dual use imagers that can:

- Capture cube with many spectral bands.
- Under semi-skilled operator guidance quickly determine application and condition specific optimum data subset combination.
- Be instantly software configurable to deliver only that optimum data subset for real time collection and analysis.

A New Method for Supervised Optimum Multispectral and Hyperspectral Data Subset Selection

End



WAY-2C