Target detection applied to detection of adulterants in powdered raw material using near infrared hyperspectral imaging

IDRC – 30 July, 2012
Neal B. Gallagher[a], Robert T. Roginski[a], Gregory A. Israelson[b]
[a]Eigenvector Research, Inc.,
[b]Nestlé Purina
nealg@eigenvector.com

Problem Statement

- Detect and classify adulterants in powdered raw materials
  - low levels of adulterants expected
  - safety and quality are primary objectives
  - quantification is secondary

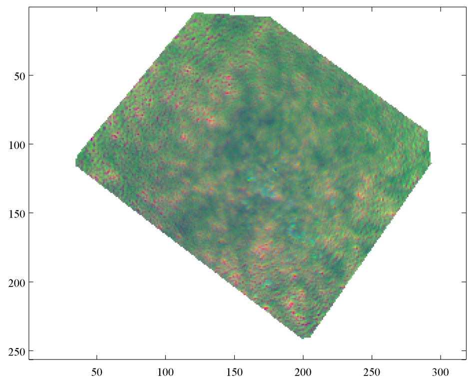
- Challenge problem …

Anomaly Detection

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}$$

$$\mathbf{T} = \mathbf{X}\mathbf{P}$$

Image of Scores on PC 1 (97.73%) , PC 2 (1.28%), PC 3 (0.56%)



NIR reflectance image of a cellulosic swipe with RDX on it.
PCA scores 1,2,3 (RGB) for mean-centered data.

a) is an anomaly apparent, and
b) where is it?
c) if seen, can the analyte be identified?

data set: courtesy OPOTEK, Inc., Carlsbad, CA

EIGENVECTOR RESEARCH INCORPORATED

Target Detection

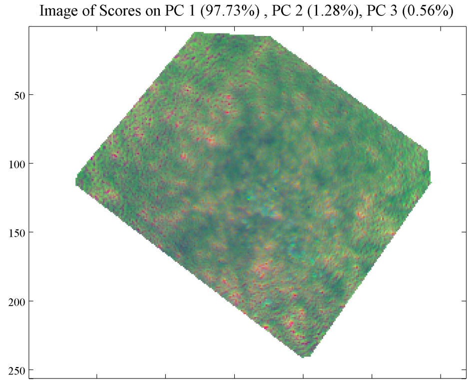
$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}$$

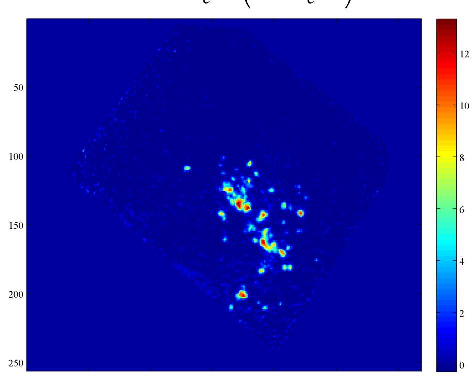
$$\mathbf{T} = \mathbf{X}\mathbf{P}$$

$$\mathbf{X} = \mathbf{C}\mathbf{S}^T + \mathbf{E}$$

$$\mathbf{C} = \mathbf{X}\boldsymbol{\Sigma}_c^{-1}\mathbf{S}\left(\mathbf{S}^T\boldsymbol{\Sigma}_c^{-1}\mathbf{S}\right)^{-1}$$

Image of Scores on PC 1 (97.73%) , PC 2 (1.28%), PC 3 (0.56%)





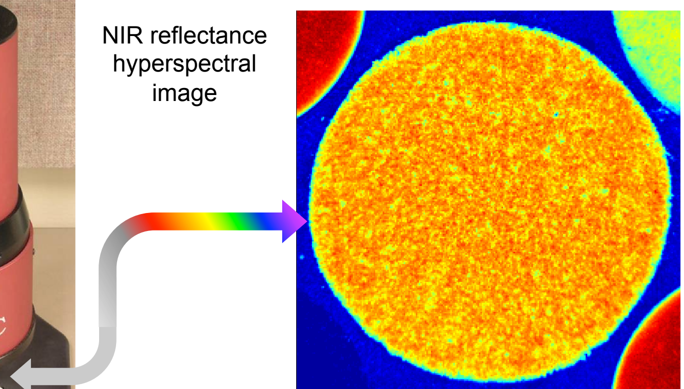
data set: courtesy OPOTEK, Inc., Carlsbad, CA

EIGENVECTOR RESEARCH INCORPORATED

Model System Studied

- ppm levels of melamine in wheat gluten

NIR reflectance
hyperspectral
image



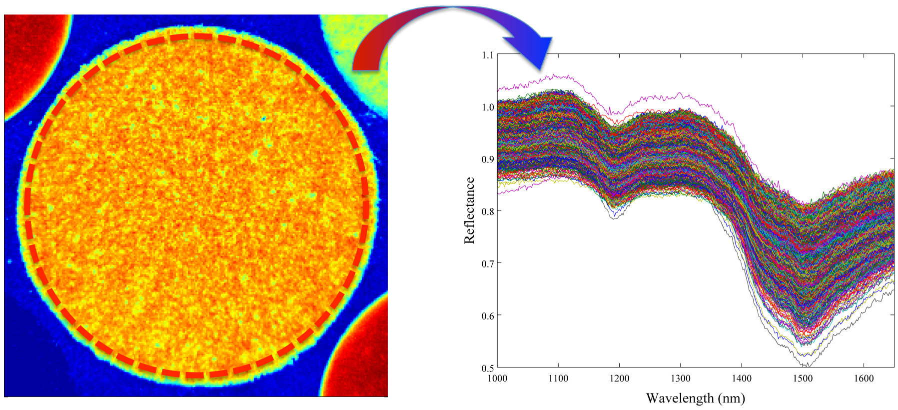
EIGENVECTOR RESEARCH INCORPORATED

Powdered Raw Materials

- Challenges
  - scattering and particle size distributions changes sample-to-sample
  - same material from a wide variety of sources
- Calibration for typical ILS models difficult
  - unlikely to acquire a calibration data set that spans all the sample variation expected to be seen
    - and if you do, the net analyte signal suffers
  - unlikely to use one ILS model from a single material for multiple raw materials

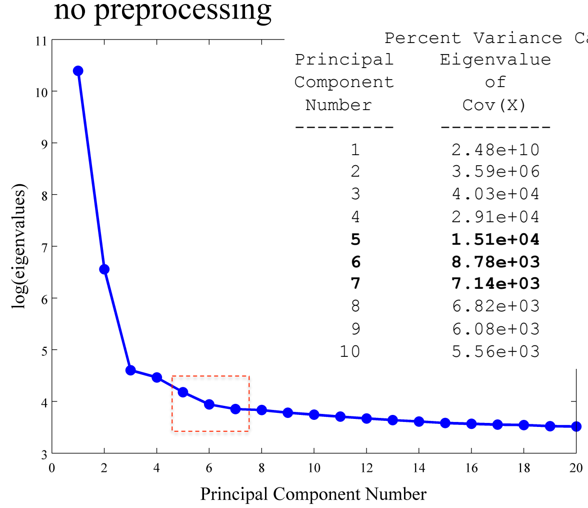
EIGENVECTOR RESEARCH INCORPORATED

Signal from the unadulterated wheat gluten is highly variable and is much stronger than the adulterant.



unadulterated wheat gluten

Eigenvalues for 200 ppm melamine in wheat gluten
no preprocessing



Percent Variance Captured by PCA Model

Principal Component Number	Eigenvalue of Cov(X)	% Variance Captured This PC	% Variance Captured Total
1	2.48e+10	99.983438	99.983438
2	3.59e+06	0.014493	99.997931
3	4.03e+04	0.000162	99.998093
4	2.91e+04	0.000117	99.998210
5	**1.51e+04**	**0.000061**	**99.998271**
6	**8.78e+03**	**0.000035**	**99.998307**
7	**7.14e+03**	**0.000029**	**99.998336**
8	6.82e+03	0.000027	99.998363
9	6.08e+03	0.000025	99.998388
10	5.56e+03	0.000022	99.998410

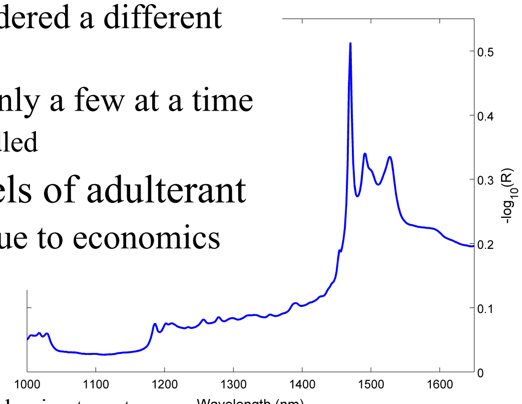
Clutter Definition

- sensor noise and signal due to interferences
  - (DOD, NATO) Unwanted signals, echoes, or images on the face of the display tube, which interfere with observation of desired signals.
  - It's measured signal unrelated to the target
    - it can be correlated or not

EIGENVECTOR RESEARCH INCORPORATED

Adulterants = Targets

- Adulterant spectrum ~constant
  - variants can be considered a different adulterant
  - tend to appear with only a few at a time
    - multiples can be handled
- Interest is in low levels of adulterant
  - higher levels likely due to economics

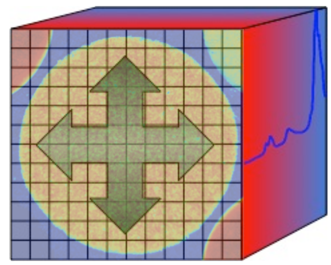


melamine target

EIGENVECTOR RESEARCH INCORPORATED

Hyperspectral

- rapid and non-invasive
- spectra measured at many pixels
  - good sampling
  - clutter can be characterized local to each sample
    - no need for extensive calibration sets!



- although adulterant quantities may be low on a mass basis, the spectral signal in individual pixels can be dominated by the adulterant.
  - detection of low quantities of adulterants via HI is less hampered by dilution effects expected from wet chemistry methods.

EIGENVECTOR RESEARCH INCORPORATED

Mixed Pixels

- Pure adulterant pixels are rare for low levels
  - Often signal from the background is also present in a pixel
- The result is that the signal is a mixture of background and adulterant
  - mixed pixels
  - linear mixture model useful
    - “sub-pixel resolution”
    - caution with interpretation!

EIGENVECTOR RESEARCH INCORPORATED

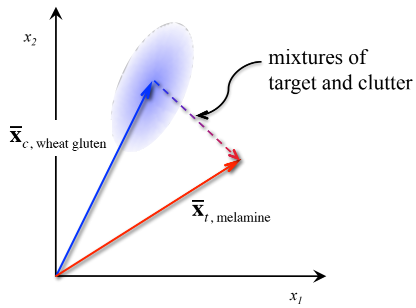
How to Account for Mixed Pixels

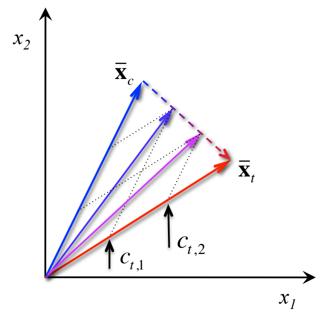
- closure: as target signal increases, background signal decreases →
- Mixed pixels lie on the line between the mean responses

$$\mathbf{x} = (1 - c_t)\bar{\mathbf{x}}_c + c_t\bar{\mathbf{x}}_t + \mathbf{e}$$

c_t is target contribution

c_c=(1-c_t) clutter spectrum





Inverse Clutter Covariance → Whitening

$$\boldsymbol{\Sigma}_c = \tfrac{1}{M-1}\left(\mathbf{X} - \mathbf{1}\bar{\mathbf{x}}_c^T\right)^T\left(\mathbf{X} - \mathbf{1}\bar{\mathbf{x}}_c^T\right)$$

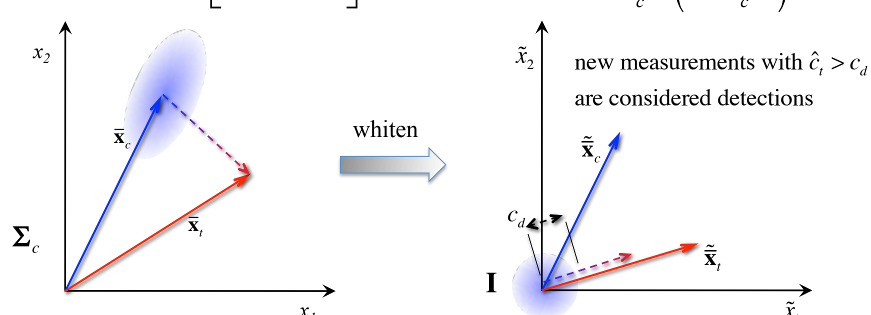
$$\tilde{\mathbf{X}} = \mathbf{X}\boldsymbol{\Sigma}_c^{-1/2}$$

$$\tilde{\mathbf{S}} = \mathbf{S}\boldsymbol{\Sigma}_c^{-1/2} \qquad \mathbf{S} = \begin{bmatrix} \bar{\mathbf{x}}_c & \bar{\mathbf{x}}_t \end{bmatrix}$$

$$\mathbf{X} = \mathbf{c}\mathbf{S}^T + \mathbf{E}$$

$$\begin{bmatrix} \hat{c}_c & \hat{c}_t \end{bmatrix} = \tilde{\mathbf{X}}\tilde{\mathbf{S}}\left(\tilde{\mathbf{S}}^T\tilde{\mathbf{S}}\right)^{-1}$$

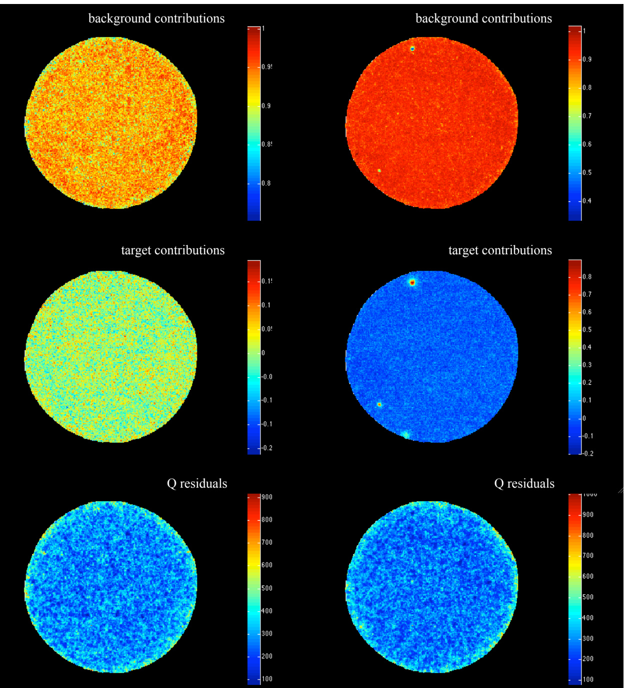
$$\hat{\mathbf{c}} = \mathbf{X}\boldsymbol{\Sigma}_c^{-1}\mathbf{S}\left(\mathbf{S}^T\boldsymbol{\Sigma}_c^{-1}\mathbf{S}\right)^{-1}$$



example contributions

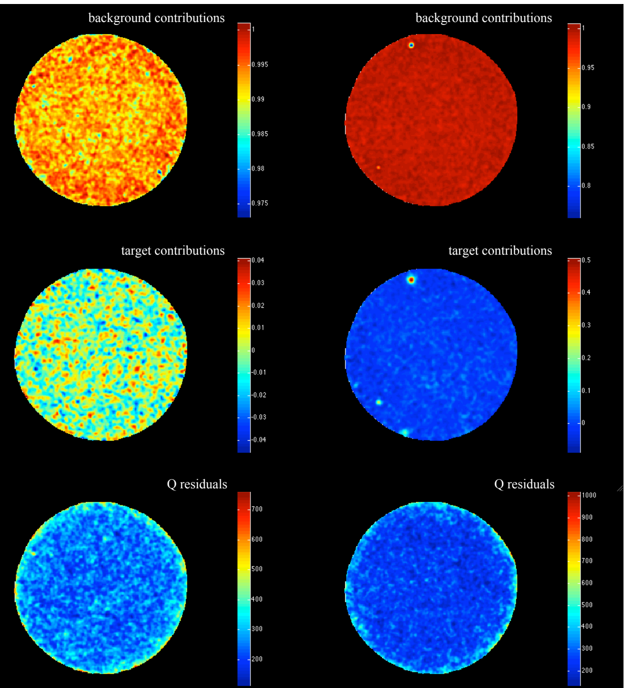
distributions are not normal (have tails) that can lead to single pixel false alarms

left = null
right = positive
200 ppm
high Q not removed



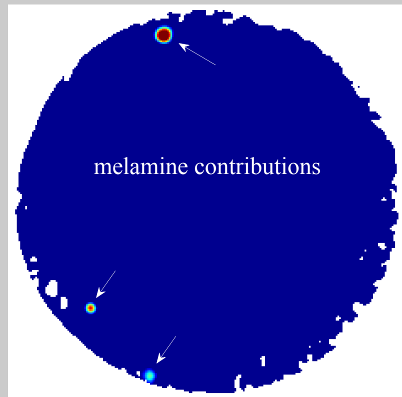
spatial filtering prior to target detection provided more normally distributed contributions

left = null
right = positive
200 ppm
high Q not removed



the more red, the higher the signal
the more blue, the lower the signal

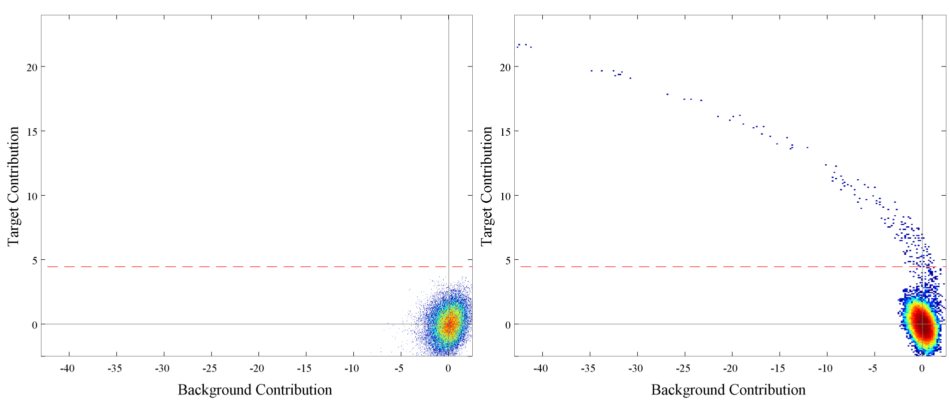
200 ppm sample



pixels w/ high sum-squared residuals are not considered in the detection
this spot was caught in anomaly detection but
was *not* melamine

EIGENVECTOR RESEARCH INCORPORATED

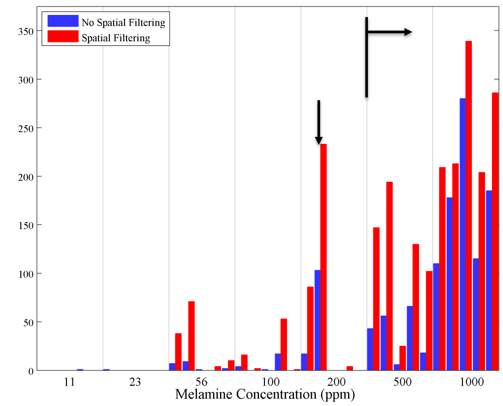
target contributions versus background clutter contributions
contributions are centered and scaled (t-stat)



200 ppm sample

EIGENVECTOR RESEARCH INCORPORATED

- Detection capability increased with concentration.
  - However, this should not be interpreted as capability to quantify.
  - The likelihood of a particle to be at, or near, the surface increases with concentration.



- Detections with spatial filtering tended to be more dramatic
  - Not clear if the ability to detect improved.
  - Parameters were set so that there were no false alarms on the five null images (at 0 ppm).

EIGENVECTOR RESEARCH INCORPORATED

Challenges: Present and Future

- Clutter
  - hyperspectral allows us to characterize and accounting for it via local modeling
- Targets
  - identifying targets, accounting for sensing physics e.g., linearization and grey-models
- Tests w/ other targets and raw materials shows that the decision limits are different for each background material but the method performs similarly
  - requires far fewer calibration samples than traditional ILS model
- Sensor fusion
  - are other methods needed to enhance detection and classification?

EIGENVECTOR RESEARCH INCORPORATED

Keep in mind that…

"The detection system being considered is the sensor that provides the measurements, the scenario in which it is to be deployed *and* the algorithm used to extract the desired information.

These must be developed concurrently for the greatest chance at success because what is learned during data analysis and algorithm development often feeds back directly to instrument design in an effort to maximize *signal-to-clutter* not just *signal-to-noise*."

EIGENVECTOR RESEARCH INCORPORATED