On the relationship between whitened principal components analysis and maximum autocorrelation factors

Neal B. Gallagher, Jeremy M. Shaver,
Randall Bishop, Robert T. Roginski,
Barry M. Wise

SSC'13 17-20, June 2013

EIGENVECTOR RESEARCH INCORPORATED

On the relationship between whitened principal components analysis and maximum autocorrelation factors

Maximum autocorrelation factors (MAF)[1] and the use of de-cluttering algorithms have been gaining popularity recently in chemometrics. Generalized least squares weighting, developed by Aitken c. 1935[2], is a popular and useful de-cluttering algorithm that can be used to "whiten" data prior to principal components analysis (PCA). The relationship between MAF and whitened PCA is elucidated with the objective of 'de-mystifying' the MAF algorithm. The insight is intended to provide a fundamental understanding of the methods such that they can be employed to best effect and that related methods can be easily developed using a similar relationship. Examples will be shown for hyperspectral images.

1. A.A. Green, M. Berman, P. Switzer, and M.D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE T Geosci Remote Sens*, **26**, 65-74, 1988.
2. Aitken, A., “On least squares and linear combinations of observations,” *Proceedings of the Royal Society of Edinburgh*, 1935, **55**, 42-48.

EIGENVECTOR RESEARCH INCORPORATED

Compare PCA and WPCA

PCA = principal components analysis
WPCA = whitened PCA

	Objective to Maximize		**Scores**
PCA	$O(\mathbf{p}) = \mathbf{p}^T \boldsymbol{\Sigma} \mathbf{p} \mid \mathbf{p}^T \mathbf{p} = 1$	$\boldsymbol{\Sigma} = \mathbf{X}^T \mathbf{X}$	$\mathbf{t} = \mathbf{X}\mathbf{p}$
WPCA	$O(\mathbf{p}) = \mathbf{p}^T \boldsymbol{\Sigma}_c^{-1/2} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_c^{-1/2} \mathbf{p} \mid \mathbf{p}^T \mathbf{p} = 1$		$\mathbf{t} = \mathbf{X}\boldsymbol{\Sigma}_c^{-1/2}\mathbf{p}$

$$\boldsymbol{\Sigma} = \mathbf{X}^T \mathbf{X}$$

$$\boldsymbol{\Sigma}_c = \mathbf{X}_c^T \mathbf{X}_c$$

$\mathbf{X}$ = data measurements
$\mathbf{X}_c$ = clutter measurements

Clutter is interferences and noise; it is measured signal not of primary interest

EIGENVECTOR RESEARCH INCORPORATED

Compare WPCA and MAF

WPCA = whitened PCA
MAF = maximum autocorrelation factors

	Objective to Maximize		**Scores**
WPCA	$O(\mathbf{p}) = \mathbf{p}^T \boldsymbol{\Sigma}_c^{-1/2} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_c^{-1/2} \mathbf{p} \;\mid\; \mathbf{p}^T \mathbf{p} = 1$		$\mathbf{t} = \mathbf{X} \boldsymbol{\Sigma}_c^{-1/2} \mathbf{p}$
MAF	$O(\mathbf{w}) = \dfrac{\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}}{\mathbf{w}^T \boldsymbol{\Sigma}_\Delta \mathbf{w}}$	$\boldsymbol{\Sigma} = \mathbf{X}^T \mathbf{X}$ $\boldsymbol{\Sigma}_\Delta = \mathbf{X}^T \mathbf{D}_1^T \mathbf{D}_1 \mathbf{X}$	$\mathbf{a} = \mathbf{X}\mathbf{w}$

EIGENVECTOR RESEARCH INCORPORATED

Define PCA, WPCA and MAF

PCA = principal components analysis
WPCA = whitened PCA
MAF = maximum autocorrelation factors

	Eigenvalue Problem	**Eigenvalue ; Eigenvector**	**Scores**
PCA	$\mathbf{\Sigma p} = \lambda \mathbf{p}$	$\mathbf{p}$; λ	$\mathbf{t} = \mathbf{Xp}$
WPCA	$\mathbf{\Sigma}_c^{-1/2} \mathbf{\Sigma \Sigma}_c^{-1/2} \mathbf{p} = \lambda \mathbf{p}$	$\mathbf{p}$; λ	$\mathbf{t} = \mathbf{X\Sigma}_c^{-1/2} \mathbf{p}$
MAF	$\mathbf{\Sigma w} = \alpha \mathbf{\Sigma}_\Delta \mathbf{w}$	$\mathbf{w}$; α	$\mathbf{a} = \mathbf{Xw}$

EIGENVECTOR RESEARCH INCORPORATED

Compare WPCA and MAF

	Eigenvalue Problem	Eigenvalue ; Eigenvector	Scores
WPCA	$\boldsymbol{\Sigma}_c^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_c^{-1/2}\mathbf{p} = \lambda\mathbf{p}$	$\mathbf{p}$; λ	$\mathbf{t} = \mathbf{X}\boldsymbol{\Sigma}_c^{-1/2}\mathbf{p}$
MAF	$\boldsymbol{\Sigma}\mathbf{w} = \alpha\boldsymbol{\Sigma}_\Delta\mathbf{w}$	$\mathbf{w}$; α	$\mathbf{a} = \mathbf{X}\mathbf{w}$
MSF	$\boldsymbol{\Sigma}\mathbf{w} = \alpha\boldsymbol{\Sigma}_c\mathbf{w}$	$\mathbf{w}$; λ	$\mathbf{t} = \mathbf{X}\mathbf{w}$
		$\mathbf{w} = \boldsymbol{\Sigma}_c^{-1/2}\mathbf{p}$	

$$\boldsymbol{\Sigma}\mathbf{w} = \alpha\boldsymbol{\Sigma}_c^{1/2}\boldsymbol{\Sigma}_c^{1/2}\mathbf{w}$$

$$\boldsymbol{\Sigma}_c^{-1/2}\boldsymbol{\Sigma}\left(\boldsymbol{\Sigma}_c^{-1/2}\boldsymbol{\Sigma}_c^{1/2}\right)\mathbf{w} = \alpha\boldsymbol{\Sigma}_c^{1/2}\mathbf{w}$$

$$\boldsymbol{\Sigma}_c^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_c^{-1/2}\left(\boldsymbol{\Sigma}_c^{1/2}\mathbf{w}\right) = \alpha\left(\boldsymbol{\Sigma}_c^{1/2}\mathbf{w}\right)$$

EIGENVECTOR RESEARCH INCORPORATED

MSF

	Eigenvalue Problem	Eigenvalue ; Eigenvector	Scores
WPCA	$\boldsymbol{\Sigma}_c^{-1/2}\boldsymbol{\Sigma}_s\boldsymbol{\Sigma}_c^{-1/2}\mathbf{p} = \lambda\mathbf{p}$	$\mathbf{p}$; λ	$\mathbf{t} = \mathbf{X}\boldsymbol{\Sigma}_c^{-1/2}\mathbf{p}$
MSF	$\boldsymbol{\Sigma}_s\mathbf{w} = \alpha\boldsymbol{\Sigma}_c\mathbf{w}$	$\mathbf{w}$; λ $\mathbf{w} = \boldsymbol{\Sigma}_c^{-1/2}\mathbf{p}$	$\mathbf{t} = \mathbf{X}\mathbf{w}$

$$\boldsymbol{\Sigma}_s = \mathbf{X}_s^T\mathbf{D}_s^T\mathbf{D}_s\mathbf{X}_s$$

$$\boldsymbol{\Sigma}_c = \mathbf{X}_c^T\mathbf{D}_c^T\mathbf{D}_c\mathbf{X}_c$$

signal and clutter can include different data and spatial operators
(or different row operators for time-series)

EIGENVECTOR RESEARCH INCORPORATED

Six wires of different alloys in epoxy, measured with EDS.

1 (top): 100% Ni

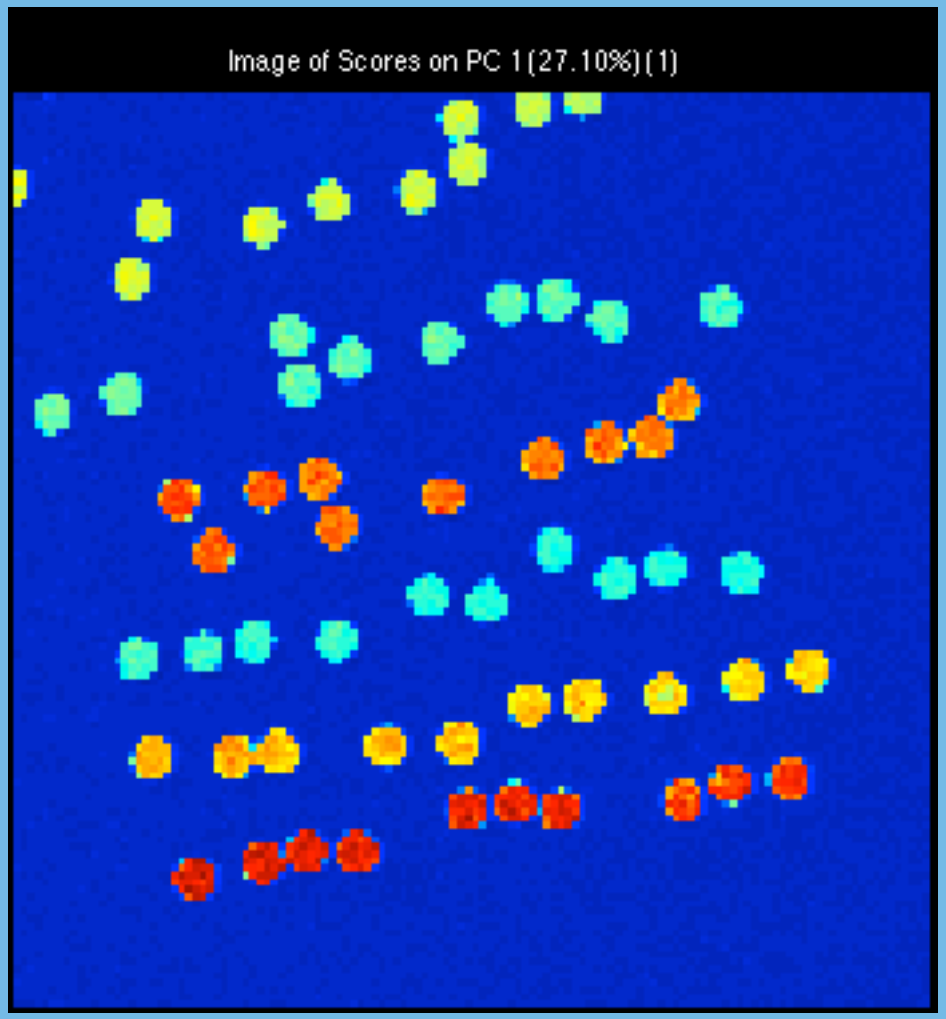
2: 36% Ni, 64% Fe

3: 70% Cu, 30% Zn

4: 16% Cr, 84% Fe

5: 13% Mn, 4% Ni, 83% Cu

6 (bottom): 100% Cu

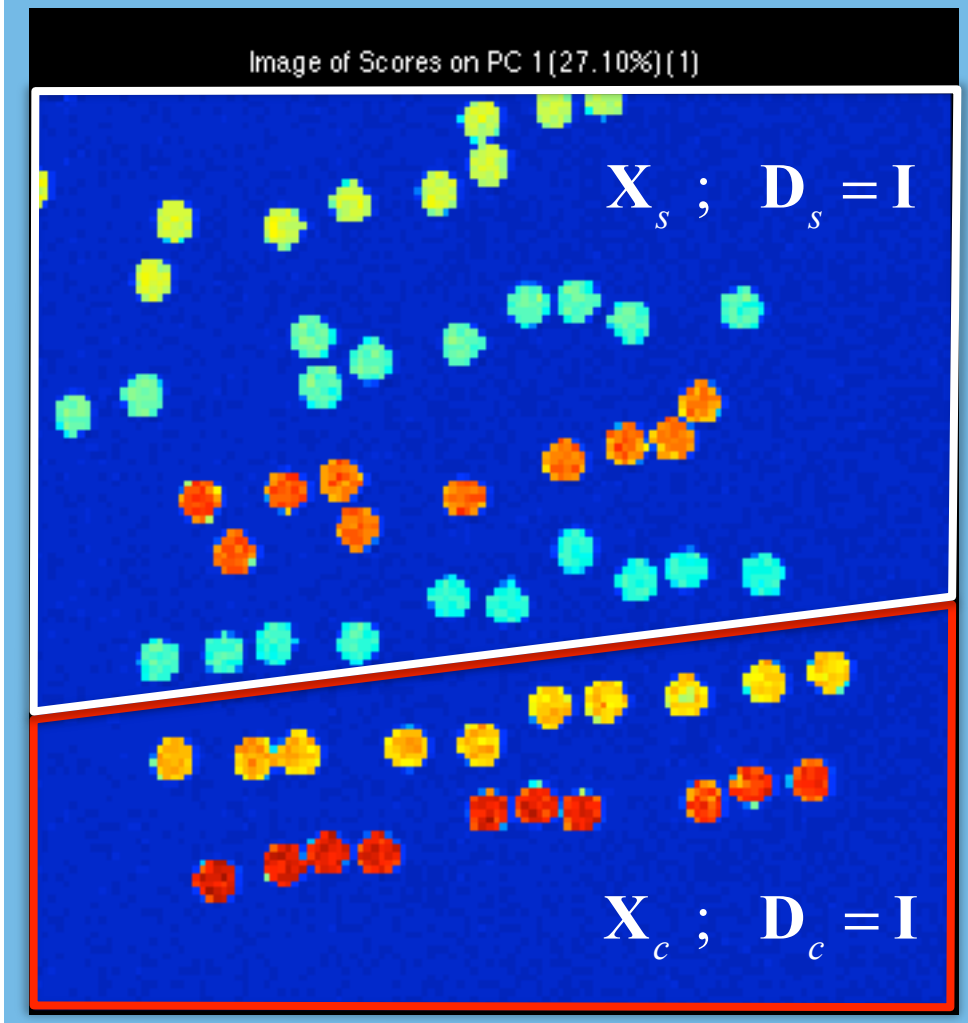


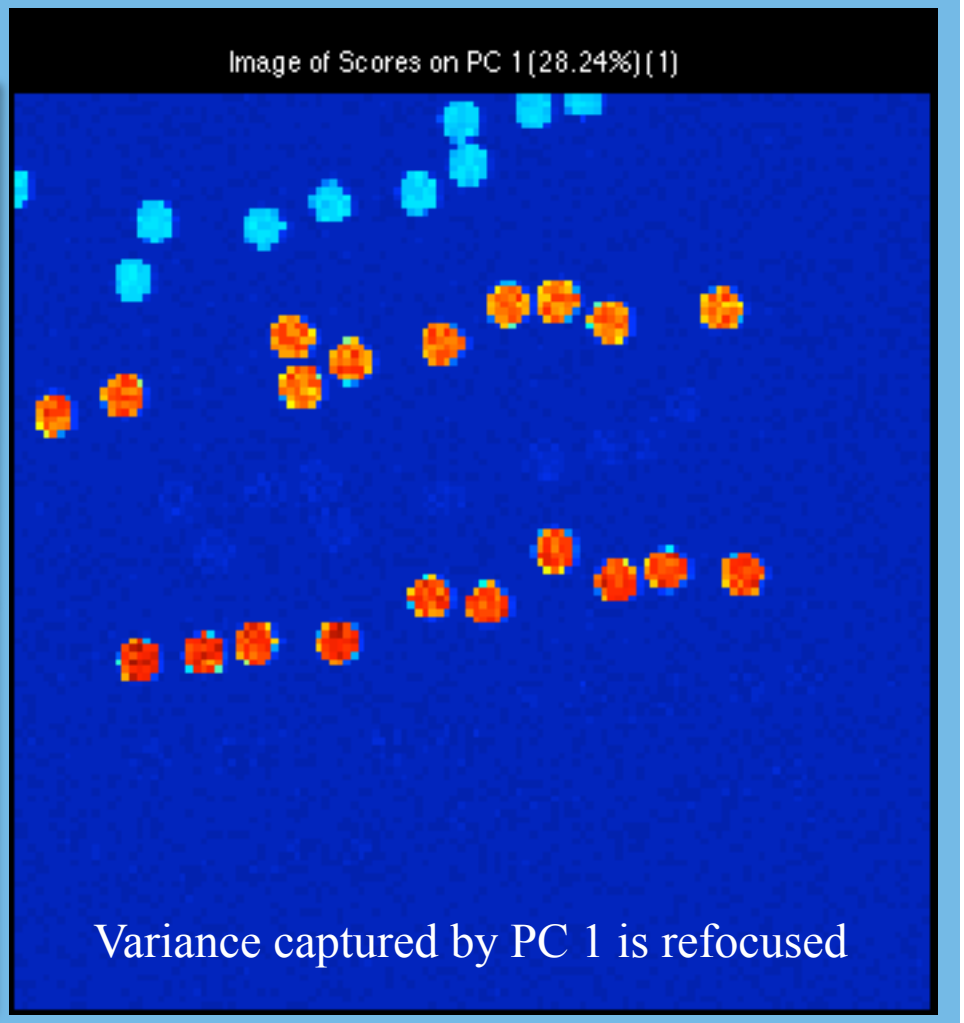
Michael R. Keenan, "Multivariate Analysis of Spectral Images Composed of Count Data," in Techniques and Applications of Hyperspectral Image Analysis, H.F. Grahn and P. Geladi, Eds (Wiley, Chichester, UK, 2007), pp. 89-126.

Paul Kotula at Sandia and Mike Keenan are gratefully acknowledged for this data

EIGENVECTOR RESEARCH INCORPORATED

Partitioning of Variance

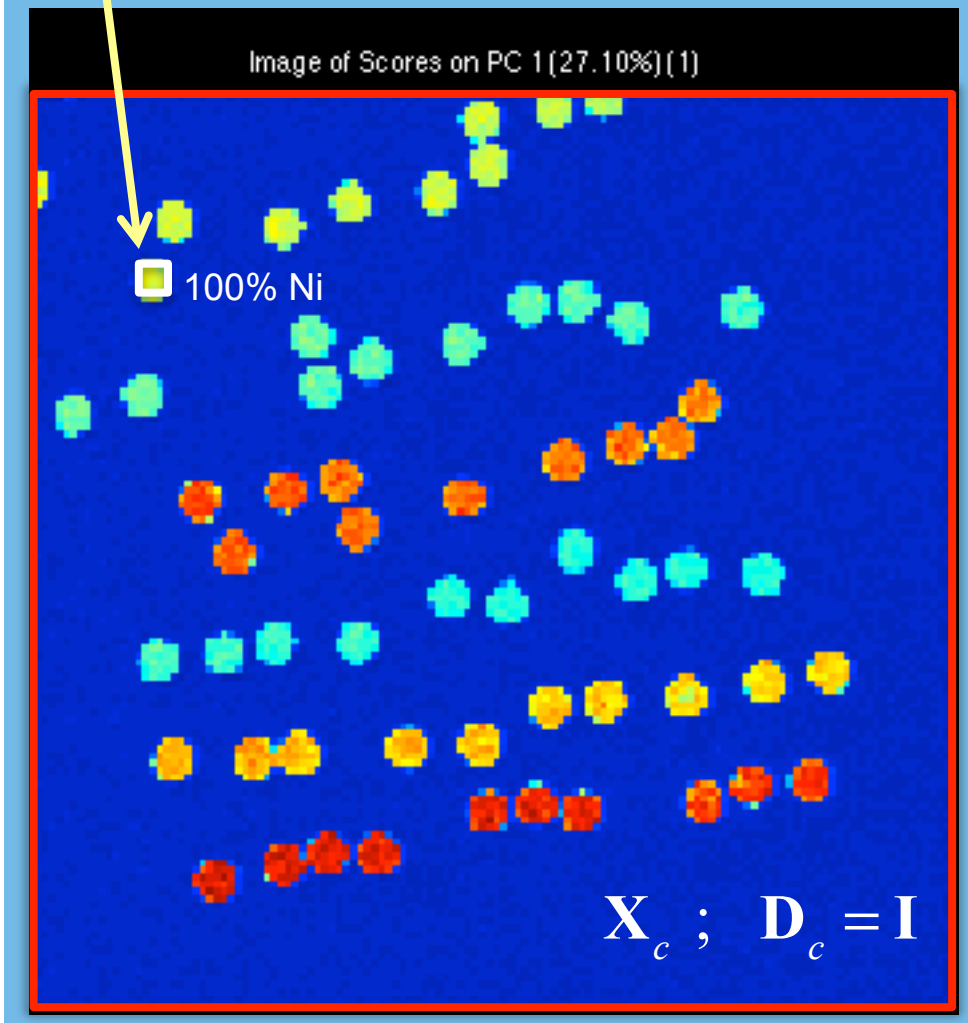


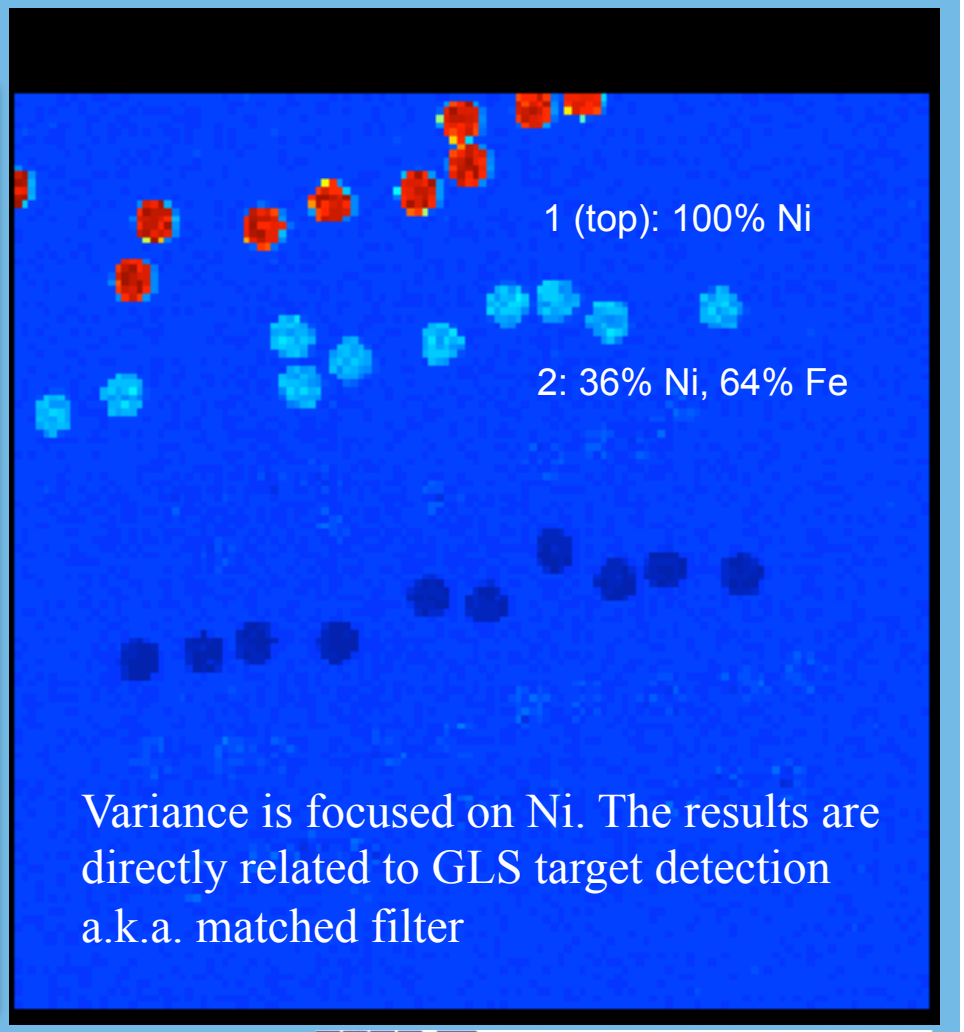


EIGENVECTOR RESEARCH INCORPORATED

Target Detection

$\mathbf{X}_s \; ; \; \mathbf{D}_s = \mathbf{I}$





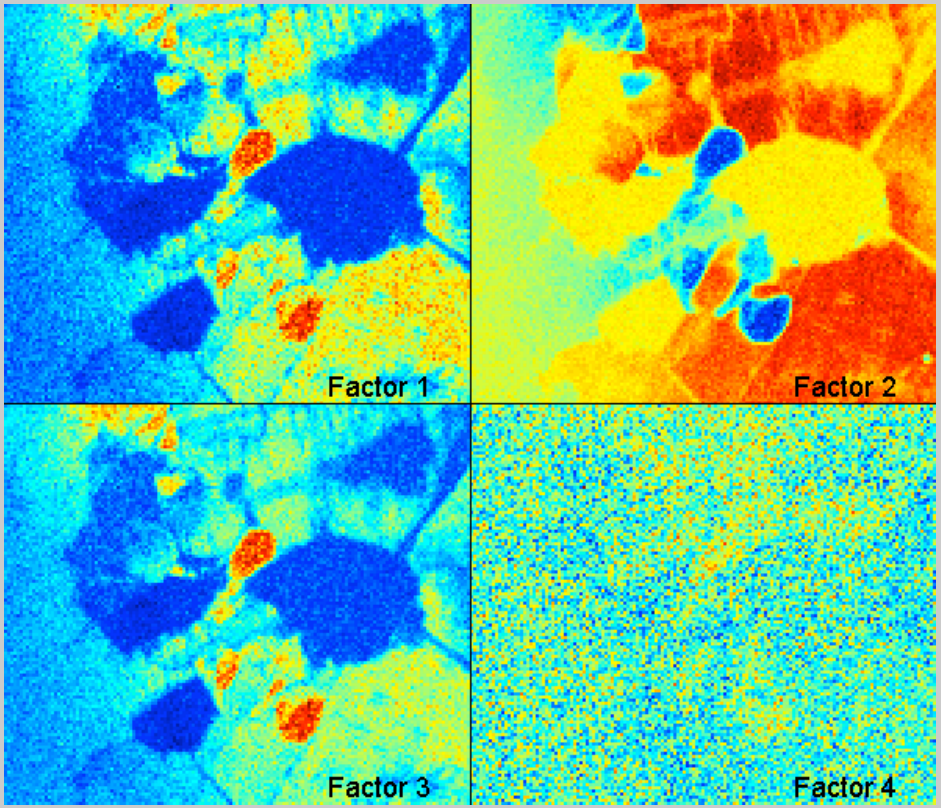
Variance is focused on Ni. The results are directly related to GLS target detection a.k.a. matched filter

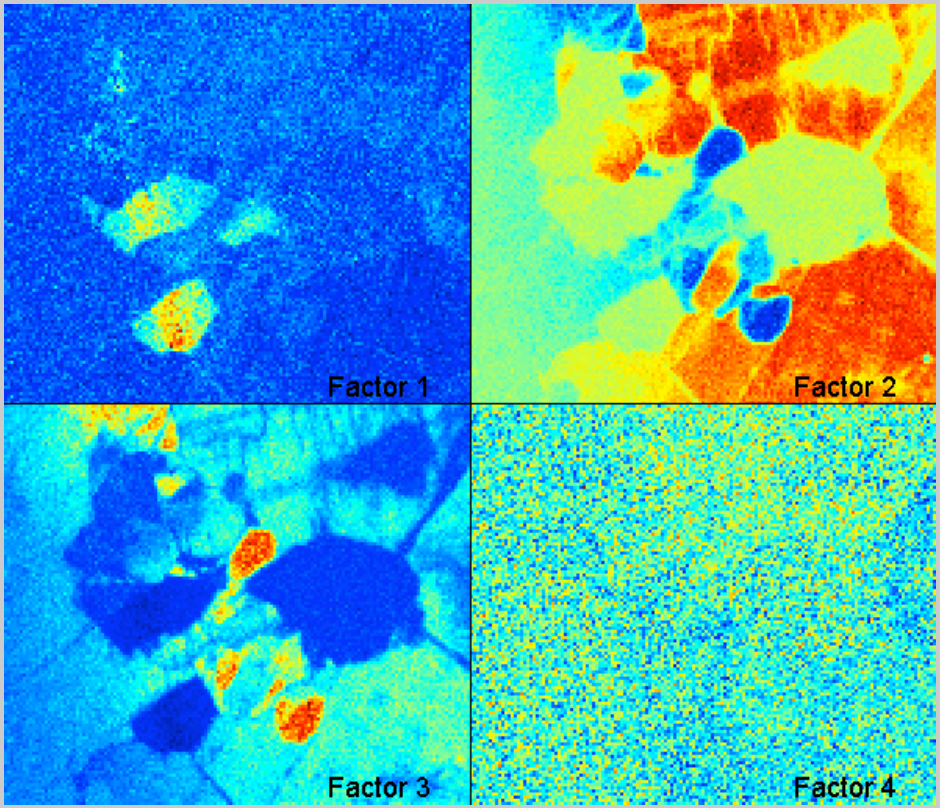
EIGENVECTOR RESEARCH INCORPORATED

PCA vs MAF

Raman spectra of Granite

$$\mathbf{X}_s = \mathbf{X}_c \; ; \quad \mathbf{D}_s = \mathbf{I} \; ; \quad \mathbf{D}_c = \mathbf{D}_1$$



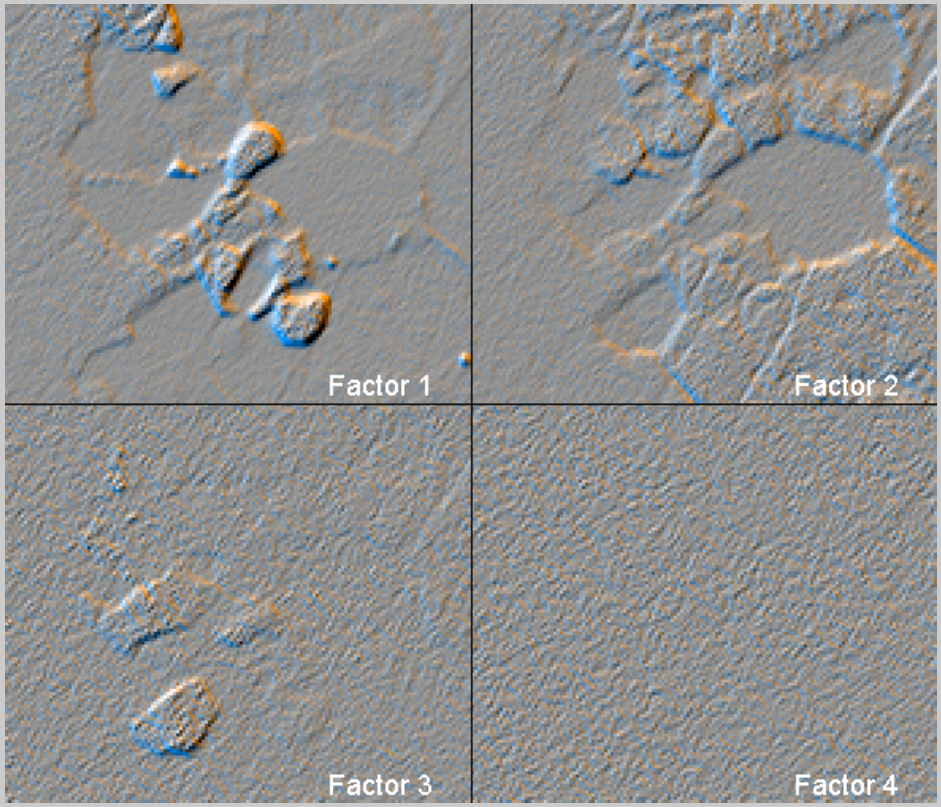


Integrating Raman and Micro X-ray Fluorescence Spectroscopic Imaging for Comprehensive Materials Characterization, George J. Havrilla,[1] Ursula Fittschen,[1] Eunah Lee,[2] Jeremy Shaver,[3] FACSS09, Oct 22, 2009.
[1]Los Alamos National Laboratory, Los Alamos, NM; [2]HORIBA Jobin Yvon, Inc, Edison, NJ; [3]Eigenvector Research, Inc., Wenatchee, WA

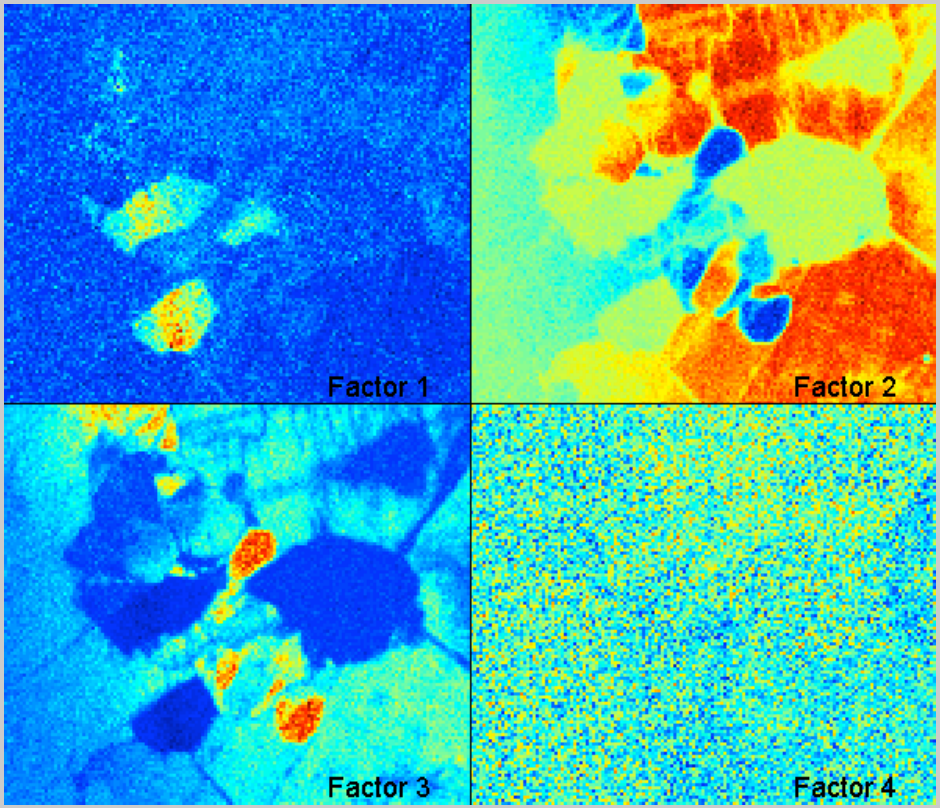
EIGENVECTOR RESEARCH INCORPORATED

MDF vs MAF

$$\mathbf{X}_s = \mathbf{X}_c \; ; \quad \mathbf{D}_s = \mathbf{D}_1 \; ; \quad \mathbf{D}_c = \mathbf{D}_2$$



$$\mathbf{X}_s = \mathbf{X}_c \; ; \quad \mathbf{D}_s = \mathbf{I} \; ; \quad \mathbf{D}_c = \mathbf{D}_1$$



Maximum Difference Factors is factor based edge detection

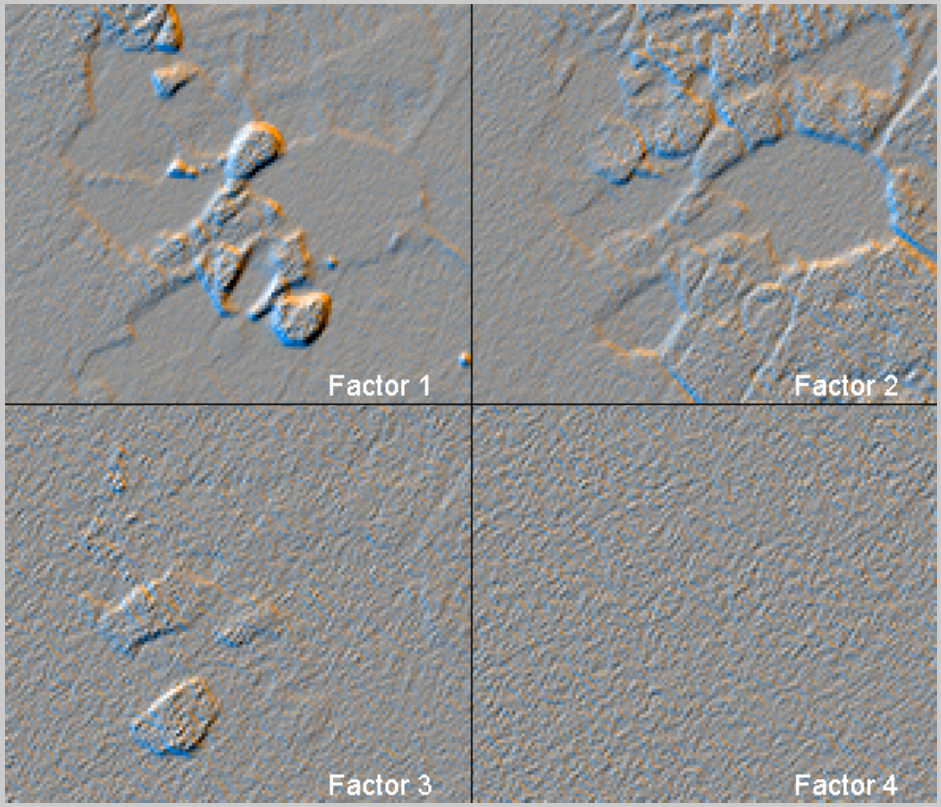
Blake, T.A., Kelly, J.F., Gallagher, N.B, Gassman, P.L., Johnson, T.J., "Passive Detection of Solid Explosives in Mid-IR Hyperspectral Images," *Anal. Bioanal. Chem.*, **395**(2), 337-348 (2009)

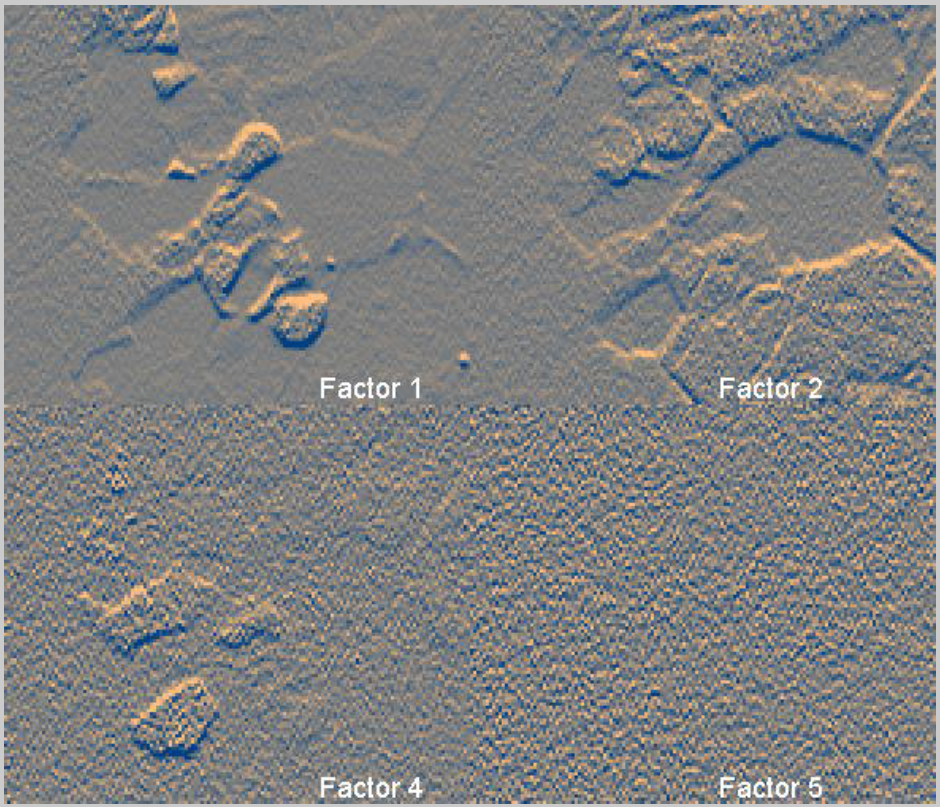
EIGENVECTOR RESEARCH INCORPORATED

MDF

$$\mathbf{X}_s = \mathbf{X}_c \;;\quad \mathbf{D}_s = \mathbf{D}_1 \;;\quad \mathbf{D}_c = \mathbf{D}_2$$

movies can be used for visualization





EIGENVECTOR RESEARCH INCORPORATED

Conclusions 1/2

- The generalized eigenvector problem (MAF, MSF, MDF…) can be reformed to a standard eigenvector problem (WPCA)
  - Whitened PCA and MAF are mathematically similar for similar definition of signal and clutter
  - PCA may be more interpretable (MAF "demystified")
- The original formulation can be memory and time intensive. The GEP problem can be reformulated using subspace projections to save time and memory, *and* focus only on eigenvalues of interest.

$$\mathbf{\Sigma}_s = \mathbf{X}_s^T \mathbf{D}_s^T \mathbf{D}_s \mathbf{X}_s \quad \Rightarrow \quad \hat{\mathbf{\Sigma}}_s = \mathbf{T}_s^T \mathbf{D}_s^T \mathbf{D}_s \mathbf{T}_s$$

$$\mathbf{\Sigma}_c = \mathbf{X}_c^T \mathbf{D}_c^T \mathbf{D}_c \mathbf{X}_c \quad \Rightarrow \quad \hat{\mathbf{\Sigma}}_c = \mathbf{T}_c^T \mathbf{D}_c^T \mathbf{D}_c \mathbf{T}_c$$

EIGENVECTOR RESEARCH INCORPORATED

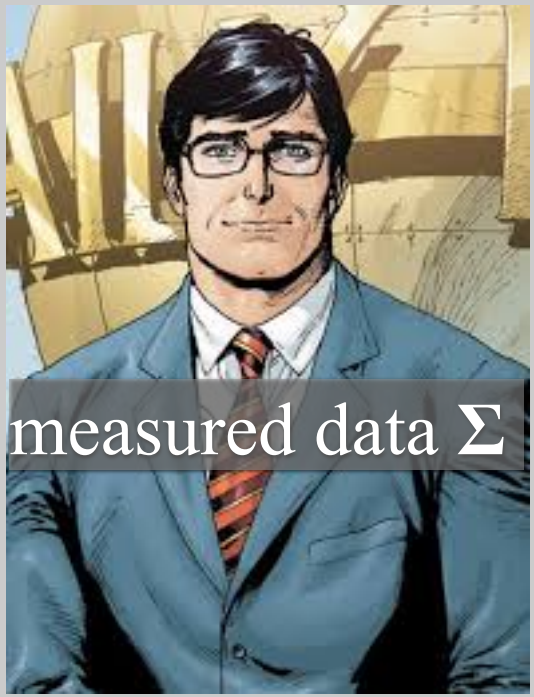
Conclusions 2/2

- The spatial operators $\mathbf{D}_i$ can be anything to focus the variance as desired
  - SavGol, wavelets, etc.
  - and can include variable operators

- Characterizing / measuring clutter is as important as characterizing the signal.

EIGENVECTOR RESEARCH INCORPORATED

WPCA and MSF partition the variance $\mathbf{\Sigma}$ through "re-weighting"



and signal of interest $\mathbf{\Sigma}_s$

to characterize clutter $\mathbf{\Sigma}_c$

EIGENVECTOR RESEARCH INCORPORATED