SAND2008-1019C

Advantages of Clustering in the Phase Classification of Hyperspectral Materials Images

Christopher L. Stork and Michael R. Keenan

Sandia National Laboratories

Albuquerque, NM

Pittcon 2008

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

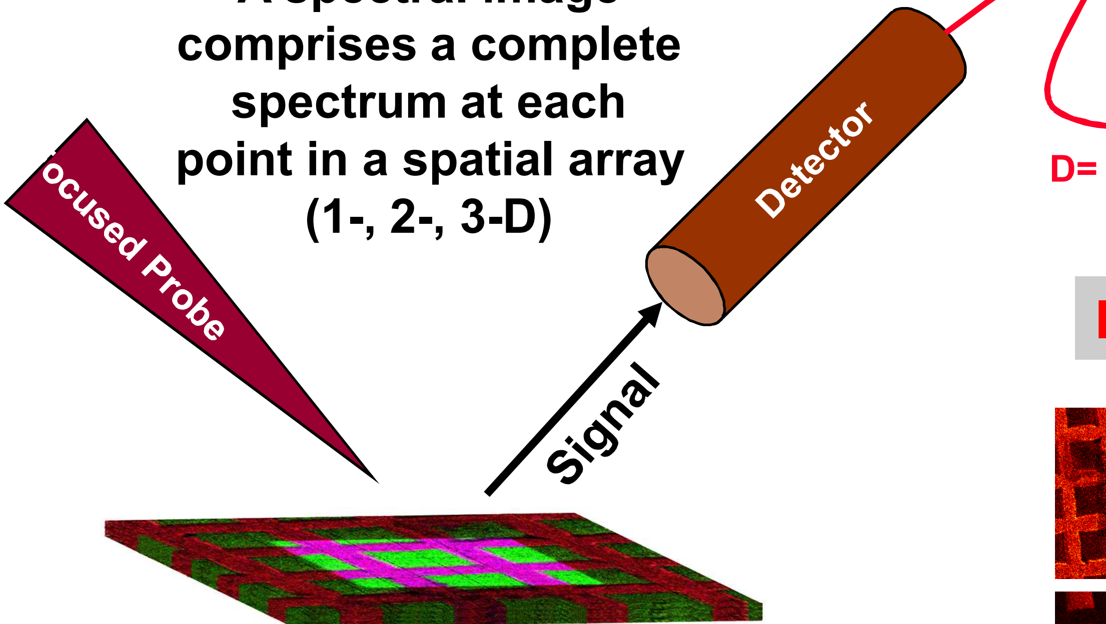
Outline

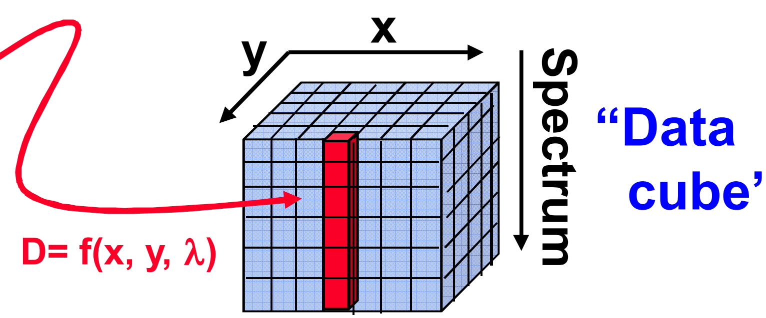
- Factor analysis (FA) of hyperspectral images
  - Can extract spectral and concentration profiles for differentiable chemical species
  - Problems in phase classification: number of components cannot exceed chemical rank
- Advantages of clustering in phase classification
  - Parsimony restriction lifted: number of components can exceed chemical rank of data
- Comparison of FA and clustering for two hyperspectral image sets
  - Solder bump data set
  - Four phase braze interface data set

Sandia National Laboratories

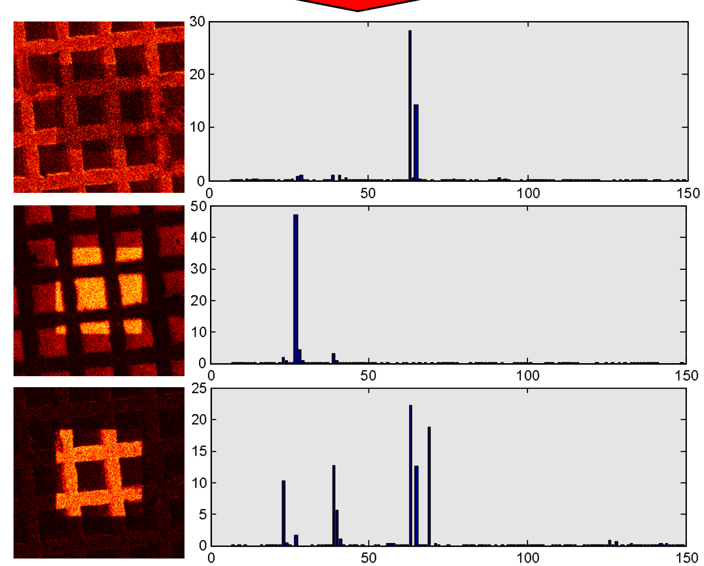
Spectral Imaging and Factor Analysis

A spectral image comprises a complete spectrum at each point in a spatial array (1-, 2-, 3-D)





Factor analysis, $D = CS^T$



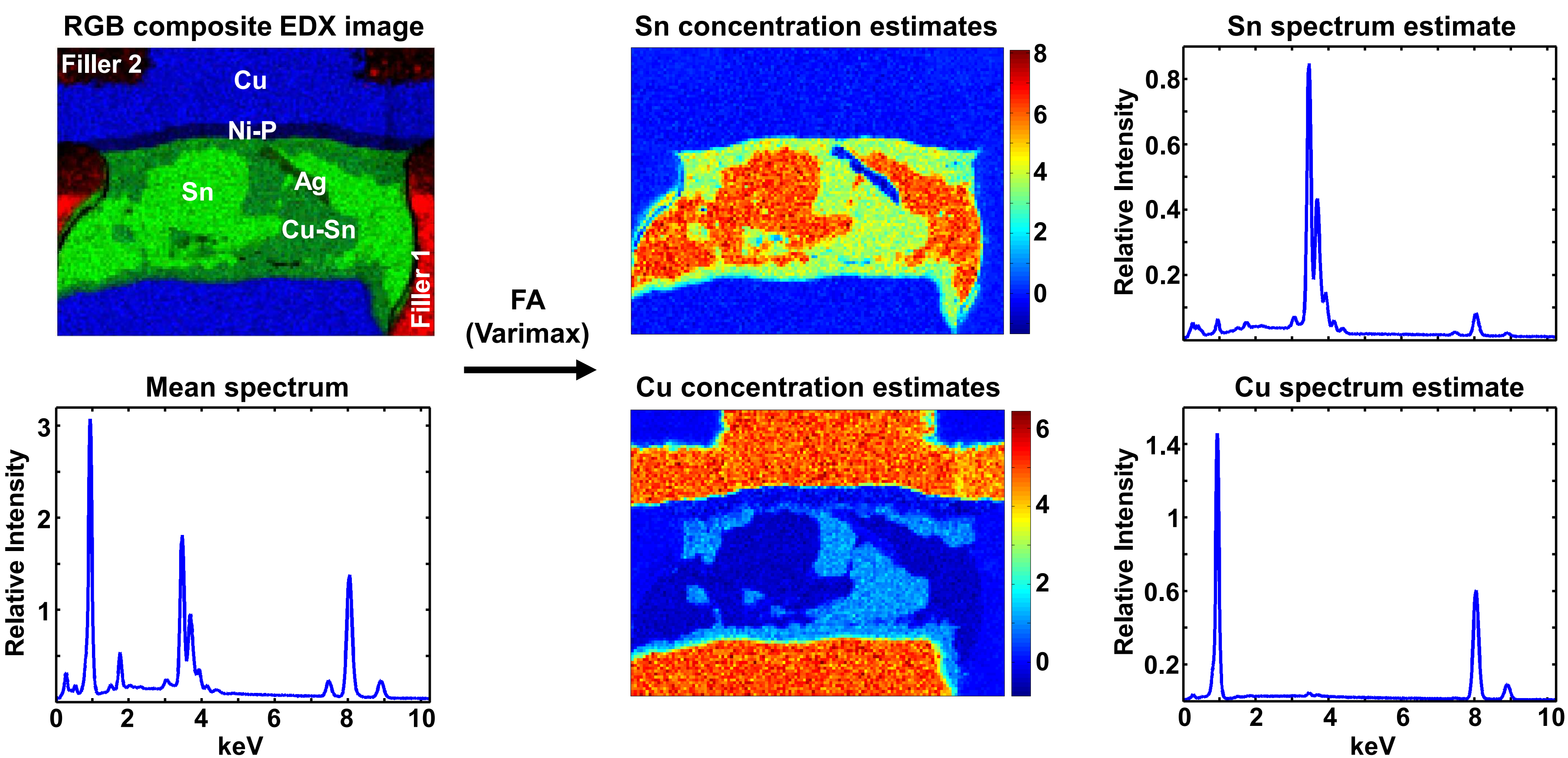
The goal of Factor Analysis it to estimate:

S, "what is it?"

C, "where is it and how much?"

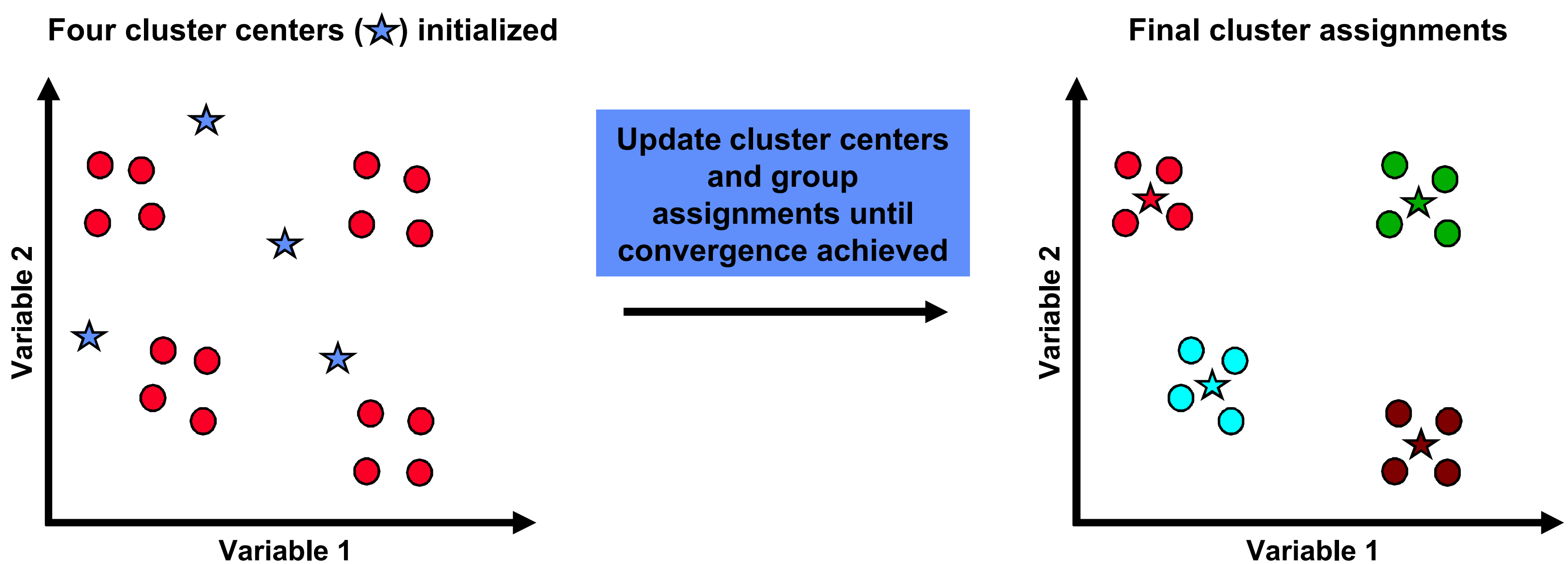
Limitations of FA for Phase Classification: Solder Bump Data

- FA parsimony restriction: number of extracted components cannot exceed chemical rank of data
  - FA fails to identify Cu-Sn intermetallic phase



Clustering: Group Data Points Based on Nearness in Multivariate Space

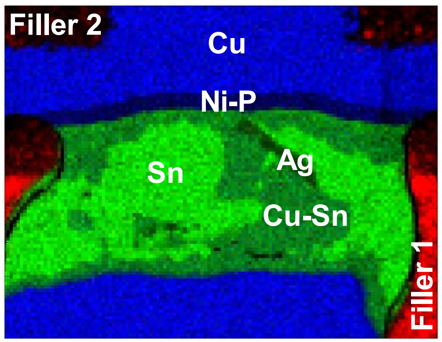
- **Partional clustering (fuzzy c-means) requires very limited prior information:**
  - **Number of clusters in data set**
  - **Initial set of cluster centers**



Cu-Sn Intermetallic Phase Identified by Clustering: Solder Bump Data

- **Clustering extracts Cu-Sn intermetallic phase: number of phases can exceed chemical rank!**

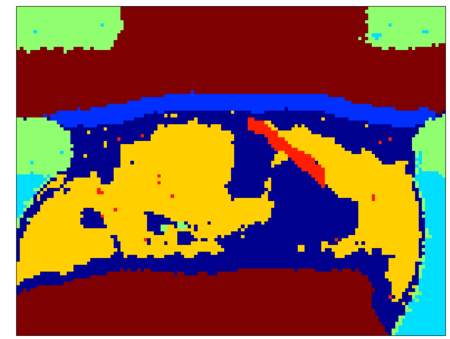
RGB composite EDX image



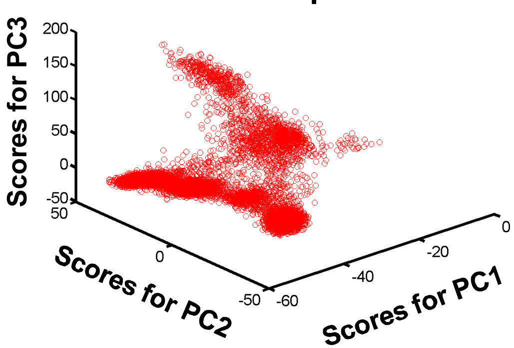
Clustering, $D = CS^T$
(Fuzzy c-means, m = 1.3)

Clustering estimates:
- **S, spectrum for each group**
- **C, group assignment for each pixel**

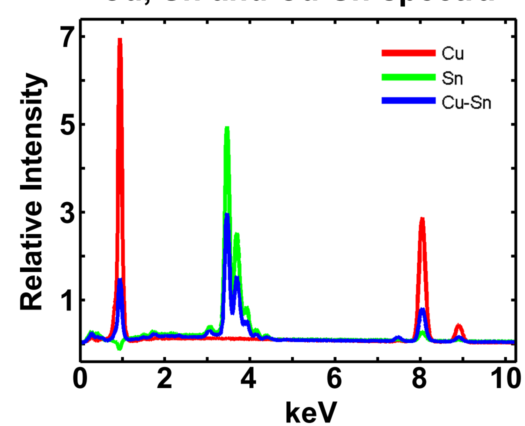
Pixel group assignments



Data points in 3-D PCA scores space



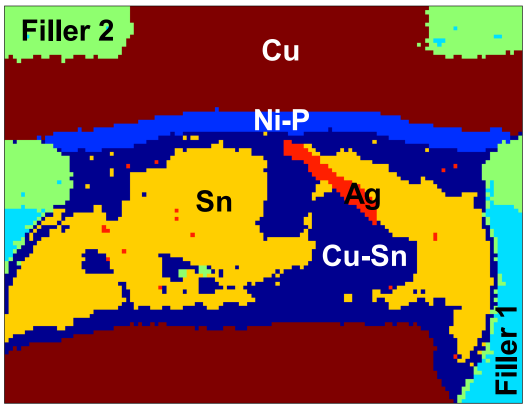
Cu, Sn and Cu-Sn spectra



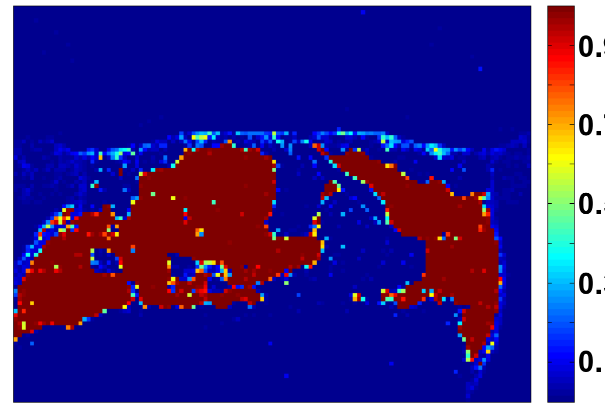
Clustering Applied to Solder Bump Data

- **Fuzzy c-means clustering provides for partial group memberships (values lie between 0 and 1)**

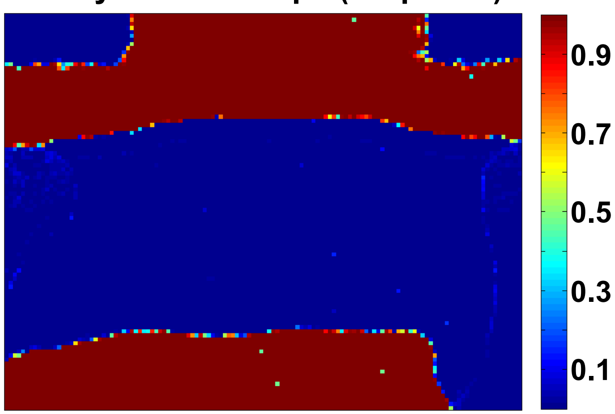
Crisp pixel group assignments



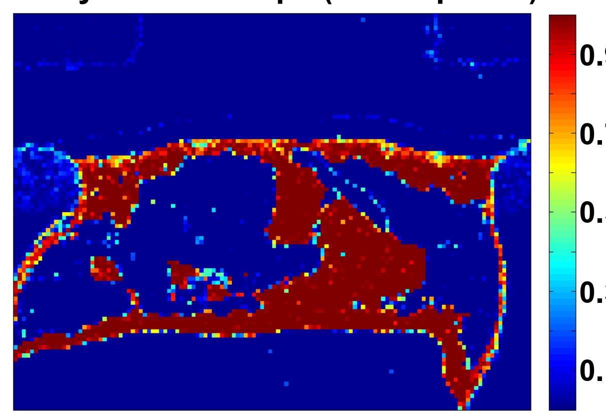
Fuzzy memberships (Sn phase)



Fuzzy memberships (Cu phase)

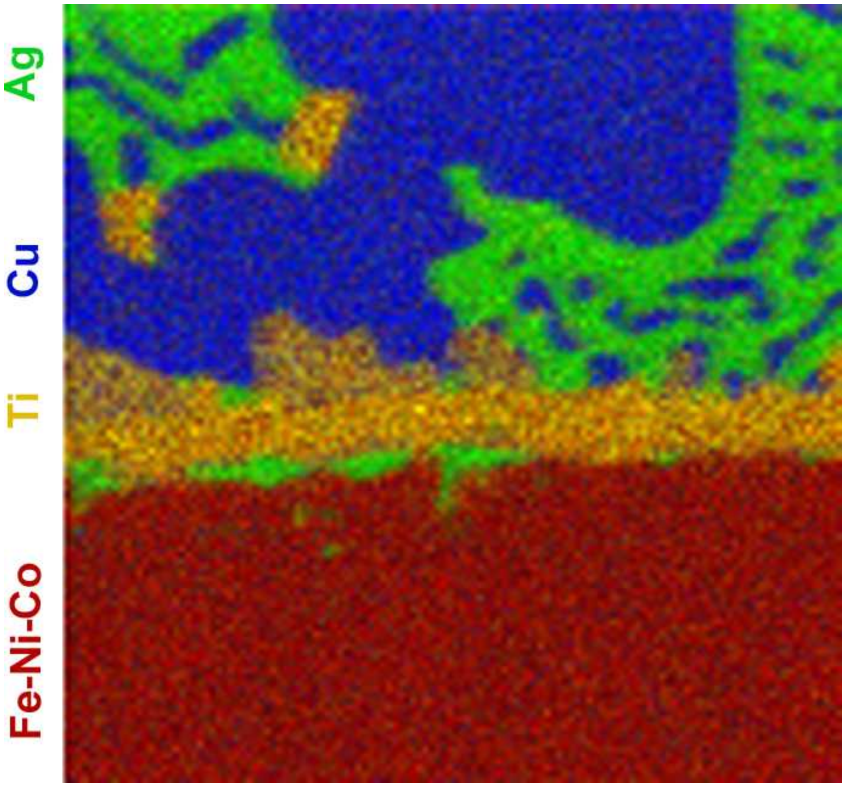


Fuzzy memberships (Cu-Sn phase)

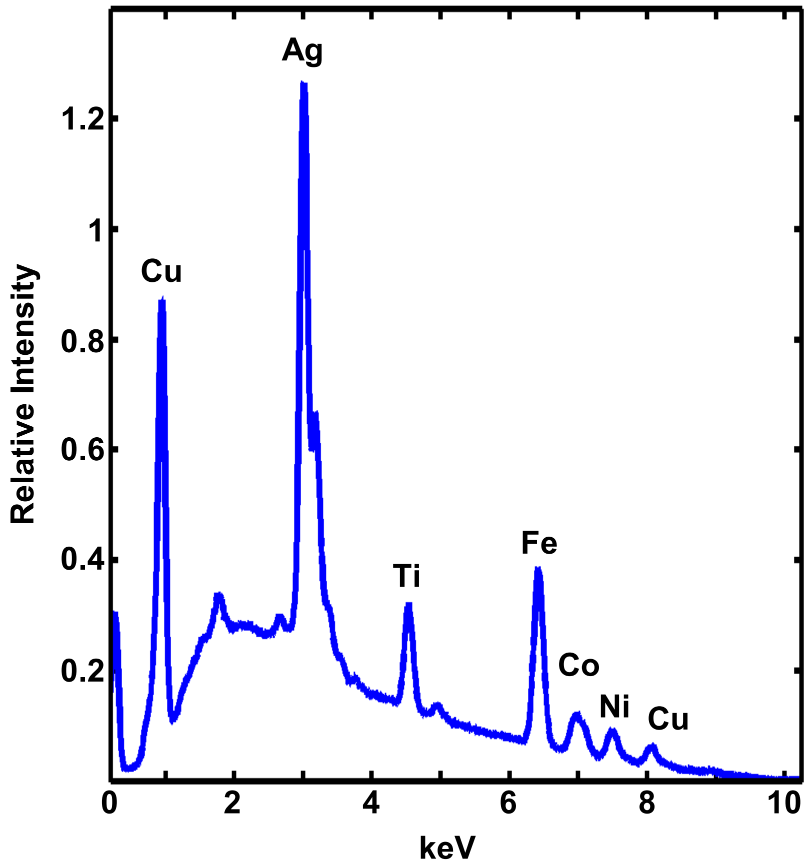


Braze Interface Data

RGB composite EDX image



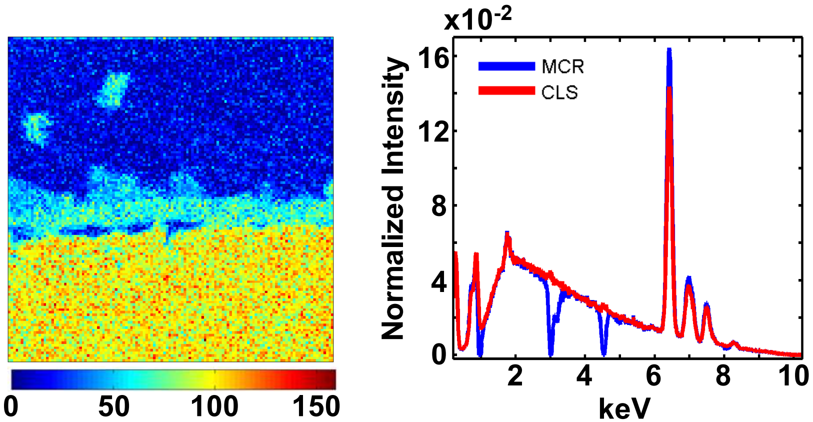
Mean spectrum



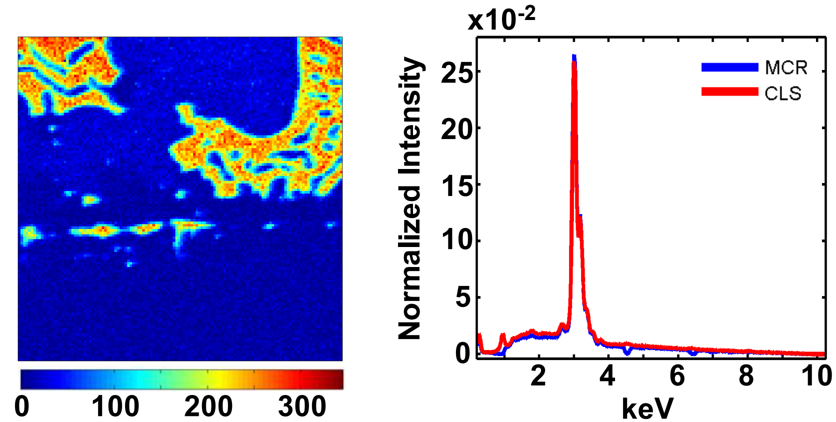
Braze Interface Data: Factor Analysis (Non-Negativity Constrained MCR)

- **Biased phase estimates obtained using FA:**

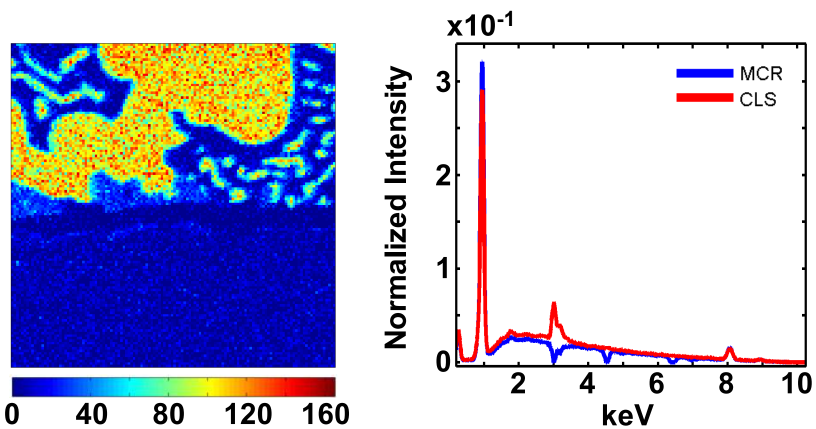
Fe-Ni-Co phase



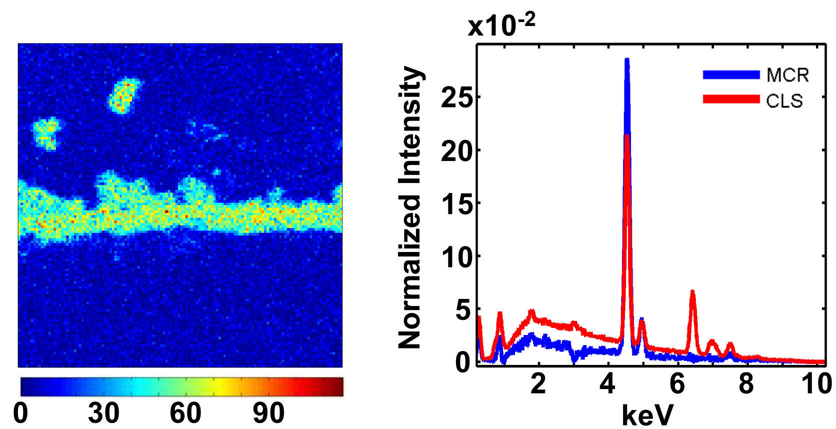
Ag phase



Cu phase

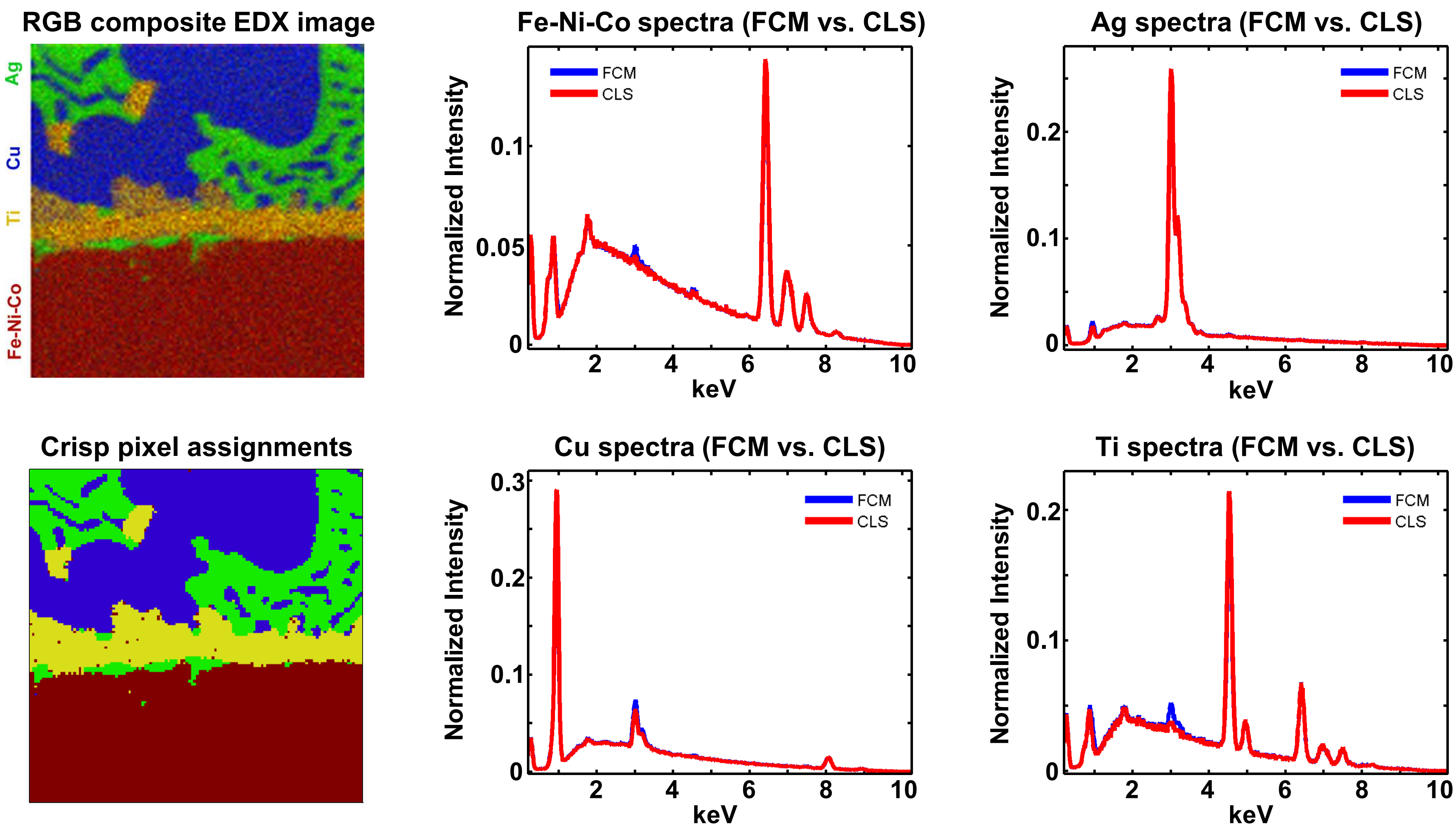


Ti phase



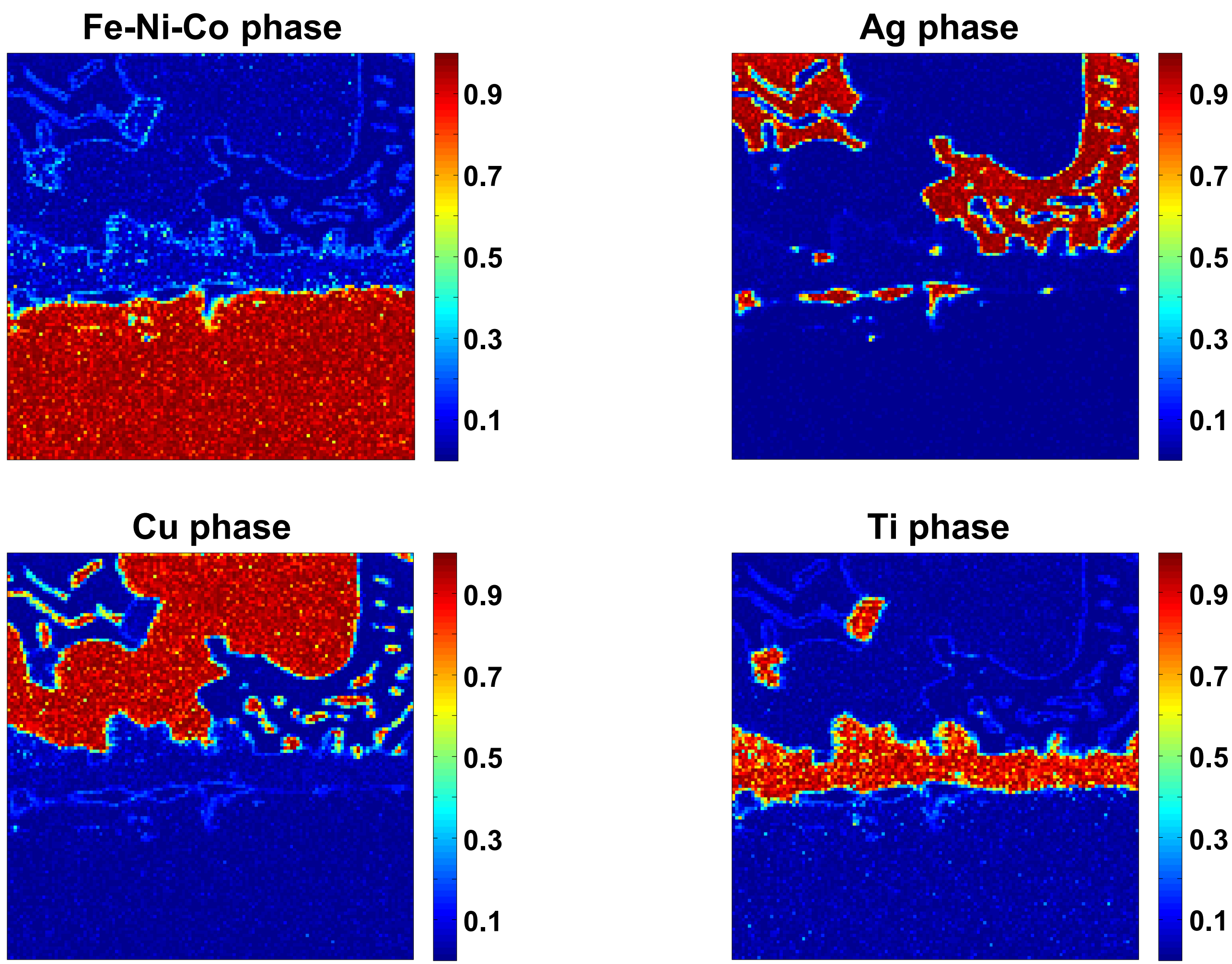
Braze Interface Data: Clustering

- **Superior phase classification results obtained using clustering (fuzzy c-means, m = 2):**



Braze Interface Data: Clustering

- **Fuzzy membership images match expectations:**



Sandia National Laboratories

Conclusions

- Clustering can solve phase problems which are intractable using FA (e.g., solder bump data)
  - In FA, number of phases cannot exceed chemical rank of data
  - In clustering, number of phases cannot exceed the number of pixels in image
- Braze interface data demonstrates that clustering can avoid bias problems encountered using FA