

Abstract No. Jaco0206

## Analysis of Soft X-ray Spectrum Images Based on Principal Components and Clusters

C. Jacobsen, M. Lerotic, and M. Feser (Stony Brook); T. Schäfer (INE-Karlsruhe); S. Vogt and J. Maser (Argonne)  
Beamline(s): X1A1 and X1A2

**Introduction:** Soft x-ray spectrum imaging involves the acquisition of a series or “stack” of images [1] across near-edge absorption features, yielding a dataset in  $(x,y,E)$  where  $(x,y)$  are position coordinates and  $E$  is photon energy. When all components with different absorption spectra in the specimen are known, curve-fitting or matrix inversion methods [2] can be used to obtain maps of component thicknesses. However, it is often the case (particularly with biological or environmental science specimens) that all of the components and their spectra are *not* known in advance. In this case, one must look to the dataset itself for clues as to the spectroscopically-significant differences in the specimen.

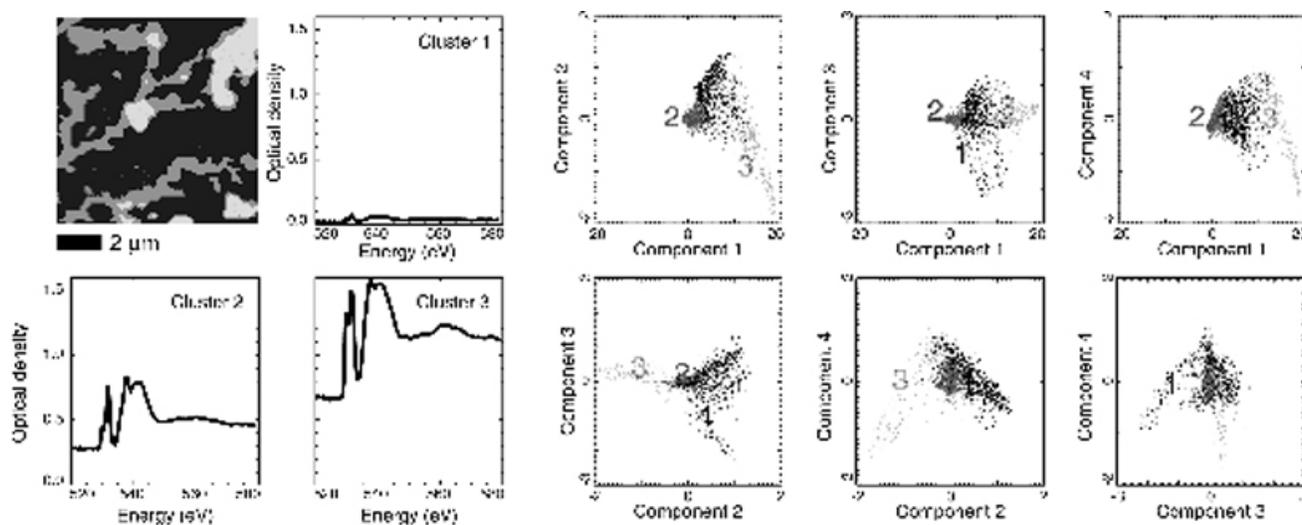
**Methods and Materials:** We have used multivariate statistical analysis methods to analyze soft x-ray spectrum image datasets. Principal component or factor analysis [3] provides a method for estimating the number of components (for example by using the factor indicator function), and determining linearly independent spectra present in the dataset. These eigenspectra can be revealing of the composition of the specimen, but do not by themselves provide spectra in a manner directly interpretable by the experimenter. We have used principal components or factors as the basis set for self-organized clustering.

**Conclusions:** Cluster analysis allows one to improve spectrum signal-to-noise by signal averaging over spatial regions with sufficiently identical spectra. It allows one to determine the real spectra of similar spatial regions in an automatic fashion, and to uncover subtle spectral changes amongst nearby regions.

**Acknowledgments:** We thank the U.S. National Science Foundation for support under grants ECS-0099893 and DBI-9986819, and the National Institutes of Health for support under grant EB00479-01A1.

### References:

- [1] – C. Jacobsen, G. Flynn, S. Wirick, and C. Zimba, *J. Microscopy* **197**, 173, (2000).
- [2] – X. Zhang *et al.*, *J. Struct. Bio.* **116**, 335 (1996); A. P. Hitchcock *et al.*, *Ultramic.* 88, 33, (2001).
- [3] – A. Osanna and C. Jacobsen, *X-ray Microscopy*, p. 350 (AIP Press, Melville, 2000); P. King *et al.*, *J. Vac. Sci. Tech. A* **7**, 3301, (1989); S. Wasserman, *J. Phys. IV* **7** (C2), 203, (1997); N. Bonnet *et al.*, *Ultramic.* **77**, 97, (1999).



**Fig. 1:** Cluster analysis results on Lu incorporated into hæmatite. At upper left we show the 3 clustered regions, with cluster 1 being the darkest and cluster 3 being the lightest; the optical density spectra for each of these regions are also shown. Cluster 2 shows a mix of Lu and hæmatite, while cluster 3 shows mostly pure hæmatite. At right we show the distribution of weights of various components for all pixels as plotted over pairs of components. These distributions show that data are clustered in a multidimensional space, so that clusters are not always readily recognizable when plotted in just two of the dimensions.