

Novel Multivariate Statistical Analysis Methods for STEM/EELS

M.C. Sarahan,^{*,**} F. de la Peña^{***}, Q.M. Ramasse,^{*} M. Walls,^{*****}

* SuperSTEM, STFC Daresbury Laboratories, Keckwick Lane, Warrington, WA4 4AD, UK

** SUPA School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, UK

*** CEA-LETI, MINATEC 17, avenue des Martyrs, 38054 GRENOBLE Cedex 9 - France

***** Laboratoire de Physique des Solides, Bât. 510, Université Paris-Sud, 91405 Orsay Cedex, France

Analytical electron microscopy is becoming increasingly limited by our ability to interpret the data obtained, rather than our ability to obtain the data. Aberration correctors enable images with such fine spatial resolution that minuscule structural features and variations are distinguishable. Similar technological breakthroughs are on the horizon for the energy resolution of EELS, opening possibilities of vibrational spectroscopy and further enhancing the value of fine-structure studies. Despite these instrumental improvements, features of interest such as point defects remain difficult to locate and characterize systematically.

Multivariate statistical analysis (MSA) is a promising development that has been increasingly applied to revealing subtle details in analytical data. By comparing correlated variations across several images or spectra, latent patterns can be extracted. Ideally, these latent patterns would directly represent physical features of interest. However, traditional methods for MSA, such as principal component analysis (PCA), have historically failed to yield patterns that are easily recognizable as any physical phenomenon [1]. In both imaging and EELS, this failure arises from the so-called orthogonality requirement among purely statistical components. As a result, MSA is more commonly applied only for denoising data and rather than exploration or interpretation of image or spectral features.

Novel methods presented here are a combination of both statistical methods and preprocessing of the data using computer vision techniques. With regards to the former, great advances in the isolation of physically realistic component EEL spectra have been realized using by relaxing the requirement of orthogonality (Independent Component Analysis, ICA) and by assuming the components must always have positive intensity (Non-negative Matrix Factorization, NMF) [2, 3]. In this work, these novel methods are compared to traditional MSA methods and to more common multiple-least squares fitting analysis methods. The unique ability to separate chemical species without any pre-determined model is shown to be particularly useful where edges from several chemical species overlap (Figs 1 and 2).

By preprocessing image data to extract peak characteristics, the statistical analysis is decoupled from image noise. Furthermore, interpretation is enhanced because the results are in the form of peak characteristics. The input data to MSA have thus far always been image intensities. As such, the output data of MSA are also intensities. These intensities can be recognizable as changes in a particular region of an image, but the quantification of those changes still requires a great deal of thought and further analysis [4]. By formulating the input to the MSA as peak characteristics (Figs 3 and 4), the output of MSA is also in the form of peak characteristics, and the relationships between changes in atomic column intensities, shapes and positions become clear without requiring comparison to average images or other components for context [5].

References

- [1] N. Bonnet, *J. Microsc.* 190 (1998) 2
- [2] N. Bonnet, D. Nuzillard, *Ultramic.* 102 (2005) 327
- [3] F. de la Pena et al., *Ultramic.* 111 (2011) 169
- [4] M.C. Sarahan et al., *Ultramic.* 111 (2011) 251
- [5] The authors are indebted to M.-H. Berger and D. Neiner for samples. This research was funded by the Engineering and Physical Sciences Research Council (EPSRC). F. de la Pena thanks the ESTEEM IP3 European Commission project, the French METSA network, the E. C. Marie Curie Action (No.MEST-CT-2004-514307) and the EOARD (Grant N° 073031).

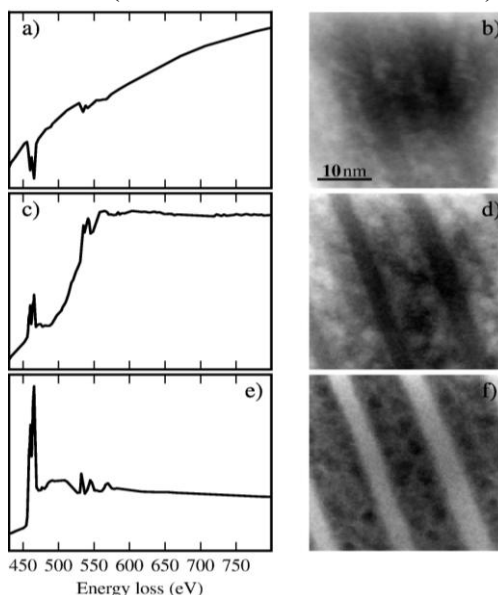


FIG. 1. PCA of a spinodally decomposed $\text{TiO}_2/\text{SnO}_2$ sample. Components are not identifiable as physical spectra.

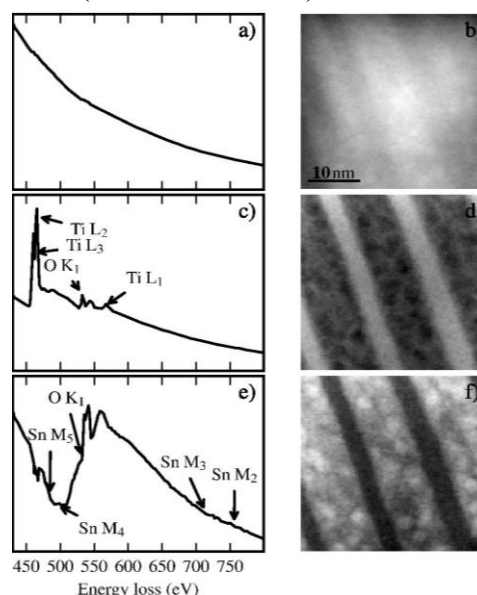


FIG. 2. ICA of the same data set. Components are readily interpreted.

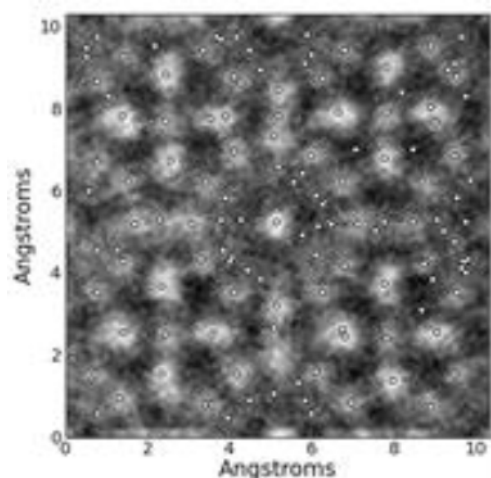


FIG. 3. Automated peak identification allows for rapid characterization. Peaks are indicated by white circles superimposed on the image.

X (Å)	Y (Å)	Height	Orient.	Ecc.
5.847639	0.482833	126	-0.407	0.000
0.751073	0.536481	110	0.307	0.000
3.648069	0.536481	211	-0.129	0.002
9.281116	1.341202	173	0.731	-0.005
2.414163	1.448498	227	-0.753	-0.008
4.077253	1.502146	158	-0.597	0.000
7.725322	3.487124	211	-0.111	0.008
0.697425	3.540773	128	-0.625	-0.001
2.521459	3.540773	216	-0.093	0.010

FIG. 4. Analysis of generic peak characteristics such as height, orientation and eccentricity simplifies interpretation of MSA results – characteristics in, characteristics out without any complicated interpretation of intensities.