

Fuzzy Clustering to Merge EDS and EBSD Datasets with Crystallographic Ambiguity

Chad M. Parish^{1*}

¹ Fundamentals of Radiation Effects Group, Oak Ridge National Laboratory, Oak Ridge, TN, USA.

* Corresponding author: parishcm@ornl.gov

A common problem encountered in electron backscatter diffraction (EBSD) analysis is that phases with similar crystal structures will differentiate poorly using standard EBSD indexing algorithms. A classical example might be Ni vs. Cu, which have the same space group (Fm-3m) and similar lattice parameters.

The obvious solution to this problem is to simultaneously acquire X-ray energy dispersive spectroscopy (EDS) spectrum images (SIs) during the scan, and then "somehow" merge the two datastreams [1, 2]. This is not as simple as it sounds. For instance, absolute counts in an X-ray line, or ratios of X-ray lines, might be used to determine which pixels will be allowed to have a certain phase in their phase list for EBSD indexing. This paradigm can work, but requires significant human input, is vendor-specific, and can be *ad hoc* and difficult to automate.

Here, I propose a nearly human-input-free method that uses fuzzy C-means (FCM) or K-means clustering. FCM was proposed for analysis of SEM-EDS SIs by Stork and Keenan [3]. Stork and Keenan proposed an excellent computational trick in which cluster analysis is performed on the reduced-rank, Poisson-noise-scaled [4] matrix that is the product of the left singular vectors \mathbf{U} and singular values' diagonal matrix Σ , which I will call \mathbf{L} where $\mathbf{L}=\mathbf{U}\Sigma$, from the rank- p singular value decomposition $\mathbf{U}\Sigma\mathbf{V}$. Stork and Keenan found that FCM of \mathbf{L} with a fuzziness coefficient ≈ 1.3 differentiates the phases in an EDS SI quite well; the cluster centers can then be easily converted to the spectral endmembers. This FCM provides crisp phase maps without the nonnegativity artifacts of multivariate curve resolution or the parsimony restriction of SVD. In this work, I expand on Stork and Keenan's innovation by appending a "one-hot" matrix of EBSD-derived phase assignments to the matrix \mathbf{L} , and then performing FCM or K-means clustering (using Python 3.7 and scikit-fuzzy or scikit-learn) to add crystallographic information.

Here, sintered TiB_2 [5] (hexagonal) that is contaminated with cubic $\text{Ti}(\text{C},\text{N},\text{O})$ and $(\text{Zr},\text{Al},\text{Ti})$ oxides was analysed by simultaneous 20 kV EDS+EBSD in a Tescan MIRA3 SEM equipped with Oxford Symmetry CMOS EBSD and Oxford UltimMax170 EDS. 80×60 pixels were acquired at 500 nm pitch, with phases indexed as "hexagonal" or "cubic," and EDS qualitatively found regions of Ti-B, Ti-CNO, and Ti-Zr-Al-O (Figure 1). Si is colloidal silica polishing media in the pores.

FCM, by itself, found four phases: Ti-B matrix, $\text{Ti}(\text{CNO})$, $(\text{Zr},\text{Al},\text{Ti})$ -oxide, and a "boundary" $(\text{Zr},\text{Al},\text{Ti})$ oxide (due to finite EDS spatial resolution), Figure 2. Appending a one-hot three-column matrix to \mathbf{L} , with a single "1" value in the 0th column for unindexed, a "1" in the 1st column for "hexagonal," and a "1" in the 2nd column for "cubic," and "0"s in all other positions. If \mathbf{L} is first scaled to a maximum absolute value of 1.0, K-means clustering of this appended matrix provides a crisp phase map that differentiates both the crystallographic and chemical identity of each phase, Figure 3.

Overall, this is an automatable method to merge EBSD and EDS information, and the only human input

needed is the number of clusters to find and implemented in freely-available Anaconda Python 3.7 [6].
References:

- [1] MM Nowell et al., Materials Science Forum, **509** (2006), p. 11.
 [2] CM Parish et al., Proc. Microsc. Microan. **19(S2)** (2013), p. 746.
 [3] CL Stork and MR Keenan, Microsc. Microan. **16** (2010), p. 810.
 [4] MR Keenan in "Techniques and Applications of Hyperspectral Image Analysis", ed. HF Grahn and P Geladi (Wiley, Chichester) p. 89.
 [5] WG Fahrenholtz et al., Scripta Materialia, **129** (2017), p. 94.
 [6] Research supported of US Department of Energy, Office of Science, Fusion Energy Sciences (via an Early Career Award), under contract number DE-AC05-00OR22725. Sample courtesy G. Hilmas & W. Fahrenholtz, Missouri University of Science and Technology. I acknowledge the Python toolboxes: scikit-learn, scikit-fuzzy, numpy, scipy, h5py, pycroscopy, pyUSID, and matplotlib-scalebar.

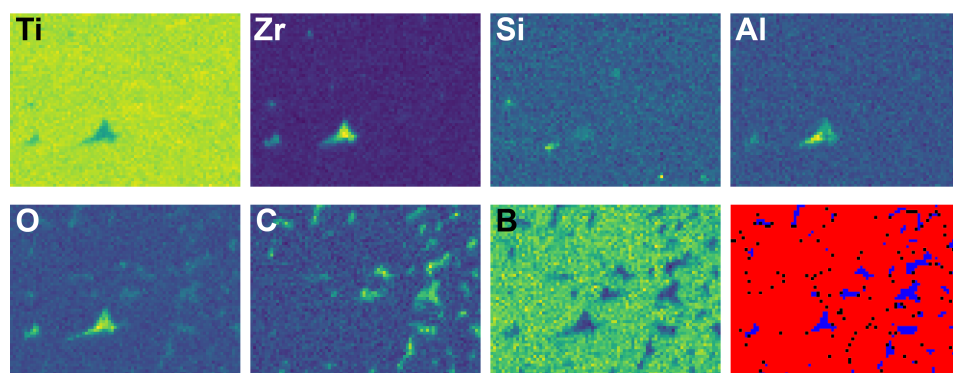


Figure 1. (left) X-ray maps and EBSD phase assignments from the TiB₂-Ti(CNO)-(Zr,Al)-oxide composite.

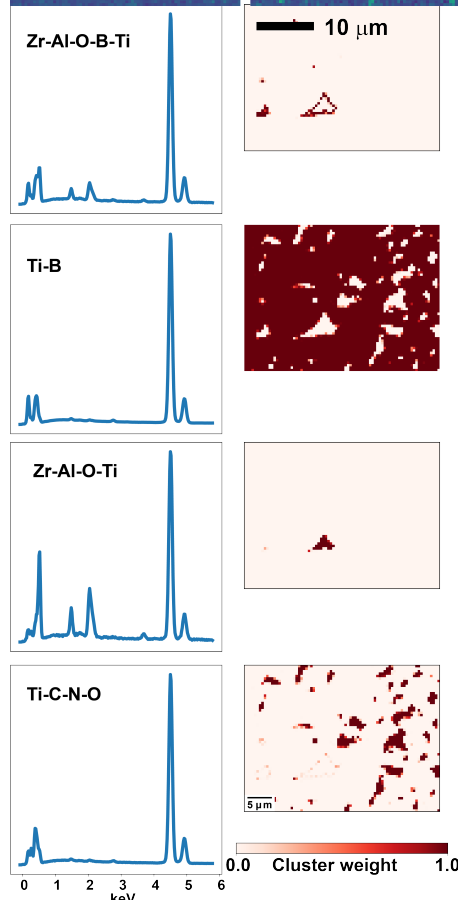


Figure 2. (left) Fuzzy C-means clustering of the EDS spectrum image into four clusters. The clusters, qualitatively, are the matrix and two types of precipitates (Ti-rich and Zr-rich), with the Zr-rich being broken up into "thin" and "thick," where the thin Zr-rich precipitate shows more matrix contribution (due to poor spatial resolution at 20 kV and 70° tilt).



Figure 3. (above) K-means cluster analysis of the merged EDS and crystallographic information. Black: Matrix, unindexed. Red, TiB₂ hexagonal. Blue: Ti(CNO), cubic. White: (Ti,Zr,Al) oxide, cubic, with strong B signal. Yellow: (Ti,Zr,Al) oxide, cubic, weak B signal.