2023
doi:10.1017/S1431927615010892 Paper No. 1010
Microsc. Microanal. 21 (Suppl 3), 2015

Phase Analysis of Large EDS Datasets with Matlab

Richard C. Hugo[1], Steven Bernsen [2], Kathy Breen[3], and Alex Ruzicka[1]

[1.] Portland State University, Department of Geology, Portland, OR, USA
[2.] New Mexico Inst. of Mining and Technology, Dept of Earth and Env. Sci., Socorro, NM, USA
[3.] US Geological Survey, Oregon Water Resources Center, Portland, OR, USA

Today's SEM experimentalist can acquire prodigious amounts of data in short amounts of time. High-stability FEG-SEMs equipped with high-throughput SDD detectors are widely available. Commercial microanalysis systems control motorized specimen stages, acquiring hundreds of spatially aligned fields of view (FOV) in an overnight or weekend SEM session [1]. The resulting mosaic of EDS datacubes can easily comprise hundreds of gigabytes.

Commercial phase analysis software is optimized for rapid, automated data exploration. In real time, powerful, proprietary algorithms reduce spectral dimensions via principal components analysis (PCA) and/or automated element identification, and identify and quantify phases via spectral and/or cluster analysis [2-4]. However, dataset size is limited by computer RAM so that algorithms can be completed in real time. Further, because PCA and cluster results are unique to a given field of view, phase analysis results cannot be applied to other specimens.

We have developed a set of software tools in Matlab that extend datacube-processing capabilities to datasets much larger than available RAM. Our process maintains a database of individual datacubes stored on disk. Subsets of this database are temporarily loaded into RAM, processing templates are applied, and the result from each FOV is added to an output database. The output database takes the form of a tiled RGB TIFF image, which can itself be retrieved and re-processed in RAM-sized chunks by Matlab, or viewed with various image processing packages. A distinct advantage of our approach is that templates created for a single dataset can be applied to subsequent datasets for direct comparison of similar specimens.

The workflow is described in Figure 1. Input data include a matched suite of BSE images, RAW datacubes, and metadata files which describe the location and dimensions of each datacube. Our data were collected with Oxford INCA software and exported to RAW form with *INCABatch*. Data flow between hard drive and RAM is managed by Matlab's *blockproc* function. Dimension-reducing options currently include PCA, spectral color mapping, and element ROI coloring. Cluster algorithms include k-means and supervised nearest-neighbor. Future capabilities may incorporate 4-channel output, hierarchical clustering, and other algorithms. After code review we will release all source code to the Matlab File Exchange website for others to use and improve [5].

References:
[1] S Burgess *et al*, Microsc. Microanal. **20-Suppl 3** (2014), p. 640.
[2] PG Kotula *et al*, Microsc. Microanal. **9(1)** (2003), p. 1.
[3] T Nylese and R Anderhalt, Microsc. Today **22(2)** (2014), p. 18.
[4] DS Bright and DE Newbury, J. Microscopy **216** (2004), p. 186.
[5] The authors acknowledge PSU Faculty Enhancement funding to Hugo. The specimen was analyzed at the PSU Center for Electron Microsc. and Nanofab. with funding from NASA grant NX10AH33G.

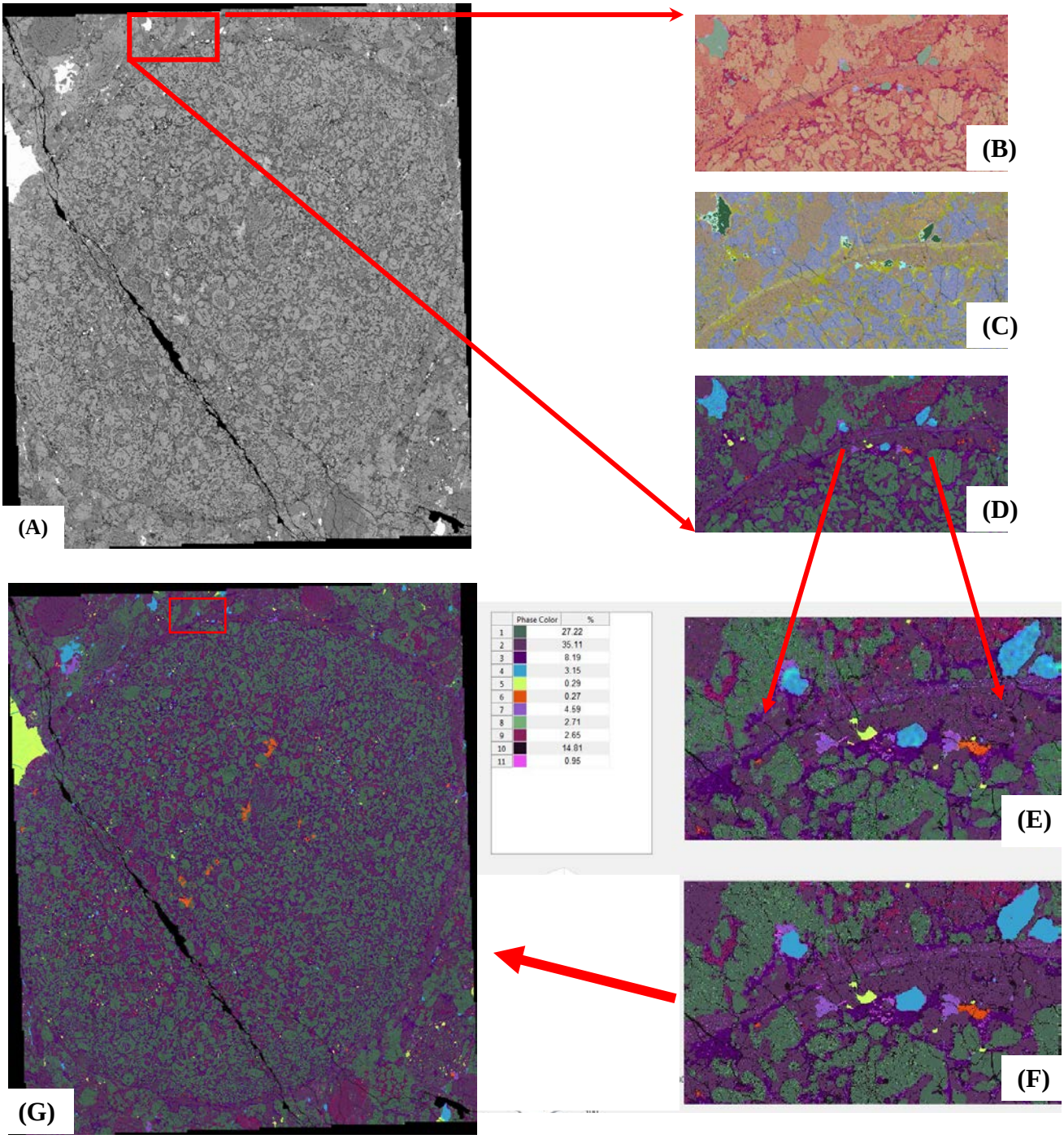

Figure 1. Workflow for a complete phase analysis of a 10k x 12k EDS mosaic. The subject is a ϕ1.6cm meteorite igneous inclusion. (A) BSE image mosaic is stitched with a cross-correlation algorithm. A representative FOV is then chosen from the BSE mosaic and dimension reducing algorithms are compared. The resulting datacube has three energy channels and is visualized as an RGB image. Energy reducing options include (B) PCA, (C) spectral color mapping, and (D) traditional energy-ROI coloring. After applying the desired dimension reducing template to the full dataset, a new representative FOV (E) is chosen for cluster analysis and a partial phase image is generated (F). Finally, the clustering template is applied to the full RGB dataset and the full phase image (G) is written in tiled TIFF format.

https://doi.org/10.1017/S1431927615010892 Published online by Cambridge University Press