Indexing of Large SEM Diffraction Data Sets Using a Dictionary Approach

Marc De Graef[1]

[1] Dept. of Materials Science and Engineering, Carnegie Mellon Univ., Pittsburgh, PA, USA

Recent years have seen a significant increase in the use of electron diffraction modalities in a scan-ning electron microscope (SEM) to cover large areas of complex microstructures, both in advanced engineering materials and natural geological samples. In addition to the well established Electron Back-Scatter Diffraction (EBSD) technique, Electron Channeling Patterns (ECP) and Transmission Kikuchi Diffraction (TKD) have seen increased use as well as progress in the quantitative modeling of the diffracted intensities. As shown in [1], these three techniques can be described theoretically in a single framework consisting of a Monte Carlo simulation followed by generation of a master pattern using the dynamical scattering theory. Combining this model with a uniform sampling of orienta-tion space $SO(3)$ as well as a geometrical detector model allows us to generate a dictionary of all possible diffraction patterns for the selected angular step size in orientation space. These dictionar-ies can then be used in combination with a similarity metric to index large scale experimental data sets.

The work flow for indexing of SEM-based diffraction data sets is shown in Fig. 1; starting from a given crystal structure and knowledge of the sample orientation and microscope accelerating voltage, a Monte Carlo and dynamical master pattern calculation are carried out. The relevant detector parame-ters (e.g., the pattern center for EBSD) can be determined by fitting a high quality reference diffraction pattern. For a given angular step size in orientation space, all unique diffraction patterns are computed and stored as a dictionary. Both experimental and simulated patterns are then pre-processed, which typically involves a high-pass filter and adaptive histogram equalization [2], to flatten the pattern inten-sity histogram. A similarity metric (dot product or mutual information) is then employed to compare each experimental pattern to all dictionary patterns; this computation is performed on a GPU. The dictionary pattern with the highest similarity value is taken to be the best match and the corresponding orientation is assigned to that experimental pattern. Since the orientations are taken from a discrete grid, a subsequent refinement step can be carried out to maximize the similarity metric (using the Bound Optimization by Quadratic Optimization algorithm [3]). For large area EBSD or TKD scans, the variation of the pattern center can be included as a post-processing step, and a second refinement then leads to the final orientation assignment. Numerical tests on synthetic data sets revealed that the orientational accuracy can be better than $0.1°$ when accurate detector parameters are known.

Since a typical dictionary indexing run produces a list of top matches for each experimental pattern, one can use this list to perform a statistical analysis of the neighborhood of the sampling point. In particular, the Orientation Similarity Map depicts how many of the top N matches each point has in common with its four nearest neighbors; such a map, which can not be obtained using the standard Hough-transform indexing techniques, reveals details of the microstructure, including grain bound-aries. The dictionary indexing (DI) technique is robust against pattern noise, and is also able to index overlapping patterns which frequently occur near grain boundaries or in materials with a substructure (e.g., nano-twins) on a length scale smaller than the size of the interaction volume. To handle such cases, one can generate composite master patterns, which consist of the original master pattern with a superimposed second master pattern that has been rotated by the twinning operation for the relevant twin system; in fcc metals, for instance, this would be a rotation of $60°$ around the $[111]$ axis.

We will discuss a number of examples based on experimental data sets: EBSD on a heavily deformed shot-peened aluminum sample; TKD for a nano-twinned diamond from a meteorite sample; EBSD from a de-wetted Ag film on a Ni substrate; and a large area EBSD scan of low angle boundaries in an Fe-3% Si sample. In all cases, we will compare the results from the DI approach with traditional in-dexing results and point out cases where the DI approach is able to provide results when Hough-based indexing fails.

Finally, we will briefly explore the use of machine learning tools (convolutional neural nets) to perform the indexing of ECPs; while the indexing rate can be fairly high (several hundred patterns per second), it is not clear at this point in time whether or not the somewhat lengthy training step of the neural net will be able to compensate for the relatively slow processing speed of the dictionary indexing. Exam-ples of CNN-based indexing of synthetic ECP data sets will be provided [4].

References:

[1] E. Pascal *et al.*, Ultramicroscopy, **187** (2018) p. 98
[2] Pizer, S.M. *et al.*, Computer Vision, Graphics and Image Processing, **39** (1987) p. 355
[3] Powell, M.J.D., Technical Report, University of Cambridge (2009)
[4] The author acknowledges ONR Vannevar Bush Faculty Fellowship support (N00014-16-1-2821) as well as the computational resources of the Materials Characterization Facility at CMU, grant MCF-677785.
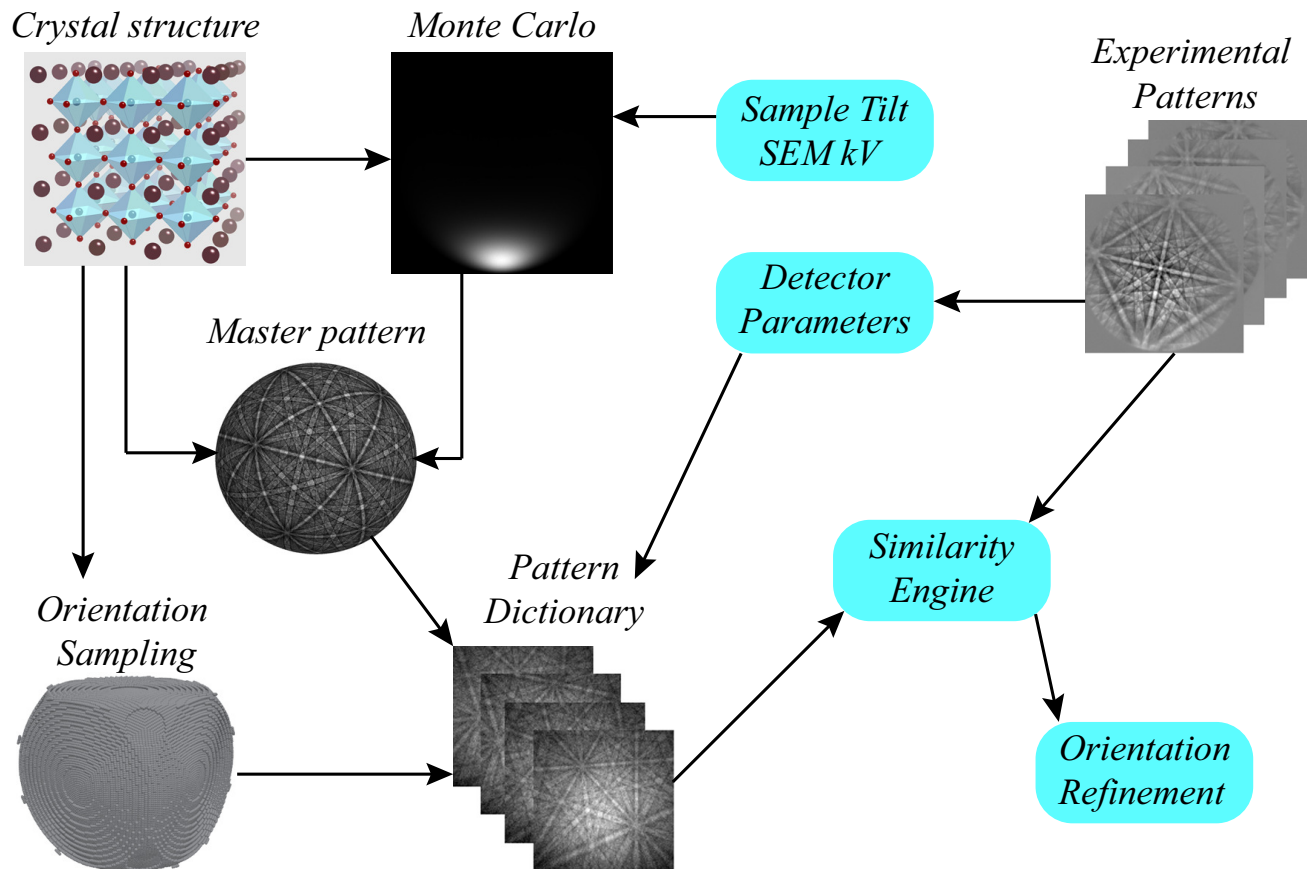
Figure 1. Schematic diagram of the dictionary indexing and orientation refinement work flow.