FEL Data Analysis Methods and Tools at NRC KI

Anton Teslyuk, Sergey Bobkov

Main areas of interest

- Diffraction images sorting and classification
- Diffraction modeling
- Data management
- 3D structure reconstruction

Diffraction Image Classification

Remove:

- Black images with background scattering
- Images from multiple particles
- Images from water droplets or some contaminant

Save:

 Images from the particle of interest



Feature vectors for image description





• Angular cross-correlation function (CCF)

- Normalization:
- Fourier decomposition:
- Feature vector for every image (m≈50):

$$F = (\bar{C}^1, ..., \bar{C}^m)$$

- Set of feature vectors was used as input for general purpose classification algorithms:
 - Principle component analysis
 - Support vector machine

Results on simulated and experimental data



SPSIM-CUDA Simulation software

Main features:

- Random particle orientation
- Beamstop and modular detector structure
- Random particle position
- Calculation on GPU (70x faster)
- More output formats
- Available at Github: <u>https://github.com/sergebb/spsi</u> <u>m_cuda</u>

S. Bobkov, A. Teslyuk, V. Ilyin. Procedia Computer Science, 66, 158-165(2015)



CXI Storage Database

- In-house software for diffraction datasets and related metadata storage and management
- Supports CXI format (derived from HDF5) used at cxidb.org
- Based on software stack:
 - GridFS + MongoDB as database layer
 - RESTful web API (Flask/Nginx) as API layer
- Currently used by our image classification tools (additional image metadata)

Structure reconstruction

- Own realization of EMC algorithm (Loh and Elser, 2009)
 - Diffraction patterns orientation determination
 - 3D structure reconstruction
 - in active development, planning release date May 2017