Department of Physics
Sungkyunkwan University

Weekly Report

Name : Donghun Jung
Number : 2020312141
E-mail :

1 Summary

1. TEM diffraction data can be regarded as two dimensional structure. Then, Finding its basis can be used to find Calibration factor and lines for intensity profile.

2. Machine Learning(SGD-momentum model) is used to find exact basis vectors. There is 0.5px difference between the real TEM data and the calculated one.

3. I'm working on labeling all the (reciprocal) lattice points with Miller indices and coefficients of linear combination automatically. Integer programming, including Lovasz' Basis Reduction Algorithm, may be helpful.

2 Silicon Diffraction Pattern

Silicon has diamond crystall structure which consists of two interpenetrating face-centered cubic lattice. That is, it can be regarded as a fcc lattice with the two point basis 0 and $\frac{a}{4}(\hat{x}+\hat{y}+\hat{z})$ where $a = 5.43[\dot{A}]$ is lattice constant. Then, its corresponding reciprocal lattice vector is given by

$$\vec{G} = \frac{2\pi}{a}(k\hat{x} + l\hat{y} + m\hat{z}) \qquad (1)$$

where k, l, m are all even or odd. Zone axis is given by [011], then $y + z = 0$. Then, electron diffraction pattern looks like the following figure.

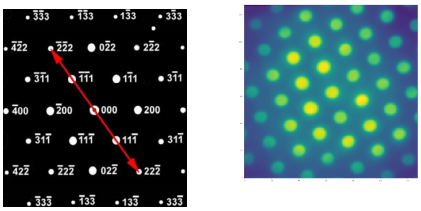


Figure 1: (a) Indexing of single crystal silicon diffraction pattern, with the beam parallel to the [011] zone axis (b) Real TEM data

As shown in the figure, the diffraction pattern is just a planar view of the reciprocal lattice structure. Then, it can be regarded as two dimensional structure with two basis vectors. Using this basis vector, we can calculate the calibration factor.

3 Calculating Calibration Factor

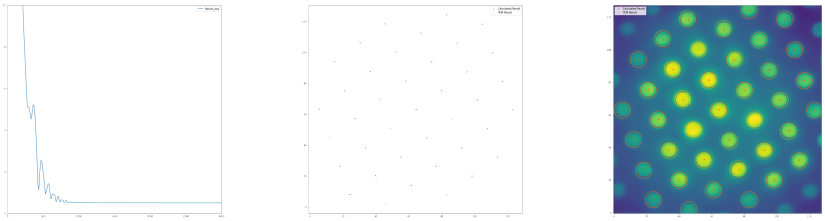
3.1 Finding Basis Vectors

All the reciprocal lattice points can be represented by linear combination of two basis vectors with integer coefficients. That is,

$$\begin{pmatrix} x_{mn} \\ y_{mn} \end{pmatrix} = (\vec{a_1}\vec{a_2}) \begin{pmatrix} m \\ n \end{pmatrix} + \begin{pmatrix} x_{00} \\ y_{00} \end{pmatrix} \qquad (2)$$

where m, n are integer. x_{mn}, y_{mn} represent the position of the lattice point whose linear combination coefficients are m, n. x_{00}, y_{00} must be the center, that is, point basis.

Suppose the position and the coefficient of the linear combination, of all the lattice points including the center, are well defined. At here, the brightest lattice



(a) (b) (c)

Figure 2: (a) Cost function (b), (c) Calculated Result and Real Data

point is chosen as a center point for convenience.

$$\begin{pmatrix} x_{00} & x_{01} & \cdots & x_{mn} \\ y_{00} & y_{01} & \cdots & y_{mn} \end{pmatrix} = (\vec{a_1}\vec{a_2}) \begin{pmatrix} 0 & 0 & \cdots & m \\ 0 & 1 & \cdots & n \end{pmatrix} + \begin{pmatrix} x_{00} \\ y_{00} \end{pmatrix} \tag{3}$$

Comparing this to the real data from TEM, we can define cost function to find the basis vectors using Machine Learning. Let X_{mn}, Y_{mn} denote the position of the lattice point whose linear combination coefficients are m, n. Subtracting them,

$$\begin{pmatrix} X_{00} & X_{01} & \cdots & X_{mn} \\ Y_{00} & Y_{01} & \cdots & Y_{mn} \end{pmatrix} - \begin{pmatrix} x_{00} & x_{01} & \cdots & x_{mn} \\ y_{00} & y_{01} & \cdots & y_{mn} \end{pmatrix} \tag{4}$$

calculating norm of each vector,

$$\begin{pmatrix} N_{00} & N_{01} & \cdots & N_{mn} \end{pmatrix} \tag{5}$$

and finally take mean value $\bar{N}$. The more accurate $\vec{a_1}, \vec{a_2}$ are, the lower $\bar{N}$ is. It becomes four variable fitting problem.

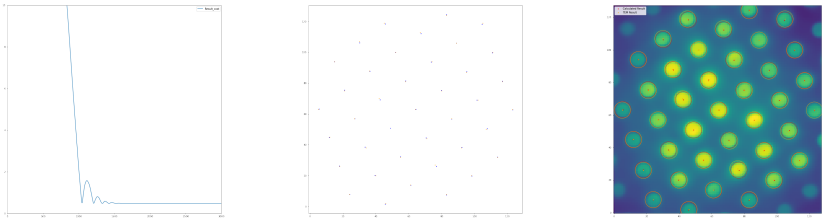
$$\vec{a_1} = \begin{pmatrix} a \\ b \end{pmatrix} \vec{a_2} = \begin{pmatrix} c \\ d \end{pmatrix} \tag{6}$$

At here, SGD-momentum model is used. The following figures show the result. the average distance between the real data and the data calculated by basis is 0.5[px]. As learning proceeds, cost function decreases. See Figure 2.

3.2 Finding Center

Suppose that the position of the center point is undetermined. In a similar manner, basis vectors and center position can be found. It is just six variable fitting problem.

Following figures show the result. the average distance between the real data and the data calculated by basis is 0.48[px]. The calculated center is $(x_{00}, y_{00}) = (48.812, 50.814)$ while the real data is $(X_{00}, Y_{00}) = (49, 51)$. See Figure 3.



(a) (b) (c)

Figure 3: (a) Cost function (b), (c) Calculated Result and Real Data

3.3 Calculating Calibration Factor

From the norm of the basis vectors or an arbitary vector obtained by linear combination of basis vectors, calibration factor can be calculated. As shown in the Figure 4, $\vec{a_1}, \vec{a_2}$ correspond to the Miller indice $(1\bar{1}1), (\bar{1}\bar{1}1)$. Calibration factor C is given by following.

$$C_1 = \frac{\sqrt{3}}{5.43} |\vec{a_1}| = 0.01616 \tag{7}$$

$$C_2 = \frac{\sqrt{3}}{5.43} |\vec{a_2}| = 0.01634 \tag{8}$$

$$C_3 = \frac{\sqrt{4}}{5.43} |\vec{a_1} - \vec{a_2}| = 0.01632 \tag{9}$$

Minhyo obtained $C = 0.01618$ in average.

4 Automatizing

4.1 Determining Coefficients of Linear Combination

There is an infinite number of sets of basis vectors. Hence, it is not easy to determine the only coefficients of linear combination. At here, using unimodular matrix U whose determinant is ± 1 and elements are all integer, we can change the basis vector and the corresponding coefficients freely. Here is a related theorem.

Theorem 1 *Two nonsingular square matrix A and B generate the same lattice if and only if there exist a unimodular matrix U such that $B = AU$.*

Once we determine a set of basis vectors, that is nonsingular square matrix $A = (\vec{a_1}\vec{a_2})$ and center(or basis point), then corresponding coefficients of linear combination are easily determined by integer programming. Using the unimodular matrix, we have

$$B = (\vec{b_1}\vec{b_2}) = AU = (\vec{a_1}\vec{a_2})U \tag{10}$$

$$\begin{pmatrix} x_{m'n'} \\ y_{m'n'} \end{pmatrix} = \begin{pmatrix} x_{mn} \\ y_{mn} \end{pmatrix} = B \begin{pmatrix} m' \\ n' \end{pmatrix} = AU \begin{pmatrix} m' \\ n' \end{pmatrix} = A \begin{pmatrix} m \\ n \end{pmatrix} \tag{11}$$

It is crucial in finding a grid and its intensity profile. Note that the basis vectors determined at first may not be desirable(or conventional). In the case above, the grids along $\vec{a_1}, \vec{a_2}, \vec{a_1} - \vec{a_2}$ are desireable to obtain intensity profile.

Determining a set of basis vectors and a center requires selecting three points. At here, basis vectors should construct a simple unit cell.

I'm working on selecting three points such that two vectors from them construct a simple unit cell.

On the other hand, Lovasz' Basis Reduction Algorithm can be used to find desirable vector basis. Reduced basis are nearly orthogonal.

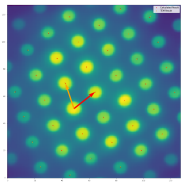


Figure 4: Basis vector, $\vec{a_1}$ is red, $\vec{a_1}$ is orange.

4.2 Miller indices

To calculate the calibration factor, Miller indices are required in some way. I'm working on how to label each point with Miller indices.