The study of effectiveness of a high-performance crystal lattice parametric identification algorithm based on CUDA technology

To cite this article: A S Shirokanev et al 2019 J. Phys.: Conf. Ser. 1368 052040

View the article online for updates and enhancements.
The study of effectiveness of a high-performance crystal lattice parametric identification algorithm based on CUDA technology

A S Shirokanev¹,², D V Kirsh¹,², A V Kupriyanov¹,²

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086
²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

e-mail: alexandrshirokanev@gmail.com, kirshdv@gmail.com, alexkupr@gmail.com

Abstract. The study of substances with a crystal structure is a complex multi-step process. The key step in the crystalline substance analysis is the unit cell parameter estimation. The estimation of the crystal lattice unit cell parameters is a particular problem that involves the search of the crystal lattice model’s parameters according to the information which can be extracted from the substance. In these recent times, the most accurate information about the substance structure can be obtained with the electron microscope whose linear resolution is high enough to observe the atomic structure of a substance. The problem of parameter estimation in this case means the reconstruction of the three-dimensional crystal lattice with 2-dimensional images received by an electron microscope, and the estimation of the crystal lattice unit cell parameters by reconstructed lattice. In the previous papers the crystal lattice parametric identification algorithms based on solving the local optimization problem were presented. However, the analysis of a large crystal lattice database requires a lot of computations. In this paper, a high-performance crystal lattices parametric identification algorithm using the CUDA technology is proposed. The investigation of the algorithm effectiveness is carried out on the GPU GeForce NVidia GTX 1070 Ti. With data dimension more than 32 translations the acceleration is higher than 70. The algorithm runs more efficiently at the use of a large number of CUDA-blocks.

1. Introduction
At present time, most of scientific studies are devoted to the three-dimensional objects reconstruction [1-4]. A particular case is the three-dimensional crystal lattice reconstruction [5-8]. Among the existing methods of three-dimensional crystal lattice parameters estimation the following methods are most common: NIST comparator [5], the identification on the base of the lattice packing density [6], the method of isosurfaces comparison [7]. The methods of the parametric identification are considered to be the most accurate methods of crystal lattice three-dimensional structure investigation [9-16]. The existing algorithms of the three-dimensional crystal lattice parametric identification are presented in scientific studies [9-14].
The crystal lattice can be described with different mathematical models but the most well-known model is Bravais lattice. A Bravais unit cell is described with three translation vectors that form the whole crystal structure [17]. There is an algorithm of the parametric identification based on the Bravais unit cell parameters estimation [12] which is the simplest one comparing with the algorithms based on the other model, e.g. the lattice model described with Wigner-Seitz cell [9-11]. Unfortunately, the corresponding problem of the parametric identification is not correctly set according to Hadamard as there can be more than one solution. This fact is the consequence of Bravais lattice ambiguity [17].

Bravais unit cell is characterized by three vectors \( \vec{r}_1, \vec{r}_2 \) and \( \vec{r}_3 \) called translation vectors. All crystall lattices are divided into 7 lattice systems according to 7 types of Bravais unit cells (Figure 1).

Presented in the paper algorithm [16] allows to improve the parametric identification accuracy comparing with algorithms presented in the scientific studies [9-12]. However the algorithm possesses the high computational complexity which sophisticates its application on large samples. The nanostructure analysis technique will be applied on large crystal structure databases. It is necessary to consider the parameters which the algorithm will possess the highest possible effectiveness with. The crystal structures differ with unit cells parameters, however, the volume of the structure (the number of nodes) can be fixed. In the paper an experimental research resulted in the development of the most appropriate structure volume providing the minimal total period of time for material parameters estimation is conducted.

2. The method of the crystal lattice parametric identification
The algorithm proposed in the paper [16], possesses the best decomposition by data than the algorithm presented in the paper [15], due to breaking the problem into three independent problems. Each of these problems is divided into small subtasks (threads) connected with particular node of the lattice. The idea is based on the characteristic stating that periodically repeated lattice plains intersect all the lattice nodes (Figure 2). These plains can be oriented to three different directions. In this case three independent threads appear.
Figure 2. Periodically repeated plains intersecting lattice nodes.

The algorithm is to perform the following steps [18]:

- Transformation of the translation vectors into vectors of independent basis (not parallelizable part);
- Solving the independent tasks of the local optimization of each vector from the received basis (parallel part on CUDA);
- Transformation of the computed vectors into translation vectors (not parallelizable part).

The problem of Bravais lattice ambiguity leads to appearance of several global minima [19]. Each global minimum is the solution of a parametric identification problem. Lattice structure distortion can lead to appearance of a new local minima, which adds extra difficulties to gradient method application. Found extrema are subjected to additional correction if found vectors are close to linearly-dependent ones. In this case the vector if it is the only dependent on two other vectors is optimized by the other performance function dependent on all the vectors – vectors of translation [15]. If there are two vectors, they are optimized the same performance function where one free vector is fixed.

3. High-performance crystal lattice parametric identification algorithm

The application of the gradient method of steepest descent possesses high computational complexity as the descent ratio $\lambda$ demands significant calculations in each iteration. In the present paper the algorithm based on the gradient method with the constant step is considered. In this case each vector algorithm thread presents the consequence of the following actions:

$$i_i = \arg \min_i \left[ \frac{\mathbf{x}_i - \mathbf{d}}{i \| \mathbf{d} \|} - i \| \mathbf{d} \| \right];$$

$$w_i = \left( \mathbf{d}_i - i \| \mathbf{d} \| \right);$$

$$\mathbf{v}_i = \left[ \mathbf{x}_i - 2i \mathbf{d} \right];$$

$$\mathbf{u}_i = \left( \frac{w_i \mathbf{v}_i}{w_i} \right).$$

The calculated coefficients are used to obtain the value of the objective function and the gradient:

$$\nabla E = 2 \sum_{i=1}^{L} w_i \mathbf{v}_i;$$

$$E = \sum_{i=1}^{L} w_i^2.$$

Each thread uses coordinates of only one lattice node. As a result, the number of threads corresponds to the total amount of three-dimensional lattice’s nodes.

Decomposition by the data leads to the fact that a separately taken thread operates with only one lattice node. Consequently the number of the threads is in proportion with the three dimensional space nodes.
After performing of all the threads, it is necessary to carry out the results reduction, that is to compute the sum of the vectors \( \vec{v}_i \), to form the performance function and gradient results. The reduction on GPU is performed according to the design paradigm “Divide and conquer” with bank conflict resolving [20].

The parallel algorithm uses the global memory for conducting the reduction operation [18]. The stop criterion is checked on CPU. That is the algorithm supposes repeatedly running kernel function on the GPU. The memory for the lattice nodes on the GPU is allocated once. At a particular iteration between random access memory and GPU memory only the optimized vector is transmitted.

GPU implements reduction operation based on the "divide and conquer" principle with the resolution of bank conflicts [18] (Figure 3).

The GPU implemented function computes the local sums for each CUDA-block. After that, CPU sums the outputs and verify the stopping criterion. Thereby CPU calls the GPU function multiple times until the stopping criteria are met.

The described vector algorithm represents an iterative process for solving one optimization problem. To estimate the parameters of a unit cell, it is necessary to solve 3 optimization problems according to the number of translation directions. These problems have no data dependencies among themselves, thus the communication overhead of their parallel implementation can be reduced to minimum.

4. The results of the large crystal lattice database analysis

The developed technique was applied for reconstruction of a big amount of the crystal lattices. Earlier the research was conducted on the GeForce NVidia GTX 840M GPU and Intel Core i7-4710MQ processor [18]. The results demonstrated that the algorithm based on the combined usage of GPU and CPU gives approximately the same result as the algorithm based on the usage of the GPU only (Figure 4).

In the figure 5, the results of the proposed high-performance algorithm acceleration depending on crystal structure dimension are presented. The number of the nodes along three axes is considered to be the crystal lattice dimension. The total number of the threads in this case is calculated as the third lattice dimension degree. The research was conducted on GeForce NVidia GTX 1070 Ti GPU and AMD Ryzen 2600 processor.

![Figure 3. Algorithm implementation scheme using CUDA technology.](image-url)
Figure 4. Acceleration of high-performance algorithms (GPU and CPU-GPU versions).

The study showed that the high-performance algorithm is very effective on higher than 30 structure dimensions: on these dimensions the algorithm acceleration exceeds the value 70. To use the algorithm with dimension lower than 7 is irrational. The work with the memory on the GPU takes quite a lot of time that is why on small dimensions the high-performance algorithm lags behind the sequential algorithm. The achievement of hundred times acceleration on dimension 35 causes a special interest: with this dimension the number of crystal structure nodes is 42875, that, generally speaking, is not an extremely high dimension. It takes the processor 12 minutes to processes this structure, for a large crystal structure database this time becomes critical. As the number of specified threads is 1024 at 1024 nodes intervals a new block on the GPU is allocated. The use of one block is executed at the 1 and 10 dimension interval. The second and third blocks are used up to dimension 16. On the last dimensions 42 blocks are used that could affect on the acceleration negatively. However the acceleration is quite high. Due to the peculiarities of cache memory work on the GPU in some situations even lower time consumption is demonstrated. Such an effect is observed on the dimension 31: the acceleration received a sharp rise. The processor on some dimensions optimized the work with cache memory so the time could be reduced on big dimensions.

Figure 5. Acceleration of the high-performance algorithm on large crystal structures.
Due to the pointed peculiarities of the GPU work the vector algorithm acceleration was analyzed at use of 256 threads. More blocks are allocated for this modification. Such modification allows to achieve a higher acceleration. Consequently for the proposed algorithm it is more effective to use as many as possible memory blocks on specified data dimensions. The optimal dimension is 32 (32768 nodes) as both variants (256 threads and 1024 threads) allow to achieve high acceleration and the restoration or the structure for presented number of nodes does not possess unacceptably high computing complexity.

5. Conclusion
In this paper, a technique for analysis of large-scale crystal nanostructures based on the high-performance crystal lattice parametric identification algorithm is proposed. Such technique will allow to perform an analysis of crystal structures more qualitative than existing crystal lattice parameters estimation methods.

The research showed that the developed technique is the most effective on higher than 30 structure dimensions. On lower than 7 dimensions the use of this technique is irrational. The optimal number of nodes is 32768 as at this structure volume reasonably high acceleration is achieved and the reconstruction will not be time consuming. The research demonstrated that 256 threads assignment appears more effective than 1024 threads one. This fact is connected to the number of used blocks on the GPU. The more the number of blocks the more effective the present algorithm operates.

6. References
[1] Fursov V A and Goshin Ye V 2014 Information technology for digital terrain model reconstruction from stereo images Computer Optics 38(2) 335-342
[2] Kotov A P, Fursov V A and Goshin Ye V 2015 Technology for fast 3d-scene reconstruction from stereo images Computer Optics 39(4) 600-605 DOI: 10.18287/0134-2452-2015-39-4-600-605
[3] Kudinov I A, Pavlov O V and Kholopov I S 2015 Implementation of an algorithm for determining the spatial coordinates and the angular orientation of an object based on reference marks, using information from a single camera Computer Optics 39(3) 413-419 DOI: 10.18287/0134-2452-2015-39-3-413-419
[4] Ilyasova N Yu 2014 Estimating the geometric features of a 3d vascular structure Computer Optics 38(3) 529-538
[5] Kessler E, Henins A, Deslattes R, Nielsen L and Arif M 1994 Precision Comparison of the Lattice Parameters of Silicon Monocrystals Journal of Research of the National Institute of Standards and Technology 99 1-18
[6] Smith W 2004 Foundations of Materials Science and Engineering (McGraw-Hill) 67-107
[7] Patera J and Skala V 2005 Centered cubic lattice method comparison 17th ALGORITMY Conference on Scientific Computing 309-319
[8] Shirokanev A S, Kirsh D V and Kupriyanov A V 2015 Researching methods of reconstruction of three-dimensional crystal lattice from images of projections CEUR Workshop Proceedings 1490 290-297
[9] Kupriyanov A V and Kirsh D V 2015 Estimation of the Crystal Lattice Similarity Measure by Three-Dimensional Coordinates of Lattice Nodes Optical Memory & Neural Networks (Information Optics) 24(2) 145-151
[10] Kupriyanov A V and Kirsh D V 2012 Estimating the similarity measure of crystal lattices by coordinates of their nodes in three-dimensional space Computer Optics 36(4) 590-595
[11] Kirsh D V and Kupriyanov A V 2015 Crystal lattice identification by coordinates of their nodes in three dimensional space Pattern recognition and image analysis 25(3) 456-460
[12] Kirsh D V and Kupriyanov A V 2015 Identification of Three-Dimensional Crystal Lattices by Estimation of Their Unit Cell Parameters CEUR Workshop Proceedings 1452 40-45
[13] Soldatova O P, Lyozin I A, Lyozina I V, Kupriyanov A V and Kirsh D V 2015 Application of fuzzy neural networks for defining crystal lattice types in nanoscale images Computer Optics 39(5) 787-794 DOI: 10.18287/0134-2452-2015-39-5-787-94
[14] Kirsh D V and Kupriyanov A V 2015 Modeling and Identification of Centered Crystal Lattices in Three-Dimensional Space CEUR Workshop Proceedings 1490 162-170

[15] Shirokanev A S, Kirsh D V and Kupriyanov A V 2016 Application of gradient steepest descent method to the problem of crystal lattice parametric identification CEUR Workshop Proceedings 1638 393-400

[16] Shirokanev A S, Kirsh D V and Kupriyanov A V 2016 Development of the crystal lattice parameter identification method based on the gradient steepest descent method Computer Science Research Notes 2603 65-68

[17] Shaskolskaya M P 1984 Crystallography: Manual for Institutes of Higher Education (Higher School) 10-14

[18] Shirokanev A S, Kirsh D V and Kupriyanov A V 2017 Development of a vector algorithm of three-dimensional crystal lattice parametric identification based on estimation of the spacing between adjacent lattice Procedia Engineering 201 690-697

[19] Shirokanev A S, Kirsh D V and Kupriyanov A V 2017 Development of Crystal Lattice Comparison Method Invariant to Bravais Unit Cell Choice IEEE Xplore 125-129

[20] Boreskov A V and Kharlamov A A 2010 Basic concepts of CUDA technology (Moscow, DMK Press) p 232

Acknowledgments
This work was financially supported by the Russian Foundation for Basic Research under grant # 19-29-01135, # 17-01-00972, # 18-37-00418 and by the Ministry of Science and Higher Education within the State assignment to the FSRC “Crystallography and Photonics” RAS.