An Analytical Method of Automatic Alignment for Electron Tomography

Shuang Wen(✉) and Guojie Luo(✉)

Center for Energy-efficient Computing and Applications,
Peking University, Beijing 100871, China
{wenshuang,gluo}@pku.edu.cn
http://ceca.pku.edu.cn

Abstract. In the imaging process for nanometer-scale electron tomography, misalignment between the actual projection parameters and the theoretical ones is inevitable due to mechanical precision of the instrument. Effective alignment remains a challenge. Currently, marker-based alignment approaches complicate the sample preparation process and worsen the sample shrinking issue. Marker-free approaches suffer from either low accuracy or long computation time.

In this paper, we formulate an analytical problem for marker-free alignment by minimizing the reprojection error. The reprojection error involves the projection operator, which is a complicated functional with the projection parameters as the variables. To solve this optimization problem, we derive a gradient-based approach by decomposing the original problem with auxiliary parameters and by linearizing a subproblem with Taylor expansion. The approach is computational friendly, especially when comparing to an exhaustively parameter tuning approach in previous practice. The results show that our method is capable of accurate alignment without fiducial markers and obtains a 16.7× speedup over the existing exhaustive approach, which makes fine reconstruction of ROI almost instantly ready after data collection. A preliminary FPGA design for the method's bottleneck process shows 6.6× speed-up over well-optimized GPU program.

Keywords: Electron tomography · Automatic alignment · Functional optimization

1 Introduction

Electron tomography (ET), a technique combining transmission electron microscopy (TEM) and computed tomography, is now widely used for acquiring high-resolution 3D structures of biological samples. To obtain the 3D structures, it is critical to reconstruct the region of interest (ROI) from projection images provided by the TEM microscope. These projection images are usually collected

L. Zhou et al. (Eds.): LABELS 2019/HAL-MICCAI 2019/CuRIOUS 2019, LNCS 11851, pp. 106–114, 2019.
https://doi.org/10.1007/978-3-030-33642-4_12

according to specific regulations called tilt geometries. A tilt geometry decides the position and attitude of a series of projection images (tilt series) for ET reconstruction.

We perform ET reconstruction given the theoretical tilt geometry and the tilt series. However, due to high magnification and low mechanical accuracy, unexpected drift and rotation of ROI happen during the image collection process. Therefore, alignment is needed for high-quality reconstruction results.

There are mainly two types of alignment methods, marker-based alignment and marker-free alignment. Fiducial marker-based alignment approaches [8,10] use high-contrast markers, such as gold beads, embedded in a sample to determine the position and attitude of the tilt series. However, this type of method is not always available, since it is difficult or impossible to embed enough fiducial markers in the ROI sometimes. Besides, the selection of marker detection algorithms is data-dependent [11], which also limits the usability of the marker-based approaches.

Marker-free alignment approaches require no embedded fiducial markers. And these methods can be further categorized into cross-correlation and feature-based methods. Guckenberger [2] brought up cross-correlation alignment method which determines the common origin of tilt series by comparing the cross-correlation coefficient. But this method is bothered by errors accumulating along with alignment going on. For this problem, Winkler and Taylor [13] proposed a method combining cross-correlation and reconstruction-reprojection to compensate accumulated errors, which is still widely used. This method will be mention below as naive exhaustive search (NES) method On the other hand, feature-based methods make use of image features as markers to do the alignment. Feature-based methods is often less time-consuming [3] but need specific detectors for different kind of datasets [11], which damages its universality.

To overcome these problems, the method presented in this paper is developed for reconstructing specimens without fiducial markers and apparent local features. Inspired by Houben and Sadan [5], method is mainly composed of a coarse alignment process by cross-correlation and a refinement process based on minimization of reprojection error. The former coarse alignment is used to provide the latter process with an initial value. The following refinement process further improves the alignment accuracy. With both procedures, we grantee the algorithm with both efficiency and accuracy. Compared with other methods, our method does not depend on fiducial markers or image features. Through reconstruction and reprojection process, both projection space and real space information are made full use of, which makes our method free from accumulated correlation errors that happen in the cross-correlation method. Compared with the iterative reconstruction-reprojection method brought up by Winkler and Taylor [13], as experimental results show, less iteration number needed and fewer operations within one iteration make our method much efficient. The experiment on a conical-tilt dataset shows our method's comparable accuracy and 16.7× efficiency compared with similar marker-free method. For the method's bottleneck process, a preliminary FPGA design shows 6.6× speed-up over a well-optimized GPU program.

2 Automatic Alignment Problem

We reconstruct the 3-D image of a specimen from a collection of 2-D TEM images (projections) using electron tomography. The process is determined by a set of parameters in θ. For a fixed tilt angle, we collect N projections of the specimen at configuration $\theta = (\theta_1, \theta_2, \cdots, \theta_N)$. The i-th projection is determined by a 5-tuple $\theta_i = (\alpha_i, \beta_i, \gamma_i, x_i, y_i)$, where the ROI center of the specimen is projected at coordinate (x_i, y_i), and α_i, β_i and γ_i are the yaw, pitch, and roll angles of the specimen, respectively.

Ideally, the 5-tuple of each projection is known at prior. Using an automatic, a semi-automatic, or a manual controller, we take TEM images at certain angles by rotating the specimen along its normal with a fixed tilt. The specimen's ideal posture at certain configuration is as shown in Fig. 1.

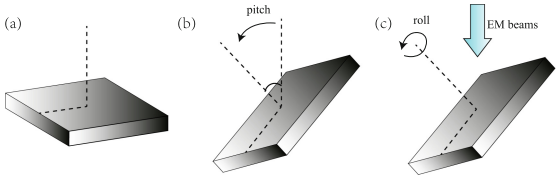


Fig. 1. (a) Original specimen. (b) Tilt specimen with a fixed pitch. (c) A series of tilt specimen with various roll angles.

At each angle, we obtain the corresponding yaw, pitch, and roll angles, and we shift, align, and refocus the ROI center before taking a TEM image.

However, due to the intrinsic random and system errors of the instrument (e.g., the controller, the motor, and the tray), the actual configuration $\theta_i^* = (\alpha_i^* + \epsilon_{\alpha_i}, \beta_i^* + \epsilon_{\beta_i}, \gamma_i^* + \epsilon_{\gamma_i}, x_i^* + \epsilon_{x_i}, y_i^* + \epsilon_{y_i})$ is different from the ideal configuration $\tilde{\theta}_i = (\alpha_i^*, \beta_i^*, \gamma_i^*, x_i^*, y_i^*)$.

Assuming f^* is the unobservable 3-D image, the TEM imaging process can be described using the projection operator $\mathcal{R}(\theta^*) = (\mathcal{R}(\theta_1^*), \mathcal{R}(\theta_2^*), \cdots, \mathcal{R}(\theta_N^*))$. And the projection data $g = \mathcal{R}(\theta^*)f^* = (g_1, g_2, \cdots, g_N)$ consists of a set of 2-D TEM images, the i-th of which is $g_i = \mathcal{R}(\theta_i^*)$.

If we use the mistakenly-believed ideal configuration $\tilde{\theta}$ to reconstruct the 3-D image by solving for $\operatorname{argmin}_f ||\mathcal{R}(\tilde{\theta})f - g||^2$, it will always generate inaccurate results, since $\mathcal{R}(\tilde{\theta})f^* \neq g = \mathcal{R}(\theta^*)f^*$. Therefore, we propose the automatic correction problem to recover the actual configuration θ^* for a high-quality electron tomography, so that we can avoid the quality degradation due to using the mistakenly-believed ideal configuration $\tilde{\theta}$ during image reconstruction. For convenience consideration, we mark $\operatorname{argmin}_f ||\mathcal{R}(\tilde{\theta})f - g||^2$ with $\mathcal{S}(\theta, g)$.

We formulate automatic correction as a functional optimization problem,

$$\begin{aligned} \min_{\theta} \quad & \|\mathcal{R}(\theta)f(\theta) - g\|^2 \\ \text{where} \quad & f(\theta) = \operatorname*{argmin}_f \|\mathcal{R}(\theta)f - g\|^2 . \end{aligned} \tag{1}$$

The parameterized operator (functional) $\mathcal{R}(\theta)$ models how the projection data g is acquired, and the image $f(\theta)$ is reconstructed at the guess of projection configuration θ.

Apparently, the actual configuration θ^* is an exact solution to this problem, such that $||\mathcal{R}(\theta^*)f(\theta^*) - g|| = ||\mathcal{R}(\theta^*)f^* - g|| = 0$. By solving the automatic correction problem to recover an estimate of the actual configuration $\hat{\theta}$, we expect to "correct" the mistakenly-believed ideal $\tilde{\theta}$, so that $||\mathcal{R}(\hat{\theta})f(\hat{\theta}) - g|| < ||\mathcal{R}(\tilde{\theta})f(\tilde{\theta}) - g||$, if not $||\mathcal{R}(\hat{\theta})f(\hat{\theta}) - g|| = 0$.

3 Automatic Alignment Methods

3.1 Naive Exhaustive Search

The basic idea of naive exhaustive search is to examine the neighborhood of known parameters in the solution space and find the best solution in this neighborhood by comparing the values of the objective function. Taking a projection series with 72 projections and each projection with 5 configuration parameters for example, the solution space dimension is $72 \times 5 = 360$. A search in 360-dimension space requires objective function $\mathcal{S}$ and $\mathcal{R}$ calculated large amount of times.

3.2 Analytical Optimization

The key of our method is the optimize of Eq. (1) using gradient descent method. This process involves computing the gradients $\nabla_\theta \mathcal{R}(\theta)$ and $\nabla_\theta f(\theta)$. The latter one is relatively difficult to write down the analytical form. To use the gradient descent method, we reformulate Eq. (1) into the following problem by introducing the auxilary varaible $\bar{\theta}$,

$$\begin{aligned} &\min_{\theta,\bar{\theta}} \quad \left\|\mathcal{R}(\theta)f(\bar{\theta}) - g\right\|^2 \quad \text{s.t.}\bar{\theta} = \theta \\ &\text{where } f(\bar{\theta}) = \operatorname{argmin}_f \left\|\mathcal{R}(\bar{\theta})f - g\right\|^2. \end{aligned} \tag{2}$$

And we apply a hybrid approach of block descent and gradient projection to solve this reformulated problem. And our algorithm is outlined as below,

Step 0. Start from $k = 0$ with initial guess $\theta^{(0)} = \bar{\theta}^{(0)} = \tilde{\theta}$;
Step 1. Solve $f^{(k)} = \operatorname{argmin}_f ||\mathcal{R}(\bar{\theta}^{(k)})f - g||^2$;
Step 2. Solve $\theta^{(k+1)} = \operatorname{argmin}_\theta ||\mathcal{R}(\theta)f^{(k)} - g||^2$;
Step 3. Update $\bar{\theta}^{(k+1)} = \theta^{(k+1)}$;
Step 4. If not converged, Set $k = k + 1$ and **Goto** Step 1.

The subproblem in Step 1 is the conventional image reconstruction problem for electron tomography. Weighted back-projection is one of the feasible methods.

We solve the subproblem in Step 2 using linearization. It is natural to start with the initial solution $\theta^{(k)}$. We then perform a Taylor expansion at $\theta^{(k)}$ and derive $\mathcal{R}(\theta^{(k)} + \Delta\theta)f^{(k)} = (\mathcal{R}f^{(k)})(\theta^{(k)} + \Delta\theta) \approx (\mathcal{R}f^{(k)})$

$(\theta^{(k)}) + \nabla_\theta(\mathcal{R}f^{(k)})(\theta^{(k)})\Delta\theta$. After solving $\Delta\theta^{(k)} = \text{argmin}_{\Delta\theta} ||(\mathcal{R}f^{(k)})(\theta^{(k)}) + \nabla_\theta(\mathcal{R}f^{(k)})(\theta^{(k)})\Delta\theta - g||^2$ by least squares, we update $\theta^{(k+1)} = \theta^{(k)} + \Delta\theta^{(k)}$.

According to the definition of operator $\mathcal{R}$, integration form of $\nabla_\theta(\mathcal{R}f^{(k)})(\theta^{(k)})$ can be expressed by

$$\nabla_\theta(\mathcal{R}f^{(k)})(\theta^{(k)}) = \nabla_\theta(\int_{L_{\theta^{(k)}}} f(\boldsymbol{r})\mathrm{d}|\boldsymbol{r}|), \tag{3}$$

where $L_{\theta(k)}$ is the integration path determined by projection configuration $\theta(k)$. We call each path of integration a ray.

With all $\nabla_\theta(\mathcal{R}f^{(k)})(\theta^{(k)})$ calculated and the precondition that the minimum of Eq. (2) is 0, the minimization turns out solving $\Delta\theta$ in a linear equation set $||(\nabla_\theta(\mathcal{R}f^{(k)})(\theta^{(k)})f) \cdot \Delta\theta|| = 0$. Each equation in this set corresponds to a different θ. Noticing the fact that the electron microscope collects all g data with same angular parameters simultaneously, data points with same $(\theta_\alpha, \theta_\beta, \theta_\gamma)$ share the same $\Delta\theta$. With the reduction of variables, the equation set is now overdetermined and its least-squares solution is what we are looking for. For every iteration, the reconstruction and corresponding Taylor expansions are recalculated to make sure the error caused by $\Delta\theta$ is always much more significant than error from minimization.

3.3 Time Complexity and Hardware-Based Improvements

According to the algorithm outline, the basic calculation unit of optimization is the solving of operator $\mathcal{R}$, $\nabla_\theta\mathcal{R}$, and $\mathcal{S}$. Compared with naive exhaustive search (NES) method, our method significantly decreases the amount of calculation. The NES method goes through the solution space to find the best configuration match. To eliminate influence by other configuration parameters, each parameter must be searched separately and one $\mathcal{R}$ and one $\mathcal{S}$ operator is carried out during every single trial. Instead, for the sake of gradient descent, our method accomplishes correcting for all configuration parameters in one descent made up with one $\mathcal{R}$ calculation, one $\nabla_\theta(\mathcal{R})$ calculation, and one $\mathcal{S}$. In addition, in order to reduce further, we find that during one descent process, the data usage and calculation structure in calculations of $\mathcal{R}$ and $\nabla_\theta(\mathcal{R})$ have much in common. We also find that during one calculation, one data point in f is used only 1–5 times for multiplication before it's discarded. All these features inspire us to an idea of heterogeneous computing and data reuse. So we firstly input data f and current configuration into an OpenCL kernel. $\mathcal{R}f$ and $\nabla_\theta(\mathcal{R})f$ are then calculated simultaneously to make full use of GPU bandwidth. The rough estimation of theoretical calculation amount is shown in Table 1. In this table, N_p represents the value number of trial for one configuration parameter. Namely, when 3 values in both sides of a given value with fixed interval are tried, N_p is 7. The 4th column in Table 1 shows a specific comparison among methods when $N_p = 7$ and the cost of $\mathcal{R}$ and $\mathcal{S}$ considered to be comparable.

For the reprojection $\mathcal{R}f$ process, which is the most expensive, we make use of Vivado High Level Synthesis (HLS) for acceleration. This tool enables accelerating Clang-base design and exporting RTL as a Vivado's IP core. To adapt for

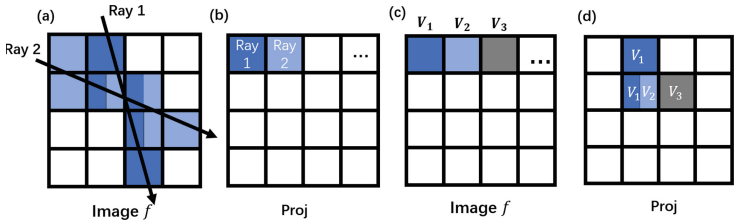


Fig. 2. (a) Ray-based: ray track in data f. (b) Ray-based: data output in data $\mathcal{R}f$. (c) Voxel-based: sequential data read from f. (d) Voxel-based: data output in data $\mathcal{R}f$.

hardware features of FPGA, we rearrange data flow of reprojection process from ray-based to voxel-based to ensure data independency of inner loop. As shown in Fig. 2(a) and (b), ray-based orienting calculate the path of rays in f and load corresponding data. The loaded data is process and then output in $\mathcal{R}f$ accordingly. However The load sections of different rays overlap a lot, which causes huge load conflicts. Those conflicts are unpredictable and make initial interval (II) unbearable. To solve this problem, we read voxel data from f sequentially and find out the rays that contain voxel data as Fig. 2(c) and (d) shows. We find that adjacent voxels always belong to adjacent rays. So we cache the output data instead of saving it until the current voxel is irrelevant to it. In this way pipeline among voxels and be carried out and the II could be reduced.

Table 1. Time complexity comparison. The value 1.4 at row.4 & col.2 is an experimental result of data reuse. The value 5 at row.2 & col.3 is the size of tuple that describe a projection.

Method	Trial cost	Trial per iter.	Specific case
NES	$1 \times \mathcal{R} + 1 \times \mathcal{S}$	$N_p \times 5$	$70 \times \mathcal{R}$
Proposed (w/o data reuse)	$6 \times \mathcal{R} + 1 \times \mathcal{S}$	1	$7 \times \mathcal{R}$
Proposed (w/ data reuse)	$1.4 \times \mathcal{R} + 1 \times \mathcal{S}$	1	$2.4 \times \mathcal{R}$

4 Experiments and Result

The dataset we use for experiment is a conical-tilt projection series collected by FEI Tecnai 12 and 2048×2048 CCD Gatan camera. The sample were tilt to 55° and then rotated by 5° interval (72 projections in total). One of its projection is displayed in Fig. 3(a). And the test is carried out on a linux platform with 2×E5-2650 v3, 64 GB Memory and Tesla K80. The synthesis tool is Vivado 2019.1 and synthesis configuration is based on Xilinx ZCU102.

Using this dataset as input, we run both proposed method and NES method for alignment. Popular feature-based method [3] does not provide a solution for conical tilt datasets, so it is temporarily excluded. For NES method, we collect the total time cost and the final alignment result. For proposed method,

Table 2. Similarity indicators comparison of different alignment methods.

Method	Avg MSE	Avg RAE	Avg NCC
Raw	92.8	0.991	0.449
NES	18.3	0.188	0.873
Proposed (one iter)	18.3	0.191	0.871
Proposed (two iters)	18.2	0.189	0.870

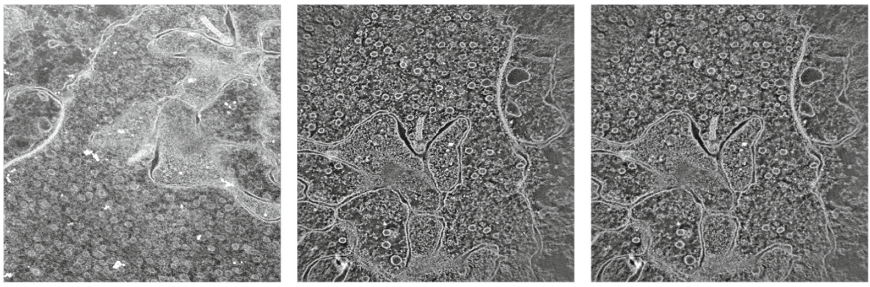
alignment results and detailed time costs after every iteration are recorded. For every alignment result, we reproject corresponding reconstruction image using configuration determined by the result to obtain reprojection images. Then the similarity between reprojection images and input conical-tilt dataset is evaluated using indicators like mean squared error (MSE), relative average error (RAE), and normalized cross correlation (NCC). The similarity indicators reveal the accuracy of the alignment. Specially, When accuracy is higher, the MSE value is lower, RAE value is lower and NCC value is higher. Table 2 shows the average indicators value of different projections. The projection process, which is the bottleneck, is C-synthesised on Vivado HLS with target ap_clk = 10ns and resource limited by Xilinx ZCU102.

Table 3. Time cost details of methods.

Method	OpenCL init (s)	Reproj (s)	Recon (s)	IO (s)	Total (s)
NES	–	–	–	–	5520
Prop (one iter)	4.5	56.2	50.6	67.0	178.3
Prop (two iters)	9.0	112.4	75.9	134.0	331.3
HLS (reproj only)	II = 1	Latency = 10.6E8			∼9.2

Except for quantitative evaluation, for visual observation, cross section of reconstruction results by proposed method and NES method is shown in Fig. 3(b) and (c). The sharpness and clarity of images by both methods is comparable.

The time cost information of methods is listed in Table 3. The result shows a significant speed-up of 16.7× between NES method and proposed method. The HLS simulation for reprojection process also has a 6.6× speed-up over our well-optimized OpenCL program on Tesla K80.



(a) One of the projections from experiment dataset

(b) Cross section of the recon by prop method after 2 iters

(c) Cross section of the recon by NES method

Fig. 3. Experimental data and reconstruction results comparison of different alignment methods.

5 Conclusion

Effective alignment for nanometer-scale electron is currently a challenge. Using the gradient-based approach, we have derived a descent method which decomposes the problem into a computational friendly optimization problem. This method is capable of accurate alignment for datasets with no fiducial markers. The experiment results show the reliability and efficiency. Compared with the NES method, our method manage to achieve comparable accuracy with 16.7× efficiency, which enables operators or researchers to get fine reconstruction of ROI almost instantly after data collection. For reprojection related process in our method, a preliminary design based on Xilinx ZCU102 shows a 6.6× acceleration compared with a well-optimized OpenCL program on GPU.

References

1. Fernandez, J.J.: Computational methods for electron tomography. Micron **43**(10), 1010–1030 (2012)
2. Guckenberger, R.: Determination of a common origin in the micrographs of tilt series in three-dimensional electron microscopy. Ultramicroscopy **9**(1–2), 167–173 (1982)
3. Han, R., Bao, Z., Zeng, X., et al.: A joint method for marker-free alignment of tilt series in electron tomography. Bioinformatics **35**(14), i249–i259 (2019)
4. Herman, G.T.: Fundamentals of Computerized Tomography: Image Reconstruction from Projections, pp. 64–68. Springer, London (2009). https://doi.org/10.1007/978-1-84628-723-7
5. Houben, L., Sadan, M.B.: Refinement procedure for the image alignment in high-resolution electron tomography. Ultramicroscopy **111**(9–10), 1512–1520 (2011)
6. Leis, A., Rockel, B., Andrees, L., et al.: Visualizing cells at the nanoscale. Trends Biochem. Sci. **34**(2), 60–70 (2009)
7. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision (1981)

8. Mastronarde, D.N., Held, S.R.: Automated tilt series alignment and tomographic reconstruction in IMOD. J. Struct. Biol. **197**(2), 102–113 (2017)
9. Natterer, F.: The Mathematics of Computerized Tomography. SIAM, Philadelphia (1986)
10. Tian, Q., Öfverstedt, L.G.,: Unit USSCB. Semi-automatically aligned tilt images in electron tomography. In: 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), pp. 71–75. IEEE (2017)
11. Trampert, P., Bogachev, S., Marniok, N., et al.: Marker detection in electron tomography: a comparative study. Microsc. Microanal. **21**(6), 1591–1601 (2015)
12. Sorzano, C.O.S., Messaoudi, C., Eibauer, M., et al.: Marker-free image registration of electron tomography tilt-series. BMC Bioinform. **10**(1), 124 (2009)
13. Winkler, H., Taylor, K.A.: Accurate marker-free alignment with simultaneous geometry determination and reconstruction of tilt series in electron tomography. Ultramicroscopy **106**(3), 240–254 (2006)