Comparative Study of Projection/Back-projection Schemes in Cryo-EM Tomography

Yu Liu and Jong Chul Ye
Department of BioSystems
Korea Advanced Institute of Science and Technology, Daejeon, Korea

ABSTRACT

In the cryo-EM tomography, the projection and back-projection are essential steps in reconstruction the 3D structure of the virus and macromolecules. Distance driven method (DD) is the latest projection /backprojection algorithm originally employed for x-ray computed tomography. This paper is mainly concerned about employing this algorithm to the cryo-EM tomography for reconstruction performance improvement. Existing algorithms used in cryo-EM are pixel-driven and ray driven projection/backprojection, etc. These methods are generally quite time consuming because of their high computational complexity. Furthermore, interpolation artifacts are usually noticeable when the sufficient view and detector samples are not available. The DD is originally proposed to overcome these drawbacks. The interpolation process in DD is done by calculating the overlap area between the detector and pixel boundaries. This procedure largely removes the interpolation artifacts, and reduces the computational complexity significantly. Furthermore, it guarantees that the projection and backprojection are adjoint to each other – a desired property to guarantee the convergence of the iterative reconstruction algorithm. However, unlike the x-ray computed tomography, the cryo-EM tomography problem generally has limited number of the projections, and projection angles are randomly distributed over 4pi steradian. Therefore, the conventional DD should be modified. Rather than computing the boundary overlap in the previous 3-D DD method, we propose a novel DD algorithm based on volume overlap. CCMV virus model is used as testing example. Results are visualized using AMIRA software. Analysis is made upon the advantages and drawbacks of both the existing approaches and distance driven method.

Keywords: Cryo-EM, Projection, Back-projection, Pixel-driven interpolation, Distance-driven interpolation

1. INTRODUCTION

Single particle reconstruction using cryo-electron microscopy (cryo-EM) is a very active area of research in structural biology due to its advantages over the x-ray crystallography [1]. In the Cryo-EM tomographic reconstruction algorithms such as weighted backprojection (WBP)[7], or algebraic reconstruction technique (ART) [5], the projection and backprojection step are essential components. The process which transforms the 3D image to 2D sinograms is called projection. Physically, this process is the line integration along the electron beams direction. And the backprojection transfers multiple views of 2D projection onto a 3D volume back through the view direction. In discrete implementation, the continuous line integral should be approximated on a discrete lattice. Depending on the approximation, different type of artifact could be expected.

Conventionally, the most popular algorithms in this field are pixel driven (PD) and ray driven (RD) methods [2]. They are widely used in the computed tomography area during the past 20 years. Pixel driven approach traces the line through the centers of voxels, and the ray driven method traces the line through the detector centers. The distance driven (DD) approach by De Man and Basu [3] is different from the PD and RD in that it uses so-called kernel operation, and a common axis or the common plane for 3D case parallel to the pixel cells. The interpolation process in DD is done by calculating the overlap area between the detector and pixel boundaries. Due to the low mathematical complexity and high sequential memory access pattern, distance driven approach achieves faster computational speed over the conventional methods. Furthermore, the distance driven method result in better performance in eliminating artifacts of reconstructed 3D image.

Yu Liu: E-mail: liuyu0817@kaist.ac.kr, Telephone: +82 (0)42 869 4360
Jong Chul Ye: corresponding autor, E-mail: jong.ye@kaist.ac.kr Telephone: +82 (0)42 869 4320

Image Reconstruction from Incomplete Data IV, edited by Philip J. Bones, Michael A. Fiddy, Rick P. Millane,
Proc. of SPIE Vol. 6316, 631606, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.682287

Distance driven approach suits well for the medical computed tomography implementation. Medical computed tomography is generally a fan beam in x-y direction with limited beamwidth along z-direction. This results in an accurate calculation of the boundary overlap.

However, unlike the x-ray computed tomography, the cryo-EM tomography problem generally have limited number of the projections randomly distributed over 4pi steradian This intrinsic difference makes it quite complicated to get a sufficiently accurate approximation of 2D overlap using the conventional distance driven method. Therefore, the conventional DD should be modified. For the Cryo-EM tomography, we propose to use the common plane parallel to the detector plane. Furthermore, each voxel is approximated as a sphere. Then, the projection of the voxel corresponds to the circle in the common plane, and the circle to circle intersection area is calculated. This proposed scheme offer better performance in the reconstruction quality over conventional DD approaches using a few set of fixed common planes. In order to speed up the circle to circle overlap area calculation, a fast approximation method is proposed using a bounding rectangle. Significant speed up of the interpolation was obtained with negligible performance degradation.

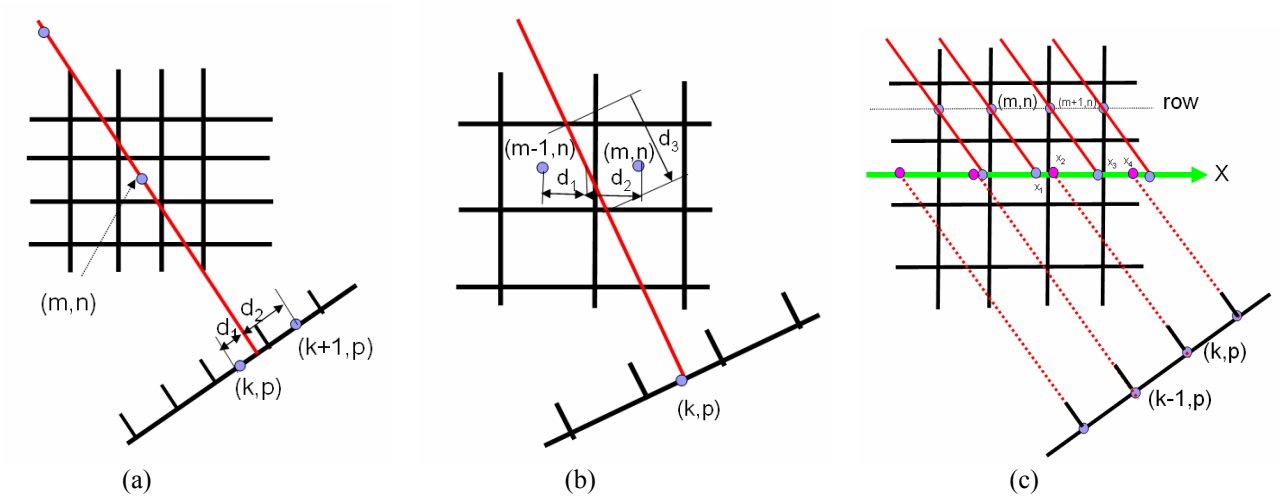


Fig 1. (a) 2D demonstration of pixel driven projection and backprojection approach. It traces the electron beam from the focal spot through the center of the pixel of interest to the detector. (b) 2D ray driven projection and backprojection. it traces the electron beam from the focal spot through the image to the center of the detector cell of interest. (c) 2D distance driven projection and back projection, the boundaries of pixel cells and detector cells are mapped to common axis.

2. REVIEW OF CONVENTIONAL METHODS

2.1 Pixel Driven and Ray Driven Approach in 2-D

As shown in Figure 1, the pixel driven projection and backprojection traces the electron beam from the focal spot through the center of the pixel of interest to the detector. The center of the pixel is mapped to the detector row. In the projection process, the detector value is calculated based on the interpolation of the pixel value using Eq. (1):

$$r(k,p) \leftarrow r(k,p) + f(m,n)\frac{d_2}{d_1 + d_2} \tag{1}$$

where f(m,n) denotes the pixel value at (m,n) of the image space, and r(k,p) is the updated sinogram values at (k,p) in the radon space, respectively. Here, d1 and d2 value denotes the distance from adjacent detector cell centers. In the backprojection, the pixel value is updated and accumulated by adding interpolated values of the detecter cells as shown in Eq.(2):

$$f(m,n) \leftarrow f(m,n) + r(k,p)\frac{d_2}{d_1 + d_2} + r(k+1,p)\frac{d_1}{d_1 + d_2} \tag{2}$$

Ray driven projection and backprojection trace the electron beam from the focal spot through the image to the center of the detector cell of interest. For each image row, the intersection is calculated. In the projection process, the linear interpolated value of the adjacent pixel values is calculated, which is then accumulated at the detector cell as shown in Eq. (3).

$$r(k,p) \leftarrow r(k,p) + f(m,n)\frac{d_1 d_3}{d_1 + d_2} + f(m-1,n)\frac{d_2 d_3}{d_1 + d_2} \tag{3}$$

In the backprojection step, relative distances from the center of the adjacent pixels are calculated and the detector values are distributed according to the weighting as shown in Eq. (5) and Eq. (6).

$$f(m,n) \leftarrow f(m,n) + r(k,p)\frac{d_1 d_3}{d_1 + d_2} \tag{5}$$

$$f(m-1,n) \leftarrow f(m-1,n) + r(k,p)\frac{d_2 d_3}{d_1 + d_2} \tag{6}$$

The intrinsic differences between the pixel driven and ray driven approaches are from the fact whether the line is connected through the center of pixels or the center of the detector cells. Due to these differences, distinct types of artifact are observed.

2.2 Distance Driven Approach in 2-D

The distance driven approach proposed by De Man and Basu [2] uses a common axis for the 2D case during interpolation step. Fig. 1 (c) shows the 2-D distance driven projection/backprojection steps. Here, x axis is the common axis, to which the boundaries of pixel cells as well as the detector boundaries are uniquely mapped. Then, the interpolation process is done based on so-called kernel operation; i.e. the intersection lengths are calculated on this axis, and the destination value is computed based on the normalized overlap length. More specifically, the projection step is given by Eq. (7) whereas the back-projection part is given by Eq. (8).

$$r(k,p) \leftarrow r(k,p) + f(m+1,n)\frac{x_4 - x_3}{x_4 - x_2} + f(m,n)\frac{x_3 - x_2}{x_4 - x_2} \tag{7}$$

$$f(m,n) \leftarrow f(m,n) + p(k-1,p)\frac{x_2 - x_1}{x_3 - x_1} + p(k,p)\frac{x_3 - x_2}{x_3 - x_1} \tag{8}$$

2.3 3D Distance Driven Approach for x-ray Computed Tomography

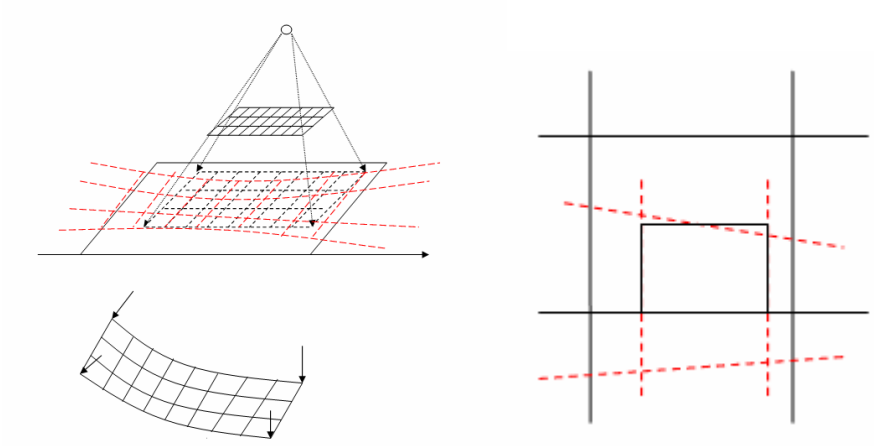
The basic idea of 3D distance driven approach has no significant differences from the 2D application. Instead of using common axis, common plane is used for 3D application. As shown in Fig. 2, the common plane is parallel to the XZ plane, the voxel boundary and the detector cell boundaries are mapped to the common plane. In practice, they only map the vertical and horizontal boundaries of the voxels and detector cells to the common plane and approximate them as rectangles. Then, the destination values(detector value for the projection process and pixel value for the backprojection process) are computed based on the normalized overlap area. This method suits well for the medical computed tomography application, where the common plane is parallel to the XZ plane. More specifically, when mapping the vertical and horizontal boundaries of the voxels to the common plane, the resultant shape on the common plane is a quite accurate approximation of the 3D voxel boundary mapping to common plane. Besides, the shape of the detector cell boundary on the common plane is close to a rectangle, so we can approximate it as a rectangle without sacrificing the accuracy of the interpolation.

3. MODIFIED DISTANCE DRIVEN ALGORITHM FOR 3D CRYO-EM TOMOGRAPHY

3.1 Circle based method

However, the intrinsic differences of Cryo-EM tomography from x-ray computed tomography make it impractical to use the conventional distance driven method. Unlike the x-ray computed tomography, in which the projecting angles are within small range, parallel beam projections used in cryo-EM tomography are randomly distributed over 4pi steradian. Suppose we assign the XY plane as the common plane similar to the conventional distance driven approach. As shown

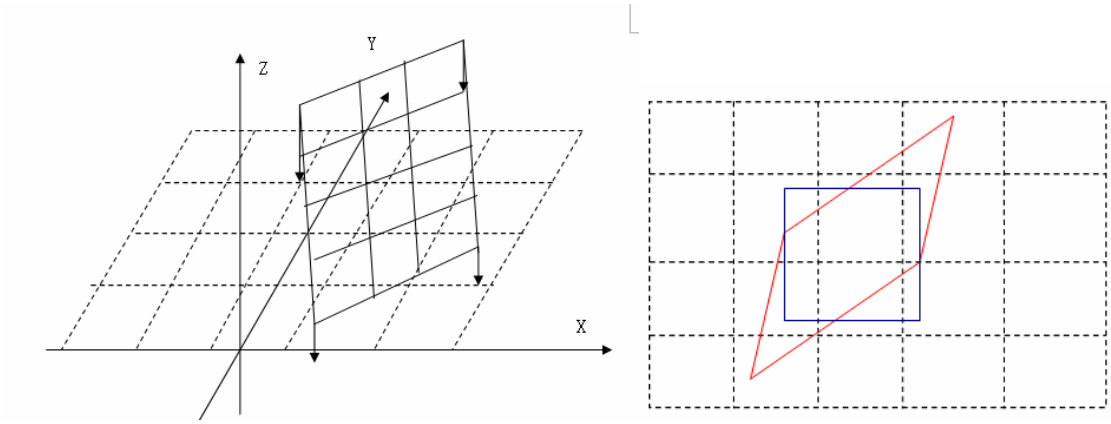
in Fig.3, when we map the boundaries of detector cells to the common plane, the resulting shape will be parallelogram; hence, it is very inaccurate if we approximate this parallelogram as a rectangle suggested by Deman and Basu [2].



(a) (b)

Fig 2. 3D implementation of distance driven approach [3]: (a) vertical and horizontal boundaries of voxels and detector cells are mapped to the common plane. The overlap lengths along x axis and z axis are calculated and then multiplied together to get the overlap area. (b) approximate the overlap area as rectangle.

In order to overcome the drawbacks, our new algorithm uses the common plane parallel to the detector plane or, equivalently, we can consider the detector plane is just the common plane. This approach appears somewhat similar to the pixel-driven approach in some sense; however, it is fundamental different from the pixel-driven approach in that the modified DD uses the area overlap as a weighting factor during the interpolation rather than the distance from the center in the PD. Then, the technical issue in this case is how to get the shape of the voxels mapped to the detector plane and how to calculate the overlap between the pixel boundaries(from voxels) and detector boundaries in a computationally efficient way.



(a) (b)

Fig 3. For the case of the common plane parallel to the XY plane (basically that means parallel to the slices of the 3D image to be projected). (a) the detector plane(real line)and common plane in 3D. The dashed grid is the pixel boundaries mapped to the common plane. (b) The detector cells are mapped to the common plane, the shape is parallelogram, this will make it difficult to calculate the overlap area with pixel boundaries(dashed line), and it will be very inaccurate if we approximate this parallelogram as a rectangle(real line)with sides parallel to the x and y coordinates.

Instead of considering each voxel as a cube voxel, we propose a new method to approximate it as a sphere. Then, for each detector cell, it is considered as a circle regardless of the view angle.

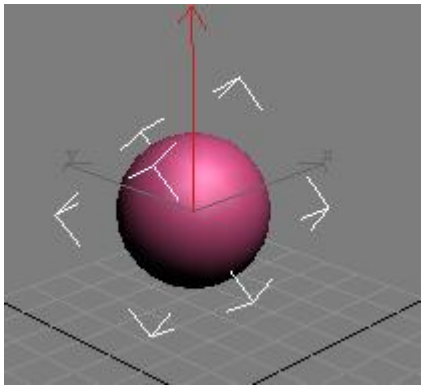


Fig 4. We define the common plane parallel to the detector plane, that means the projecting direction is always perpendicular to the common plane. Here, the voxel is approximated as a sphere. When the sphere is mapped to the common plane, the shape is always circle regardless of the rotation angle.

We also approximate that the detector elements as a circle. Then, the overlap computation accuracy benefits greatly from this approach, because no matter what view angle is applied, when voxel is mapped to the detector plane, the shape is circle. So the problem turns to be the circle and circle overlap area calculation.

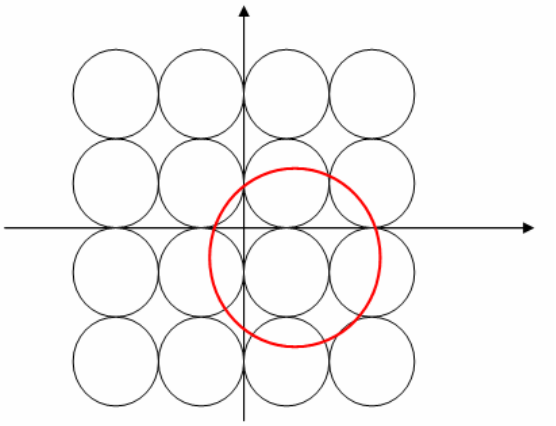


Fig 5. Overlap situation between the one pixel cell (big circle) and the detector cells (small circles) on the common plane.

More specifically, the projection operation can be described by

$$r(i,j,\theta) \leftarrow r(i,j,\theta) + f(l,m,n)\frac{O}{A} \tag{9}$$

where $r(i,j,\theta)$ denotes the (i,j)-th detector value at the view angle θ, f(l,m,n) is the voxel value at the (l,m,n)-th voxel, O denotes the overlap area between the circles, and A is the detector cell area. Eq. (9) should be computed for all voxel f(l,m,n) whose overlap area O is not zero. Similarly, the back-projection operation can be described by

$$f(l,m,n) \leftarrow f(l,m,n) + r(i,j,\theta)\frac{O}{A} \tag{10}$$

where O is the overlap area between two circles, and A is the projected voxel area on the detector plane, Note that Eq. (9) and (10) are the kernel equation for our approach, and they are adjoint operation to each other.

In the modified distance driven approach, the overlap area between the detector circles and pixel circles can be calculated according to the formula Eq. (11).

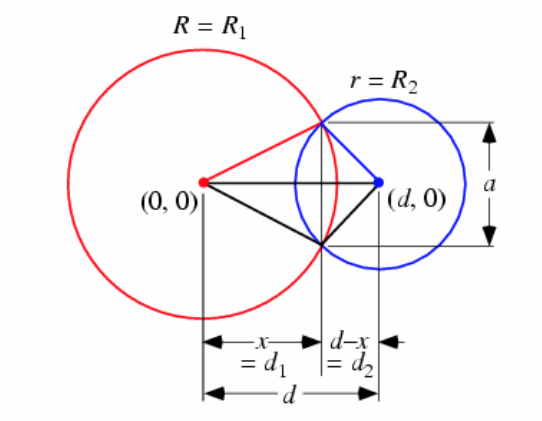


Fig 6. Overlap area computation

$$O = r^2 \cos^{-1}\left(\frac{d^2+r^2-R^2}{2dr}\right) + R^2 \cos^{-1}\left(\frac{d^2+R^2-r^2}{2dR}\right) - \frac{1}{2}\sqrt{(-d+r+R)(d+r-R)(d-r+R)(d+r+R)} \tag{11}$$

3.2 Bounding box for algorithm simplification

However, the intrinsic complexity of Eq. (11) mainly due to the calculation of arcos() makes this algorithm somewhat time consuming. Hence, we introduce the bounding box method to simplify this approach. For the simulation results and differences of the circle based approach and simplified one, the detailed discussion will be presented later in this paper.

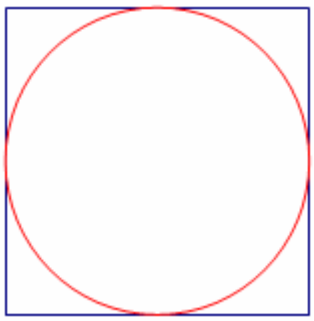


Fig 7. Circle and its bounding box

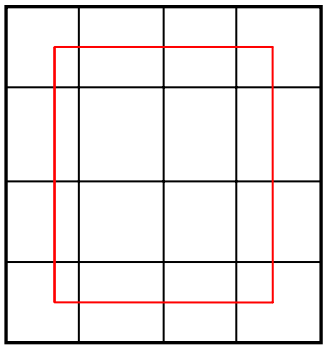


Fig 8. Square overlapping

For each circle, including pixel circles and detector circles, we make a bounding box with the same center, and the side length equal to the diameter of the circle. Then, the implementation is exactly the same as Eq. (9)(10) except that the overlap area is calculated by

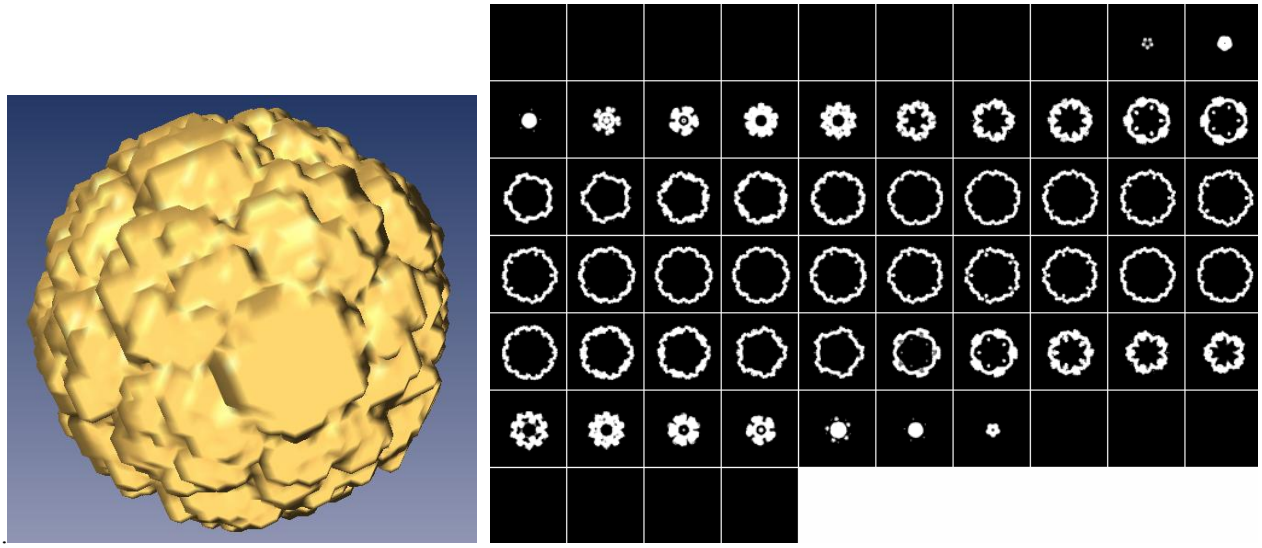
$$O = lx \times ly \qquad (12)$$

where lx is the overlap length in x axis direction, ly is the overlap length in y axis direction. Similarly, the area A is calculated as the size of the bounding box. Compared to the circle to circle overlapping, this simplified method greatly reduced the computing complexity.

4. IMPLEMENTATION AND RESULTS

4.1 Quality and artifacts

We use synthetic icosahedrons model of CCMV virus for testing, as shown in Fig 9. Icosahedron is very common structure of viruses. We will make detail comparison between the pixel driven results and distance driven results. The results of pixel driven approach are generated by Xmipp package[8], which is popular software package in electron microscopy area. We implement our own approach based on Linux OS with GCC 4.0 compiler, and moderate optimization is applied. The computation is done on a computing server with dual Xeon 3.2 GHz CPU and 2 GB RAM



(a) (b)

Fig 9. 3D icosahedrons model: (a) 3D visualization from AMIRA[12]. (b) visualization by slices

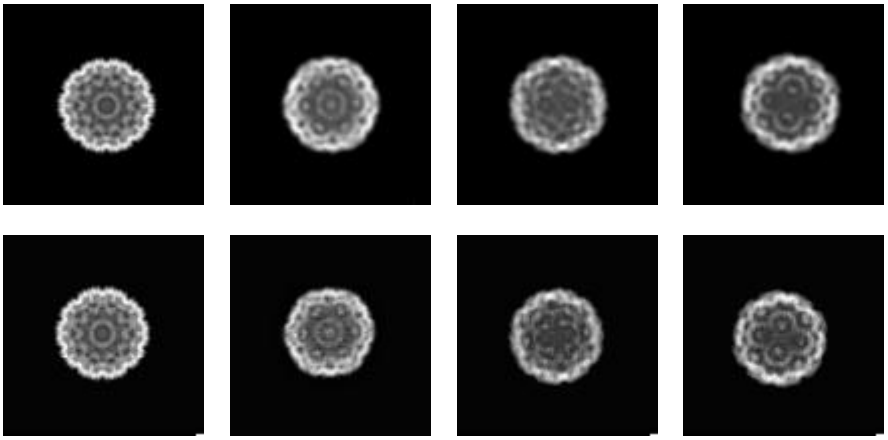
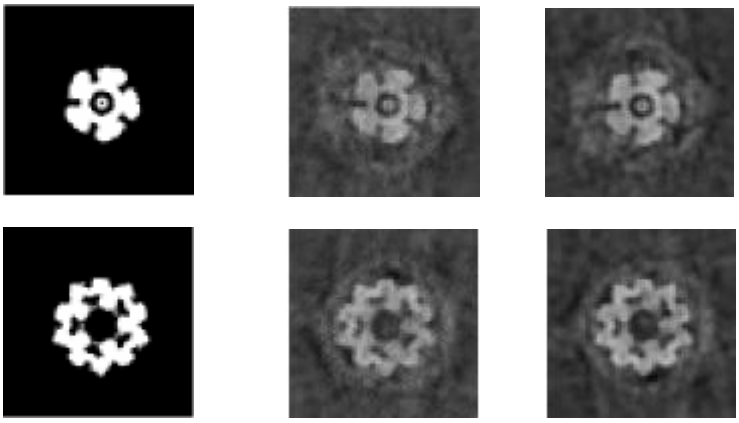


Fig 10. Projection sinograms: upper row shows the results from pixel driven approach, bottom row shows the results from distance driven approach. The two figures at same column are generated with same euler angle. 100 by 100 for each image size.

We applied 3 euler angles to describe the rotation of the model (or the projection view). We uses same set of rotation angles for different approaches, the Euler angles are uniformed randomly distributed. Fig 10 shows the results from pixel driven approach (from Xmipp package) and the modified distance driven approach. Even though the differences are not very obvious, still we can see the distance driven approach provide sharper projection image with high contrast.

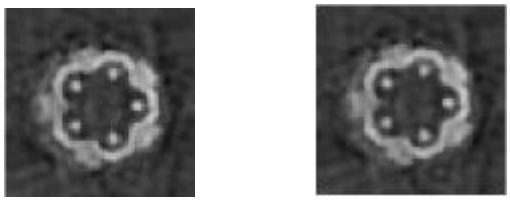
For reconstruction, we use WBP (weighted back projection) for both approaches. Here, the same ramp filter is applied before the back projection.



(a) (b) (c)

Fig 11. Artifacts comparisons between same slice(the reconstruction results in (b) and (c) are from 64 projections): (a) from original 3D image (b) from Xmipp package (c) our approach based on circle overlapping

As shown in Fig. 11(b), which is the result generated by Xmipp package, the image background is noisy, but in (c), which is the result from our approach, the image quality is better, the noises are eliminated and the contrast looks higher. Our approach provider clearer structure.



(a) (b)

Fig 12 (a) circle based distance driven method (b) simplified method based on bounding box

Fig 12 (b) is generated by the simplified approach using bounding box method. Compared to Fig 12 (a) from circle overlapping based approach, this simplified method provides very similar result. This result proved that this simplified approach is more practical for implementation. Fig 13 and Fig. 14 shows the slice-by-slice view of the reconstructed CCMV virus model using XMIPP and the modified distance driven method, both are based on 1000 projections. For large number of view samples, similar reconstruction results are observed as expected. Fig 19 shows the 3D visualization result of the reconstruction results using distance driven back projection approach. Accurate 3D structure is obtained..

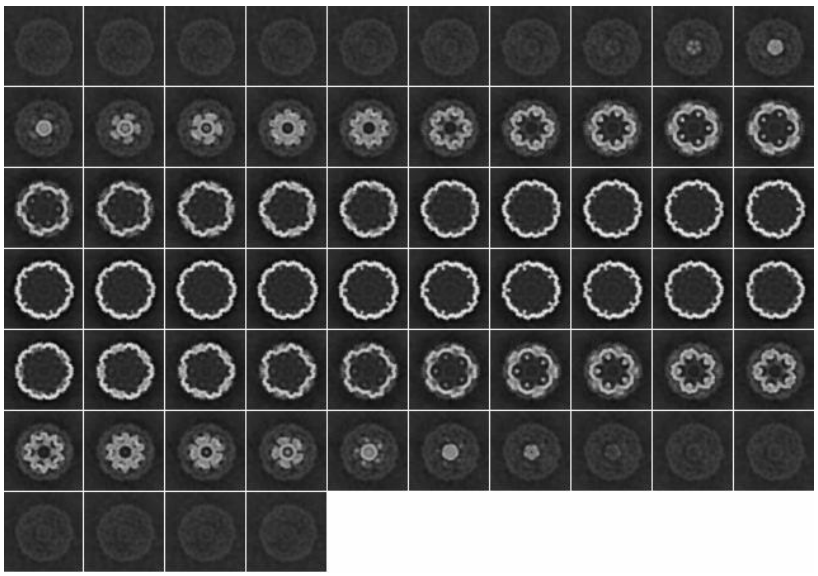


Fig 13. Reconstruction result from Xmipp WBP (1000 projections)

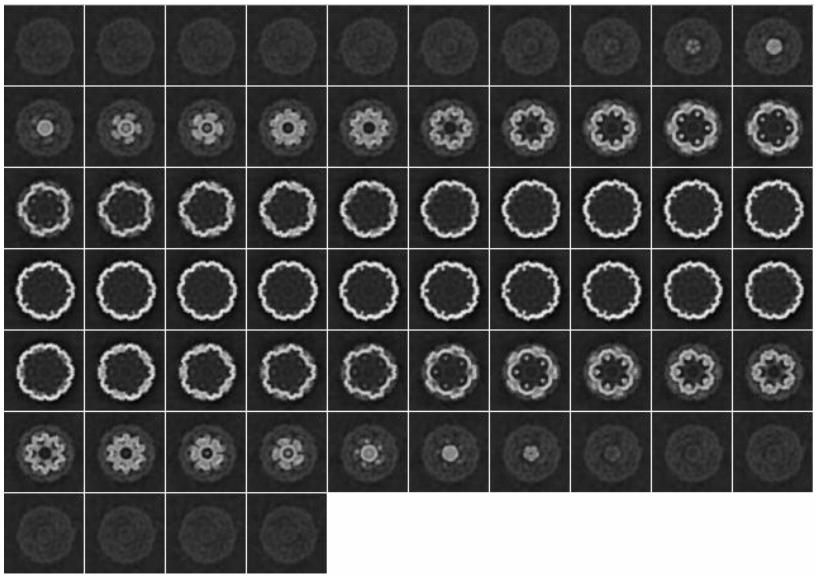


Fig 14. Reconstruction result from the simplified distance driven approach (bound box method with 1000 projections)

4.2 Computational Performance

In order to compare the computational burden of the modified algorithm compared to the XMIPP, 64 by 64 by 64 icosahedrons model is used for testing. Size of projection image is 100 by 100. 1000 projection images are generated in projection process and used for back projection. The computations of projection and backprojection take very similar time. Compared to the Xmipp package, in the projection process, our approach takes 25second, and for the same process, Xmipp package take 108 second. This means the computation complexity of our approach has much better computational performance that the pixel driven method in Xmipp package.

5. CONCLUSION

The overlapping shape approximation is essential issue in distance driven approach. In Cryo-EM tomography, projection angles are randomly distributed over 4pi steradian. But in x-ray computed tomography, large number of

projecting angles are distributed within a small range. Hence, the direct use of conventional distance driven approach results in inaccurate approximation and the reconstruction artifacts. The modified distance driven method successfully overcomes the problem with better result than that of pixel driven approach. The low arithmetic complexity in the simplified method with bounding box make the new distance driven approach provide high performance in computation speed. We can conclude that in cryo-EM tomography, the modified distance driven algorithm is one of good approaches for accurate reconstruction with high computation performance.

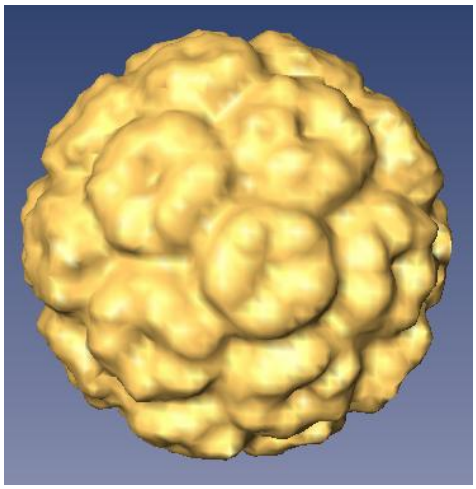


Fig 15. 3D Visualization by AMIRA (result from 1000 projection using our approach)

ACKNOWLEDGEMENT

This work was supported by a grant No. R01-2005-000-10035-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

REFERENCES

[1] Joachim Frank, "Single-particle imaging of macromolecules by cryo-electron microscopy", *Annual Review of Biophysics and Biomolecular Structure*, Vol. 31: 303-319, 2002

[2] Herman G , Image Reconstruction from Projections, Academic Press, 1980,

[3] Bruno De Man and Samit Basu, "Distance-driven projection and backprojection in three dimensions", *Physics in Medicine and Biology*, 2463–2475, 2005

[4] De Man B and Basu S , " Distance-driven projection and backprojection", *IEEE Nuclear Science Symp. Medical Imaging Conf. (Norfolk), 2004*

[5] Gordon R, Bender R, Herman GT. "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography", *J Theor Biol*, Dec;29(3):471-81, 1970

[6] A.H. Andersen and A.C. Kak, "Simultaneous Algebraic Reconstruction Technique (SART): a superior implementation of the ART algorithm," *Ultrason. Img.*, vol. 6, pp. 81-94, 1984.

[7] Joachim Frank. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Academic Press, 2006

[8]C.O.S. Sorzano, R. Marabini, J. Velazquez-Muriel, J.R. Bilbao-Castro, S.H.W. Scheres, J.M. Carazo, A. Pascual-Montano, "XMIPP: a new generation of an open-source image processing package for Electron Microscopy", *Journal of Structural Biology* 148(2) pp. 194-204, 2004

[9]A. C. Kak and M. Slaney, Principles of computerized tomographic imaging, IEEE Press, New York, 1988.

[10]Crowther, R. A., DeRosier, D. J., and Klug, A, "The reconstruction of a three-dimensional structure from projections and its application to electron microscopy", *Proc. R. Soc. London Sect A* 317**,** 319–340, 1970.

[11]Salvatore Lanzavecchia, Pier Luigi Bellon, and Michael Radermacher, "Fast and accurate three-dimensional reconstruction from projections with random orientations via Radon transforms", *Journal of Structural Biology* 128, 152–164 , 1999

[12] http://www.amiravis.com/