

Learning Frame Interpolation for Tilt Series Tomography

Alexander Rakowski, Jovany Merham, Lingge Li, Pirre Baldi and Joesph Patterson

University of California-Irvine, Irvine, California, United States

Electron tomography (ET) aims to reconstruct a 3D object from a series of 2D electron micrographs. This may be achieved in multiple ways, the most common of which are tilt series, slice and view and single particle reconstructions. Here we focus on tilt series tomography, where the 3D object is reconstructed by taking images at different tilt angles. Electron tomography, which is a subset of tomography, is founded on the work of Johann Radon, who showed that an n dimensional object can be reconstructed from an infinite number of $n-1$ dimensional projections.[1] However, experimental datasets are not infinite in size, which necessarily introduces artefacts and limitations in the reconstruction.[1]

In transmission electron microscope (TEM)-ET the primary limitations on the reconstruction stem from the tilt range and interval. Tilt range is typically limited to $\pm 70^\circ$ due to constraints imposed by the geometry of the pole piece, and specimen holders. This introduces an artefact known as the ‘missing wedge’, and results in an elongation in z (we define the x and y to be the plane normal to the electron beam at 0° tilt, with y as the tilt axis). The missing wedge may be reduced experimentally by specialised holders through novel tilt strategies [2], or in the data post processing [1]. Recent efforts include machine learning (ML) approaches which aim to fill in the missing information.[3]

A limited number of projections reduces the resolution in x and z direction, with more projections yielding a better reconstruction (if the projections are well aligned and of good quality). The choice in step size is dictated by a number of factors, most notably dose and time. The tilt series acquisition must be sufficiently speedy to be experimentally practical, and minimise sample drift which lowers the quality of the reconstruction. While the electron dose must be budgeted effectively such that each image has sufficiently high signal-noise ratio, and provide maximal number of projections, while being sufficiently low so as to minimise beam induced changes and damage in the sample. These factors make optimal experimental design challenging, and limit the ability capability of tilt series tomography for beam sensitive materials. In this work we focused on the limitations imposed by the tilt interval, and demonstrate the use of ML for frame interpolation.

(Video) Frame interpolation has been a well-studied problem in the field of computer vision and deep neural networks have demonstrated success on various benchmark data sets using different approaches.[4,5] Current popular approaches for deep neural networks are to learn optical flow between images frames or to predict convolution kernels over image frames. While these approaches can create realistic interpolations for natural images, these trained models cannot be directly used on electron microscopy data. This is due to certain symmetries and unique features to the micrographs, which are not present in the training data for these current networks. Consequently, it was necessary to take a different approach for the electron microscopy domain.

The aim of this work was to create a ML algorithm which given two frames for a tilt series was able to robustly interpolate the intermediate frame e.g. given images for 0° and 2° , produce 1° . This would allow for collection of fewer experimental projections, permitting increased flexibility in allocation of electron beam dose, and experimental time. The overall objective of this task and the approach taken is illustrated in Figure 1. In this paper we concentrate on metallic nanoclusters, as these are some of the most industrially important particles where tilt series ET is routinely used.

The model was trained on a dataset of simulated high angle annular dark field HAADF scanning TEM (STEM) images of metal nanoclusters. The clusters were generated with different symmetries (octahedral, icosahedral, decahedral), elemental compositions (Pt, Ag, Au, Cu), sizes, and with random orientations (x , y , and z). The

number and variety of clusters represent a sufficiently large diverse dataset to show ensure the model generalizable. For each cluster HAADF STEM images were simulated for each tilt angle in $\pm 90^\circ$ with a tilt interval of 1° . The images were simulated using PyQSTEM[6], a python wrapper for QSTEM[7], on a HPC cluster to efficiently generate a large dataset. Poisson noise was added to the images post simulation.

The tilt series were first split, into a training, validation and test datasets, and further subdivided by removing intermediate frames (Figure 1). The removed frames served as the ground truths for the model to compare to its interpolated output frames. The model is composed of convolutional layers which down sample the image, and dense layers. The output of the dense layers are subsequently reshaped in to the interpolated image.

The performance of the model was compared against reconstruction with fewer frames (i.e. without intermediate frames), and a simple linear interpolation of two neighbouring frames. Several performance criteria were assessed, including robustness to noise, performance vs. number of missing frames (i.e. larger tilt intervals).

The bespoke model was able to interpolate intermediate frames, for metallic nanoclusters of different shapes and elemental compositions. It was shown to be robust to noise and offers promise as useful tool for tilt series electron tomography, by reducing the number of required projections to be collected. This work should be expanded to train on more complex nanoclusters such as bimetallic, doped, core shell and hollow nanoparticles etc. Furthermore, as nothing in the model was tailored to be specific for metallic nanoclusters, it should be transferable to other specimen types.

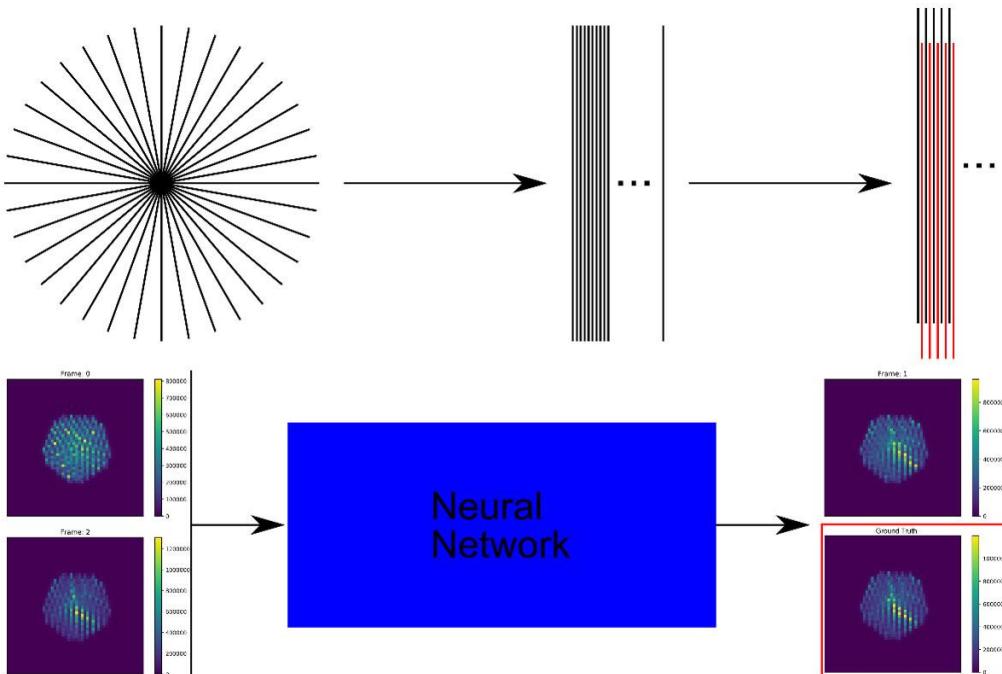


Figure 1. Top) Illustrative example of how the tilt series tomography HAADF STEM images are stacked and split prior training, the red frames are removed and the models aims to interpolate these missing frame. Different splice intervals can be used. Bottom) Illustration of the input and output of the model. Two neighbouring frames are input, and a single interpolated frame is output. The interpolated and ground truth are compared and error is feedback to train the model.

References

1. Midgley, P.A. and Weyland, M., 2003. 3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography. *Ultramicroscopy*, 96(3-4), pp.413-431.
2. Arslan, I., Tong, J.R. and Midgley, P.A., 2006. Reducing the missing wedge: High-resolution dual axis tomography of inorganic materials. *Ultramicroscopy*, 106(11-12), pp.994-1000.
3. Ding, G., Liu, Y., Zhang, R. and Xin, H.L., 2019. A joint deep learning model to recover information and reduce artifacts in missing-wedge sinograms for electron tomography and beyond. *Scientific reports*, 9(1), pp.1-13.
4. Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., Van Der Smagt, P., Cremers, D. and Brox, T., 2015. Flownet: Learning optical flow with convolutional networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2758-2766).
5. Niklaus, S., Mai, L. and Liu, F., 2017. Video frame interpolation via adaptive separable convolution. In *Proceedings of the IEEE International Conference on Computer Vision*(pp. 261-270).
6. <https://github.com/jacobjma/PyQSTEM>
7. Koch, C.T., 2002. *Determination of core structure periodicity and point defect density along dislocations*.