Applying denoising algorithm to improve alignment for burst imaging data

Xiaochen Du & Madhurima Vardhan
Duke University, Durham, NC, USA
E-mail: `xiaochen.du@duke.edu`
E-mail: `madhurima.vardhan@duke.edu`

26 November 2019

Abstract.
The development of burst imaging in cryo-electron microscopy (cryo-EM) has resulted in a whole new realm of possibilities and challenges for researchers and practitioners in the field. One such challenge is local drift alignment, whereby particles move between consecutive frames. Current approaches adopted by cryo-EM packages utilize a global drift alignment algorithm. Such a method takes the average of all frames and compares the cross-correlation of each frame gradually shifted in transverse axes with the average frame before determining optimal shift with respect to the average. This process results in a new average image and the process is repeated until convergence. However, such a process does not always yield optimal results. We propose applying a demonising algorithm to multiple frames as noted in cryo-EM movies. In particular, we specifically include the Block-matching and 3-dimensional (BM3D) filtering method as part of the alignment loop. To this end, we denoise the complete micrographs from individual frames and use the frame-to-frame alignment approach to determine the average frames and 3D reconstruction. The findings of this study demonstrate that our proposed method is effective to improve 3D reconstruction. Furthermore, this method has the potential to be applicable to any macromolecular particles, irrespective of their size and type. This proposed approach is easily extensible to determine local alignment and local translations on a per-particle basis.

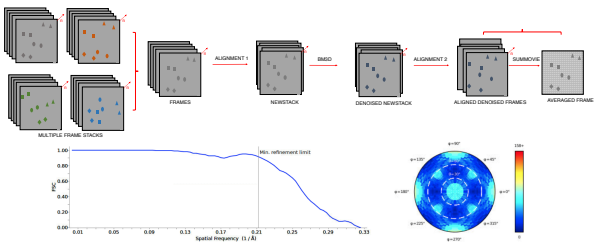

Figure 1. Central illustration. Top panel demonstrates iterative alignment and denoising approach. Bottom panel demonstrates achieved resolution in 3D reconstructed Apoferritin data set

Keywords: Movie alignment, BM3D denoising algorithm, Single particle reconstruction

1. Introduction

Structural determination of macromolecules continues to be revolutionized with single particle cryo-electron microscopy (cryo-EM). The workhorse behind reaching near atomic resolutions of many different protein complexes [1, 2] is the advent of direct detector devices and remarkable improvements in image-processing methodologies [3]. Specifically, the use of complementary metal oxide semiconductors in the detectors lends itself to the acquisition of exposure series 'movies' [4]. Such developments have yielded single particle reconstruction methods to be regarded as highly promising approaches in structural design and drug discovery.

The three-dimensional single particle reconstruction requires the projections of several-thousand high-quality particles. These particles are identified from micrographs acquired through detector device cameras, which suffer from low signal-to-noise ratio (SNR) and contamination artifacts. The low SNRs in electron micrograph frames can be explained by the low exposure rates, which are typically on the order of 1 to 3 $e^-/\mathring{A}^2/frame$ on the specimen, and corresponds to 2 to 5 $e^-/pixel/frame$ on the detector [5].

Interestingly, the direct-detector devices allow for micrograph movie acquisition due to two primary reasons 1) specimen stage drift resulting from microscope instabilities and 2) electron beam-induced movement. The latter is the consequence of the deformation of the ice layer, which results in somewhat non-homogeneous local movements within a particular micrograph and causing the images to blur [4, 5].

As high-quality particle projections are required for high-resolution 3D reconstruction, individual movie frames need to be aligned. However, this process of movie frame alignment suffers from 1) low SNR, 2) presence of fixed pattern noise in images and 3) errors in sensor gain normalization [5, 6].

One such alignment method is rigid body translational motion alignment. This approach has been previously investigated to determine the 'optimal' frame-to-frame translations employing mathematical concepts of over-determined linear equations and matrix algebra [7]. While the rigid body translation determination has been used to reconstruct high-resolution macromolecular structures [7, 8], it is not extensible to entire cryo-EM movie frames of large particles. The reason for that lies in the fact that beam-induced movement of ice-embedded proteins is dependent on local translations of different particles in different frames [4]. This explanation can be further expanded by analyzing tilt pairs of images which reveal that rotation of specimens occurs due to shifts in the ice layer [9, 5]. In summary, it is important to keep in mind that the beam-induced sample motion comprises of two parts, uniform whole-frame motion and non-uniform local motions. Thus, it is important to develop methodologies that take into account both these approaches and are not limited by the size and type of the particles being reconstructed.

Here in this work, we propose a generalizable methodology for determining optimal frame alignments. We couple alignment algorithm with the image denoising algorithm to calculate optimal alignment on a frame-to-frame basis in cryo-EM movie datasets. As such, this approach is extensible to small and large-sized particles. Furthermore, it can be easily extended to determine local alignments on a per-particle basis. However, in this work we limit our denoising and alignment approach for frame-to-frame assessment. To determine the ground-truth of our methodological advancement, we compare the resolution of resulting 3D reconstruction with appropriate controls. Overall, we believe that by revisiting established methods in image-denoising and alignment allows us to unlock the tremendous untapped potential of image-processing and further advance the field of single particle reconstruction using cryo-EM.

2. Related Work

Several studies have investigated methods to determine optimal alignment of beam-induced motion and obtain improved resolution from cryo-EM [5, 7, 6, 10, 11]. Li *et al* use a least squares method whole frame alignment to determine frame-to-frame translation by creating a system of over-determined linear equations using cross-correlation functions [7]. Despite the application of this approach, the authors note that this method could not align image regions smaller than 2000x2000 pixels for movies. Furthermore, for particles larger than 1 MDa local specimen rotation resulting from beam-induced motion are significant and must be taken into account [9].

To address local specimen motion, another study aligned frames on a per-particle basis and is presented in the software package *Relion* [12]. Briefly, the central idea of this study is based on computing rolling averages of frames and ultimately resulting in high SNR. However, without necessarily taking into account true particle trajectories and fitting linear trajectories with constant velocities for small particles, this approach is prone to errors as linear trajectories may not be a suitable approximation [5, 12]. Furthermore, the implementation is contingent to use of the *Relion* package, limiting wider adoption.

Yet another study proposed the use of an objective function comprising of the Fourier transform correlation of individual frames. The partial derivatives of the objective function is then used in

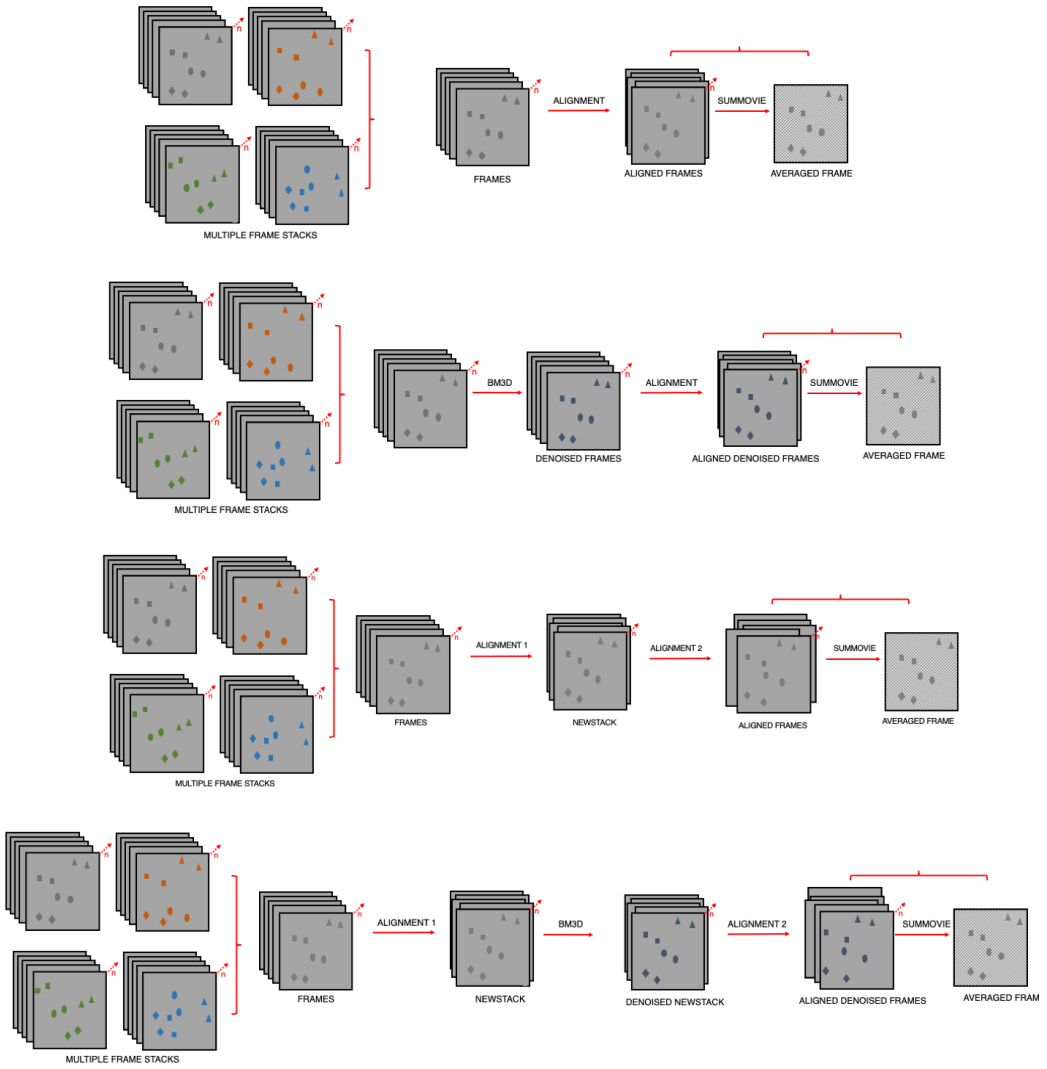

Figure 2. Experimental approach. For 20 different Apoferritin frame stacks (left side of each panel) successive rounds of alignment and denoising were performed. Top two panels demonstrates one round of alignment using UnBlur and denoising with BM3D along with the respective control. Bottom two panels demonstrates two iterative round of alignment using UnBlur and denoising with BM3D along with the respective control.

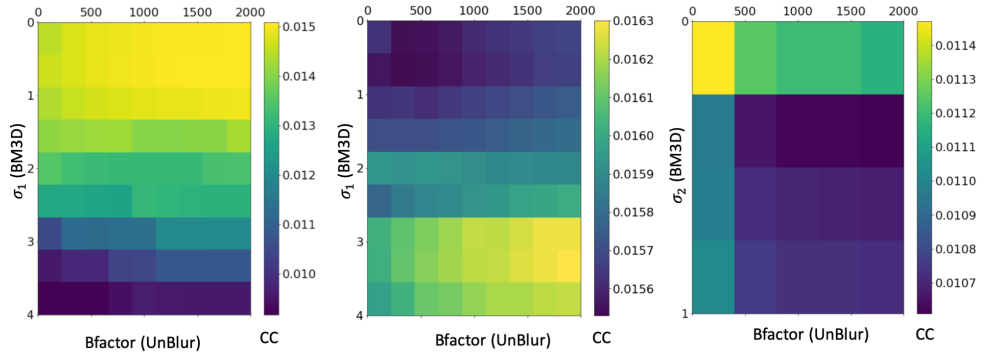

Figure 3. Heatmap of cross-correlation varying the σ_1 and B-factor parameters for different stacks. Representative stacks are shown for different experiments performed with a) Denoising and alignment with half BM3D b) Iterative alignment and denoising with half BM3D c) Denoising and alignment with full BM3D

Movie stack #	Denoising and alignment with half BM3D			Iterative alignment and denoising with half BM3D		
	Control (σ_1=0)	Half BM3D + UnBlur + summovie		Control (σ_1=0)	UnBlur + newstack + half BM3D + UnBlur + summovie	
	B-factor	σ_1	B-factor	B-factor	σ_1	B-factor
1	2000	2.5	2000	0	3.0	0
2	2000	2.5	2000	2000	3.0	2000
3	2000	1.5	2000	2000	2.5	2000
4	2000	1.0	2000	2000	1.0	2000
5	1750	0.5	1750	0	3.0	0
6	2000	0.5	2000	1750	3.5	1750
7	2000	1.5	2000	1250	3.5	1250
8	500	0	500	0	3.0	0
9	500	2.0	500	500	4.0	500
10	2000	0	2000	2000	4.0	2000
11	250	0	250	250	4.0	250
12	0	0	0	0	4.0	0
13	250	1.0	250	0	3.5	0
14	0	1.5	0	0	3.5	0
15*	N/A	N/A	N/A	N/A	N/A	N/A
16	2000	0	2000	2000	3.5	2000
17	2000	0	2000	500	4.0	500
18	250	1.5	250	0	4.0	0
19	1250	2.5	1250	1000	4.0	1000
20	0	1.0	0	0	4.0	0

Table 1. Combinations of σ_1 and B-factor used to obtain best cross-correlation score for each frame. *Excised after CTF estimation.

an iterative optimization algorithm to compute the per-frame translation values. The optimized objective function is then used to calculate per-frame basis local alignment of individual particles [5]. This study improves over the method by Li *et al* as it is able to work with smaller image boxes up to 256x256 pixels [5]. Despite these advantages, this technique assumes particles in close proximity to have similar motion trajectories and corrects for these motions early in the cryo-EM pipeline before 3D refinement [5, 10].

A different alignment registration strategy was proposed in [13] by Grant and Grigorieff, wherein the shifts were iteratively refined against current frame with average sum reduced until a converged solution was reached. In this approach all low frequencies are permitted by using strong low-pass B-filter. The particle trajectory is measured using spline-smoothing. Such an assumption where drift between individual frames is considered to be a smoothed spline works well for charge-driven particles, but does not hold for particles following Brownian motion [13, 6]. To further this approach, in the *Zorro* software package sub-frame registration was evaluated to determine alignment performance on stack registration, however this implementation did not yield expected performance gains [6].

More recent attempts used an experimental data-driven approach by describing sample motion as local deformation and computing the projection of local motion as a time-varying 2D polynomial function [10]. In this (*MotionCor 2*) approach, local patches are further used to iteratively compute local translations which are fit to a polynomial function and remapped to each pixel. This study demonstrated unequivocal results that local particle alignment can improve overall 3D reconstruction resolutions without adding tremendous compute time. However, the necessary dependence on experimental data instills subjectivity in obtained results. To alleviate this subjectivity, the *Warp* package [11] uses a very similar technique to *MotionCor 2*, but it relies only on the parameter grid resolution assumption.

All the above strategies are concrete blocks that have revolutionized cryo-EM analysis and pushed the achievable resolution limits. This work builds on these ongoing efforts. However, we take a different approach in contrast to existing techniques by adding a denoising algorithm and thereby improving cryo-EM image processing. The Block-matching and 3D filtering (BM3D) method is a state-of-the-art frequency-space denoising algorithm. Since its introduction in 2007, it's been widely used in many areas and applications of image denoising [14, 15]. To the best of our knowledge, a coupled denoising and alignment approach has not yet been applied in cryo-EM image alignment.

		Denoising and alignment with half BM3D		Iterative alignment and denoising with half BM3D	
		Control (σ_1=0)	Half BM3D + UnBlur + summovie	Control (σ_1=0)	UnBlur + newstack + half BM3D + UnBlur + summovie
# of particles	Picked	1978	1980	1976	1980
	Refined	1919	1866	1863	1894
Iteration #	1	3.45	3.49	3.49	3.41
	2	3.49	3.45	3.45	3.45
	3	3.45	3.45	3.45	3.45

Table 2. Achieved resolutions of 3D reconstructions obtained with different refinement levels using combined denoising and alignment with respective controls.

3. Methodology

3.1. Denoising algorithm

We use the Block-matching and 3D filtering (BM3D) method as our denoising algorithm. For many applications of image processing, the BM3D algorithm is considered as an effective baseline. To summarize: BM3D groups similar and non-local image patches into a 3D array. The array is then filtered in the transform domain using different levels of threshold. After filtering, the image patches are then restored to their original positions and reweighed to form a denoised image. For a detailed description and understanding of the BM3D algorithm, we refer the reader to [14, 15]. Here in this work, we use the BM3D algorithm as our preferred denoising algorithm since this algorithm works better with image patches that have self-similarity.

The complete BM3D algorithm in Python was obtained from `https://github.com/legendzhangn/blog/tree/master/color_bm3d`. The original BM3D algorithm goes through two rounds of collaborative filtering in the transform domain: 1) the first round, which uses a hard thresholding whereby coefficients below a certain value are dropped to 0 and 2) the second round, which uses a Wiener filter-based thresholding to weigh each of the coefficients. The resulting improvement in real images, using the first round of filtering significantly improves the peak signal-to-noise ratio of the image, and using the second round further improves the denoising quality.

Here in this work, we adopted the collaborative filtering in the transform domain section of the algorithm for our project. In this implementation, we used discrete cosine transform (dct). However, our method is extensible to different pipelines and allow for either the hard thresholding (half BM3D) or both hard thresholding and Wiener filtering (full BM3D). For our work, we vary the σ_1 (hard thresholding) and σ_2 (Wiener filtering) values of the BM3D denoising portion of our algorithm.

3.2. Cryo-EM datasets

We used the 5G Apoferritin data set (EMPIAR-10146, 20 movies, $1.5A$/pixel) with a binning factor of 2 for our experiments (giving 3.0 A/pixel). We originally used this test data set in order to obtain a working pipeline; however, with the limits on pixel size we believe we may not observe significant improvement using our proposed framework. After obtaining a concrete pipeline, we removed the binning factor and used original high resolution images in the pipeline, which were subsequently used for 3D single particle reconstruction.

3.3. Denoising and Alignment

The UnBlur package from *cis*TEM [16] is used for global movie alignment. We used the following parameters: image pixel size = 3.0, output binning factor = 1, not applying exposure filter, set expert options, varying the B-factor. The lower the B-factor value, the more higher-resolution information is used in alignment. All other settings used were based on default values.

summovie (*cis*TEM) and newstack (IMOD [17]) were used for shifting movie frames based on results from UnBlur. summovie gives an average but newstack returns a shifted stack, allowing for further processing. For summovie, we used the settings: image pixel size = 3.0, do not apply dose filter.

Results were measured based on the correlation score function of ctffind (*cis*TEM) using the following parameters: pixel size = 3, acceleration voltage = 300, spherical aberration = 0, amplitude contrast = 0.07, min resolution = 40, max resolution = 6. All other settings were default values.

We performed three experiments with all 20 movies.

1) **Experiment 1**: half BM3D + UnBlur + summovie
   (a) $\sigma_1 = 0$ (control), B-factor from 0 to 2000 in 250 increments

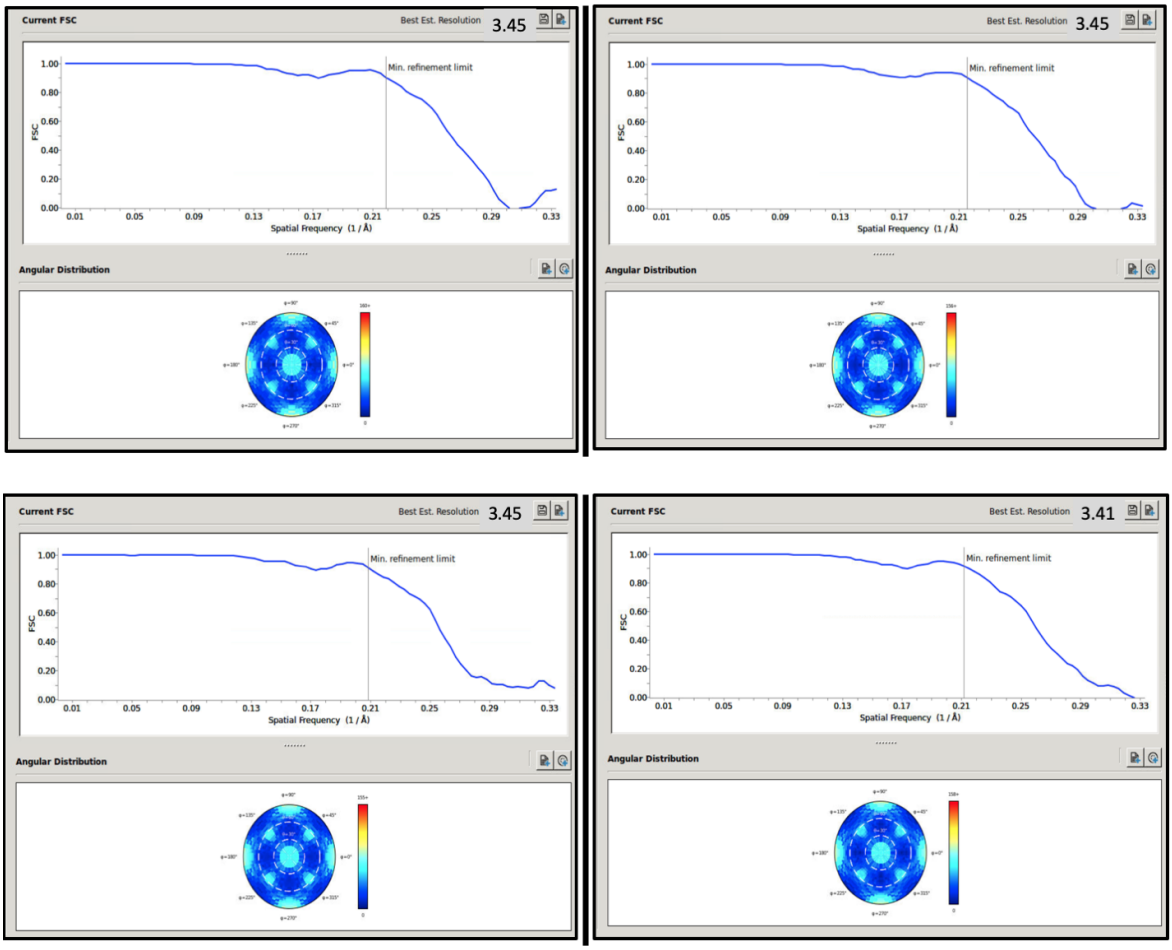

Figure 4. 3D Reconstructions. Top panel demonstrates achieved resolution with one round of alignment using UnBlur and denoising with BM3D along with the respective control. Bottom panel demonstrates achieved resolution with two iterative round of alignment using UnBlur and denoising with BM3D along with the respective control.

   (b) σ_1 from 0.5 to 4.0 in 0.5 increments, B-factor from 0 to 2000 in 250 increments

2) **Experiment 2**: UnBlur + newstack + half BM3D + UnBlur + summovie
   (a) $\sigma_1 = 0$ (control), B-factor from 0 to 2000 in 250 increments
   (b) σ_1 from 0.5 to 4.0 in 0.5 increments, B-factor from 0 to 2000 in 250 increments
3) **Experiment 3**: UnBlur + newstack + full BM3D + UnBlur + summovie
   (a) $\sigma_1 = 3.5, 4.0, 4.5$, $\sigma_2 = 0$ (control), B-factor from 0 to 2000 in 250 increments
   (b) $\sigma_1 = 3.5, 4.0, 4.5$, σ_2 from 0.25 to 1.0 in 0.25 increments, B-factor from 0 to 2000 in 250 increments

3.4. 3D Reconstructions

We used the most promising results from Section 3.3 i.e. taking the best result for each movie from experiment 1 and experiment 2 for 3D reconstruction. For the 3D reconstructions, as previously mentioned we used the original images without applying a binning factor. The 3D reconstruction provide the ground truth to compare achieved resolution by different experiments in Section 3.3. We used the *cis*TEM package for 3D reconstructions to provide an objective comparison between different sets of experiments; the *cis*TEM package offers a graphical user interface to obtain high resolution 3D reconstructions by processing cryo-EM images. Here, we briefly describe the employed 3D reconstruction pipeline and for a more in-depth understanding we refer the reader to cistem.org [16].

Broadly the *cis*TEM GUI comprises of 'assets' and 'actions'. Assets consist of data files and actions consist of all different processing steps. As shown in Figure 1, the 20 average .mrc image files were imported for case 1 and 2, along with their control. Special focus was kept on maintaining the controls

for each of the described experiments as can be noted in each of the experiments. Here we refer to controls as image files when BM3D denoising was not applied by setting the variable σ_1 to 0. After importing the files for all the cases, reconstructions were successively performed. The voltage was set to 300kV, pixel size 1.5$\mathring{A}$, spherical aberration to 0 and exposure per frame to 2.0 $e^-/\mathring{A}^2$. Since alignment (using UnBlur) and averaging (Summovie) were performed outside of *cis*TEM, we performed CTF estimation by setting number of frames to 2, box size 512 pixels and amplitude contrast 0.07. We then visually inspected to determine detectable resolution of Thon rings. Following this visual inspection, the 'bad' images were filtered out and the set of 'good' images were selected and used in particle picking. Particles were automatically picked using the 'Find Particles' action and setting maximum particle radius and characteristic radius to 50$\mathring{A}$ and peak threshold height to 4.0.

A refinement package was created as template and 2D classification was performed by setting particle weight to 440.0 KDa, particle diameter 120$\mathring{A}$ and 192 box size, O for octahedral symmetry and 10 for number of classes. Subsequently, ab-initio reconstruction and 3D refinement were performed changing the initial resolution to 8$\mathring{A}$. The resulting MRC formatted reconstructions with associated resolution and Fourier shell correlation curves were saved to analyze results.

4. Results

For denoising and alignment with half BM3D (experiment 1), we observed no clear trend on whether using BM3D improves alignment quality. Our heatmaps indicate, for some movies, BM3D in fact worsened the cross-correlation score. From Table 1, some of our best results have σ_1 of 0, indicating the best alignment was achieved without any BM3D denoising. However, there were other movie stacks that showed improvement in alignment with significant BM3D denoising (up to σ_1 = 2.5). There was also no clear trend on how B-factor affects BM3D performance. Nevertheless, most of the movies showed the best performance with a B-factor of 2000.

For iterative alignment and denoising with half BM3D (experiment 2), we note a clear trend that BM3D improved denoising performance. The best score results for all the movies were observed with σ_1 ranging from 3.0 to 4.0, with multiple best results coming from σ_1 = 4.0. These values are significantly higher than the values in experiment 1. Similar to experiment 1, there was no obvious trend in how B-factor affects performance. However, there are more movies that have a B-factor of 0 and fewer movies with a B-factor of 2000 compared with experiment 1. This observation is as anticipated because after one round of UnBlur alignment, it might be possible to obtain more useful information from the higher frequencies.

Since most of the best alignments in experiment 2 had σ_1 = 4.0, we used a range of 3.5 to 4.5 to do an initial exploration of σ_2 values for iterative alignment and denoising with full BM3D (experiment 3). However, the results suggest that using Wiener filtering only decreases the performance of the algorithm.

The best results for experiments 1 & 2 were obtained in Table 1, which were then used to obtain the 3D reconstruction results in Table 2. We find that between experiment 1a and experiment 1b, there was not a clear difference in performance, with each of the three refinement iterations giving the same set of FSC results at 3.45$\mathring{A}$. Such a result is not surprising given the results of experiment 1 described earlier in this section. For experiment 2, we did notice a significant improvement from the control when we added in BM3D denoising, boosting our resolution to 3.41$\mathring{A}$ from 3.49$\mathring{A}$.

5. Discussion

This study builds a framework to improve the alignment by calculating frame-to-frame translations in burst imaging data. The major findings of this work are: 1) Coupling iterative alignment and denoising relative to alignment alone improves achieved resolution of 3D reconstructions 2) Excessive denoising potentially masks the translations between successive frames in movie data. 3) While in this study applied to whole frames, this approach offers seamless integration to align frames on a per-particle basis.

We were able to achieve resolution up to 3.41$\mathring{A}$ using iterative alignment and denoising, compared to 3.49$\mathring{A}$ with its control that did not have denoising implemented. This result provides compelling evidence that resolution improvements are possible by including denoising in the alignment loop. In contrast, without iterative alignment no significant difference in achieved resolution was noted. There can be multiple reasons for this observation. One viable explanation is that one round of alignment prior to BM3D denoising improves the similarity of patches, thereby improving denoising performance.

For our second finding, visual inspection of the BM3D denoised movies revealed that when larger values of σ_1 and σ_2 were used, the translation between successive frames becomes vanishingly diminished. This finding while counter intuitive is not entirely unexpected as the BM3D denoising algorithm works well on patches with notable self-similarity. In the absence of significant translation, such as in the

Apoferritin data set, these translations get masked. This finding might also explain why iterative alignment and denoising with full BM3D (experiment 3) showed a sharp decrease in alignment in the Wiener filtering step. Further investigation through fine tuning σ_2, e.g. in 0.01 increments might yield more conclusive insights into the denoising behaviour of BM3D.

Some caveats to bear in mind with presented results would be that while better reconstruction resolution with UnBlur + BM3D can be obtained, the higher frequencies could be compromised as we observe in Figure 4. We also note that for the 3D reconstruction, due to technical difficulties, we did not fix the number of particles picked/2D classes resulting in different number of particles used for the reconstruction step.

We also believe that our framework can be extended by first naively dividing the frames into smaller regions and then further fine tuning to a single particle level per frame. Such an implementation allows for increasing the scale of improved translations by several folds and thereby enhancing the resulting 3D reconstruction resolution. This expansive implementation would require more intricate integration of our existing pipeline with the *cis*TEM binaries, but would otherwise be a feasible option for future investigations.

6. Conclusion

In this study, we proposed coupled denoising and alignment framework to improve frame-to-frame translations. We used the Block-matching and 3-dimensional (BM3D) filtering method as part of the alignment loop. We test the applicability of our method by using the Apoferritin data set and assess the achieved resolution using our proposed framework. We note that denoising without iterative alignment does not change the achieved resolution, however, coupled with iterative alignment improves the resolution. This study builds on the existing multitude of alignment approaches and attempts to solve the grand challenge of achieving atomic resolution of different macromolecules.

7. Acknowledgments

This research was conducted as part of the CS 590 coursework at Duke University. We would like to thank Prof. Alberto Bartesaghi for providing this novel project idea and valuable discussions. We would also like to thank Hsuan-Fu Liu and Ye Zhou for overcoming key implementation challenges.

8. References

[1] Merk A, Bartesaghi A, Banerjee S, Falconieri V, Rao P, Davis M I, Pragani R, Boxer M B, Earl L A, Milne J L *et al.* 2016 *Cell* **165** 1698–1707
[2] Bartesaghi A, Merk A, Banerjee S, Matthies D, Wu X, Milne J L and Subramaniam S 2015 *Science* **348** 1147–1151
[3] Nogales E 2015 *Nature methods* **13** 24
[4] Brilot A F, Chen J Z, Cheng A, Pan J, Harrison S C, Potter C S, Carragher B, Henderson R and Grigorieff N 2012 *Journal of structural biology* **177** 630–637
[5] Rubinstein J L and Brubaker M A 2015 *Journal of structural biology* **192** 188–195
[6] McLeod R A, Kowal J, Ringler P and Stahlberg H 2017 *Journal of structural biology* **197** 279–293
[7] Li X, Mooney P, Zheng S, Booth C R, Braunfeld M B, Gubbens S, Agard D A and Cheng Y 2013 *Nature methods* **10** 584
[8] Cao E, Liao M, Cheng Y and Julius D 2013 *Nature* **504** 113
[9] Henderson R, Chen S, Chen J Z, Grigorieff N, Passmore L A, Ciccarelli L, Rubinstein J L, Crowther R A, Stewart P L and Rosenthal P B 2011 *Journal of molecular biology* **413** 1028–1046
[10] Zheng S Q, Palovcak E, Armache J P, Verba K A, Cheng Y and Agard D A 2017 *Nature methods* **14** 331
[11] Tegunov D and Cramer P 2019 *Nature methods* 1–7
[12] Campbell M G, Veesler D, Cheng A, Potter C S and Carragher B 2015 *Elife* **4** e06380
[13] Grant T and Grigorieff N 2015 *Elife* **4** e06980
[14] Dabov K, Foi A, Katkovnik V and Egiazarian K 2007 *IEEE Transactions on image processing* **16** 2080–2095
[15] Lebrun M 2012 *Image Processing On Line* **2** 175–213
[16] Grant T, Rohou A and Grigorieff N 2018 *eLife* **7** e35383 ISSN 2050-084X URL `https://doi.org/10.7554/eLife.35383`
[17] Kremer J, Mastronarde D and McIntosh J 1996 *Journal of structural biology* **116** 71—76 ISSN 1047-8477 URL `https://doi.org/10.1006/jsbi.1996.0013`

9. Appendix

Source code and results are available at: `https://github.com/dxc2007/cs590project`