Characterization of whole-frame motion correction algorithms for movie-mode images collected by direct detection devices

James Michael Bell[1,2], Isaac Forrester[2], Joanita Jakana[2], Steven J. Ludtke[1,2]

Structural, Computational Biology and Molecular Biophysics program, Baylor College of Medicine[1];
Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine[2]

Abstract

The recent emergence of routine near-atomic resolution structures in CryoEM can largely be attributed to the development of direct electron detection devices (DDD) as replacements for older fiber-optic coupled scintillator-CCD technology. In addition to dramatically improved modulation transfer characteristics, these new detectors are also capable of acquiring data at frame rates sufficiently high for individual electron event counting, which can then be averaged to produce movies at typically 10-40 frames/s. These movies can then be corrected for stage and specimen motion to further improve the high-resolution signal present in the final averaged image. While conceptually similar, the algorithms used to perform this motion correction vary widely. At present there are six algorithms in widespread use; however, because each alignment routine uses different criteria to guide, smooth, and otherwise bias frame translations toward the optimal alignment, results vary, sometimes significantly. In this study we compare alignments from each algorithm. Because frame alignment ultimately determines the obtainable resolution in single particle analyses, assessing the strengths and weaknesses of each algorithm may have a critical impact on the selection of a specific algorithm, and ultimately on quality of the final reconstructions. Through understanding the reasons behind the disagreements among packages, we may achieve insights to better design the next generation of alignment software.

Whole-frame motion correction

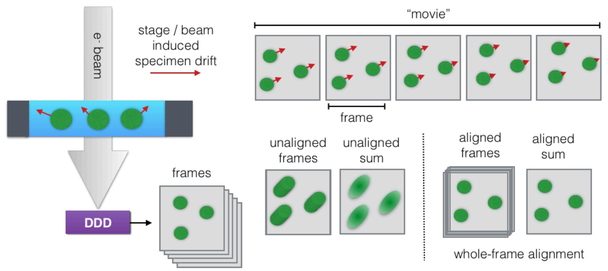

Figure 1: Visual summary of DDD movie acquisition and motion correction. *DDD movies consist of a stack of electron micrographs recorded in rapid succession with short exposure times. The resulting frames are aligned computationally to compensate for stage and beam induced motions of the imaged specimen (Brilot et al, 2012). In the absence of such motion correction, high resolution information is blurred.*

Most whole-frame motion correction algorithms rely on cross correlation methods to obtain translation vectors for each frame to bring them into register with a stationary frame or the frame average (Ripstein & Rubinstein, 2016). A non-exhaustive list of algorithms used in the field include Unblur (Campbell et al, 2012), MotionCorr (Li et al, 2013), DE process frames (Wang et al, 2014), IMOD's alignframes (Kremer et al, 1996), EMAN2's e2ddd_movie.py (Bell et al, 2016), and alignframes_lmbfgs (Rubinstein & Brubaker, 2014). Below, we provide preliminary benchmark data for each algorithm run on the same hardware except for differences in CPU/GPU requirements. Parallel results included as well where threading has been made possible. In the next panel, we visually display an assortment of data and frame trajectories measured by these algorithms.

Program	Package	G/CPU (#)	Walltime (min)
DE_process_frames-2.8.1.py	DE	CPU (32)	1.9 ± 1.3
e2ddd_movie.py	EMAN2	CPU (32)	5.9 ± 1.4
alignframes	IMOD	GPU	0.5 ± 0.2
alignframes_lmbfgs.exe	LMBFGS	CPU (1)	5.1 ± 2.2
unblur_openmp_7_17_15.exe	UNBLUR	CPU (1)	2.1 ± 0.3
dosefgpu_driftcorr	UCSF (Motioncorr v2.1)	GPU	2.3 ± 0.6

Table 1: Algorithm runtimes on super-resolution movies. *CPU hardware: Intel ® Xenon ® CPU E5-2650 v2 @ 2.60 GHz (16 physical cores, arch x86_64). GPU hardware: nVidia Tesla C2070 2.0 with 5301Mb memory. 128GB DDR3 RAM. Memory cache cleared between trials.*

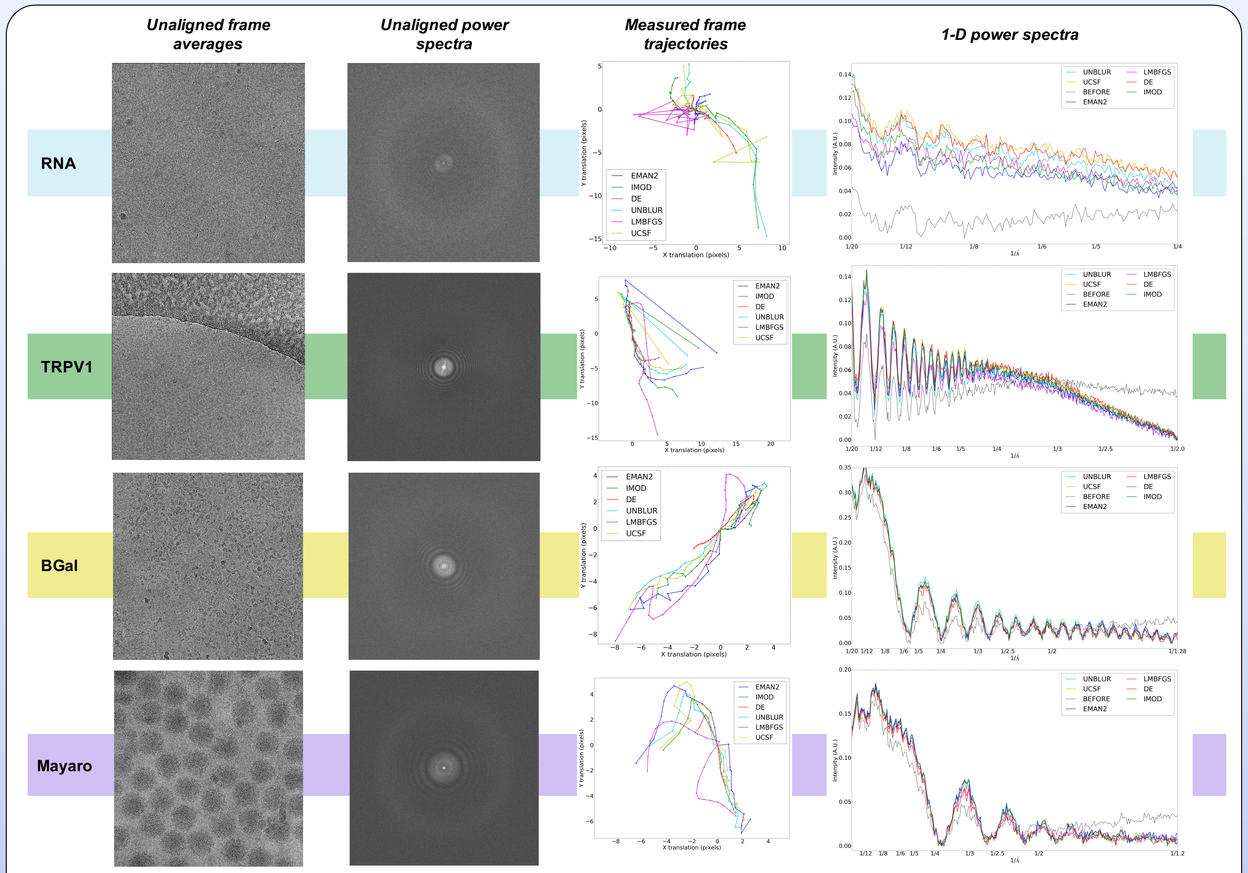

Figure 2: Comparative alignments of movie mode data. *Here we show results obtained by applying each motion correction algorithm to select movies from the EMDB and unpublished data collected by members of the NCMI. Note that all frame trajectories are plotted such that (0,0) corresponds to the middle frame of the movie. Power spectra computed using 2048 pixel tiles. These results show discrepancies in the results of the various alignment algorithms we tested**. We observe differences on the order of 2-8 pixels, corresponding to a distance discrepancy of 0-4Å. Given an Å/pixel of 0.6, differences greater than ~3Å can blur information beyond 2/3 Nyquist frequency. IPS are discussed in Fig. 3.* Because runtime parameters chosen on the basis of peer guidance and methods publications, we welcome any recommendations that might lessen the reported discrepancies.

Specimen	HIV DIS dimer RNA Kaiming Zhang unpublished	TRPV1 EMD-5778 Liao et al, 2013	β-galactosidase EMD-5995 Bartesaghi et al, 2013	Mayaro Virus Jason Kaelber unpublished
Weight (kDa)	30	380	470	52000
Microscope	JEOL JEM3200FSC	Polara 300	Titan Krios	JEOL JEM3200FSC
Detector	K2 (super resolution, bin x2)	K2 (super resolution)	K2 (super resolution)	K2 (super resolution)
Å/pix	1.198	0.61	0.64	0.65
Total dose	50	41	45	35
Dose / frame	2.0	1.37	1.2	1.4
Frames / sec	2.5	5	2.5	5

Table 2: Parameters linked to stage and beam-induced specimen motion. *It is currently understood that mechanical stage drift, electron radiation damage, and local charging, causes specimen motion observed in micrographs and DDD movies. Here we have curated a collection of 4 datasets with varying molecular weight and image recording parameters. All samples were applied to plasma cleaned/glow discharged Quantifoil holey carbon grids with the following geometries: HIV DIS dimer RNA (R2/1, 200mesh), TRPV1 (R1.2/1.3, 400 mesh), ß-galactosidase (R2/2, 200 mesh), Mayaro virus (R2/2, 200 mesh).*

Visual comparison

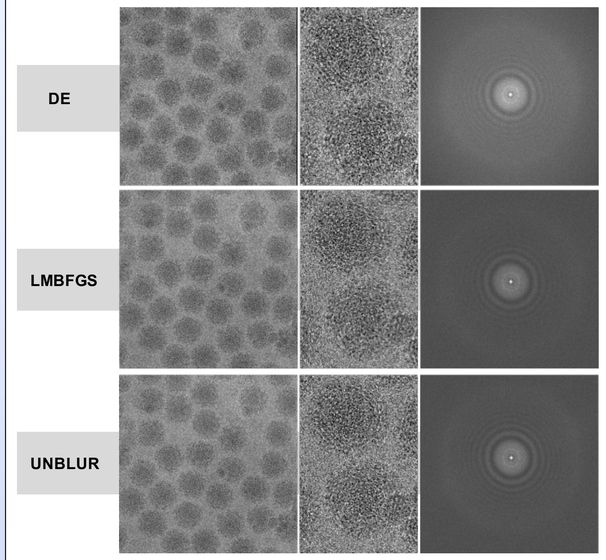

Figure 3: Representative frame alignment results from mayaro virus movie data. *Results chosen on the basis of their difference in motion measured by the 6 alignment programs referenced in Table 1.* Note that it does not appear possible to visually identify one of these results as superior to the others, despite the significant alignment differences. However, it is also clear that differences at the observed level cannot be equivalent

Conclusions

We have found that the existing widely used algorithms for CryoEM whole frame motion correction produce significantly different results on a significant fraction of data. From our current testing over a range of specimens, large discrepancies are observed in ~10% of movies, with smaller, yet significant discrepancies observed in as much as 30-40% of movies. Clearly frames which do not exhibit much motion, and hence do not require much correction, are not impacted by this problem. This difference is observed not only for small particles with very low contrast, but even on frames containing high-contrast virus particles. Some of this effect may be due to the need for local corrections in these frames, and the variation due to how this local motion is interpreted globally by each algorithm, but it is impossible to assess whether this is truly the case since the algorithms in question do not presently do local correction. The observed differences can be many pixels, meaning they extend to very low resolution, and may be dramatically impacting the quality of the affected images. While it has clearly been possible for CryoEM to achieve high resolution structures despite this issue, it does mean that there is substantial room for improvement in most projects, and that the community is in need of a more in-depth study of the reasons for these discrepancies.

Acknowledgements

This research is supported by a training fellowship from the Keck Center of Gulf Coast Consortium (GCC) Houston Area Molecular Biophysics Program (HAMBP, T32 GM008280) and the National Institute General Medical Sciences (NIGMS, R01GM080139, P41GM103832). Additionally, I am grateful to Kaiming Zhang and Jason Kaelber for their data contributions as well as Muyuan Chen for insightful discussion.