Original Article

Ultra-Fast Electron Microscopic Imaging of Single Molecules With a Direct Electron Detection Camera and Noise Reduction

Joshua Stuckner[1]* [iD], Toshiki Shimizu[2], Koji Harano[2]*, Eiichi Nakamura[2]* and Mitsuhiro Murayama[1]

[1]Material Science and Engineering Department, Virginia Polytechnic Institute and State University, 109A Surge, 400 Stanger Street, Blacksburg, VA 24060, USA and [2]Department of Chemistry, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Abstract

Time-resolved imaging of molecules and materials made of light elements is an emerging field of transmission electron microscopy (TEM), and the recent development of direct electron detection cameras, capable of taking as many as 1,600 fps, has potentially broadened the scope of the time-resolved TEM imaging in chemistry and nanotechnology. However, such a high frame rate reduces electron dose per frame, lowers the signal-to-noise ratio (SNR), and renders the molecular images practically invisible. Here, we examined image noise reduction to take the best advantage of fast cameras and concluded that the Chambolle total variation denoising algorithm is the method of choice, as illustrated for imaging of a molecule in the 1D hollow space of a carbon nanotube with ∼1 ms time resolution. Through the systematic comparison of the performance of multiple denoising algorithms, we found that the Chambolle algorithm improves the SNR by more than an order of magnitude when applied to TEM images taken at a low electron dose as required for imaging at around 1,000 fps. Open-source code and a standalone application to apply Chambolle denoising to TEM images and video frames are available for download.

Key words: denoising, image analysis, image processing, single-molecule imaging, transmission electron microscopy

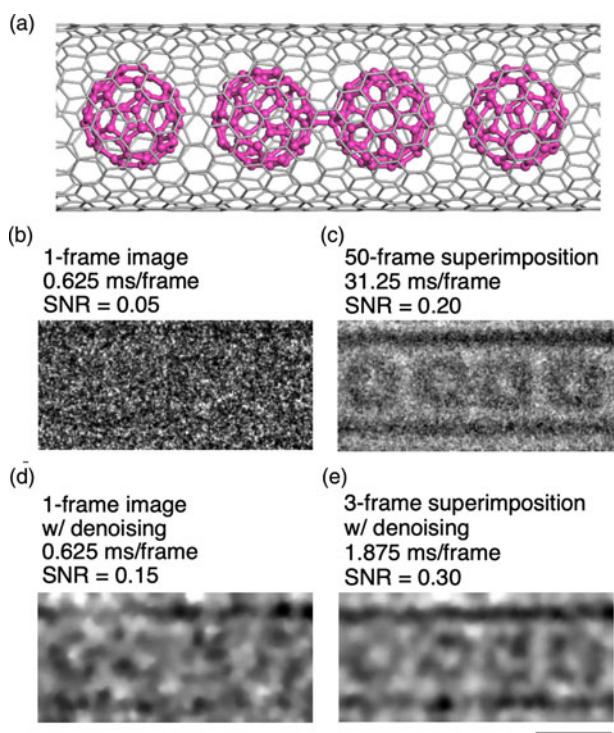(Received 20 December 2019; revised 15 June 2020; accepted 15 June 2020)

Introduction

Video recording of molecular motions and chemical reactions with a single-molecule atomic-resolution real-time transmission electron microscopic (SMART-EM) technique has emerged as a new technology for the study of mobile molecules and nanoscale assemblies (Nakamura, 2017). The SMART-EM technique enables us to record the time evolution of individual chemical events occurring in a one-dimensional (1D) test tube of a single-walled carbon nanotube (CNT) and on an outer surface of a CNT. This method can be used to perform statistical analysis of atomistic structure and dynamics of molecules over several hundred molecules. The *in situ* kinetic study of chemical reactions (Okada et al., 2017), mechanistic investigation of molecular crystal formation (Harano et al., 2012), and capturing and analyzing minute reaction intermediates (Xing et al., 2019) have illustrated the potential of the SMART-EM methodology in chemistry and nanoscience. This video technology has posed a new challenge of acquiring video images of fast moving or reacting molecules, so that we can visually and quantitatively study the dynamics of the observed chemical events. To this end, we need the highest possible frame rate with the highest possible image contrast.

The most advanced electron-counting direct-detection complementary metal oxide semiconductor (CMOS) cameras are capable of taking as many as 1,600 fps (Liao et al., 2014), but their latent potential has not, so far, been fully realized, because a high frame rate reduces the electron dose per frame, lowers the signal-to-noise ratio (SNR), and renders the molecular images practically invisible. For example, the 1,600-fps CMOS camera (K2-IS, Gatan) at 400,000× on screen magnification receives an electron dose per frame of only 10,000 electrons/nm^2 even at the detector-safe electron dose rate (EDR) of 160×10^5 electrons/nm^2/s. Hence, a single-frame molecular image of two [60] fullerene (C_{60}) molecules and their [2 + 2] cycloadduct (C_{120}) in a CNT (Figs. 1a, 1b) is entirely obscured by noise with a signal-to-noise ratio (SNR) of 0.05. For fast imaging of mobile molecules and molecular clusters, we cannot use the widely used contrast enhancement methods for static objects such as symmetry-imposing and lattice-averaging protocols (Zhu et al., 2017). To solve the problem, we decided to explore computational image processing of raw transmission electron microscopy (TEM) images (Kushwaha et al., 2012). There has been, however, no systematic study on denoising algorithms suitable for low SNR TEM videos. We report herein that the Chambolle total variation denoising algorithm (Chambolle, 2004; Munezawa et al., 2019) significantly improves the SNR (SNR = 0.05 in Fig. 1b to 0.3 in Fig. 1e) of low electron dose TEM videos and images. When combined with a 1,600 fps camera, this technique allows us to routinely record molecular motions as slow as 1.9 ms/frame while maintaining a high SNR. The potential of the camera with 0.625 ms/frame is realized nearly to its maximum as illustrated

Fig. 1. Noise reduction of single-molecule TEM images by superimposition and denoising algorithm. (**a**) Molecular model of two C_{60} molecules and a [2 + 2] dimer in a CNT. (**b**) A single-frame image of C_{60} molecules and a [2 + 2] dimer in a CNT at 1,600 fps on a scintillator-free CMOS sensor (K2-IS, Gatan). (**c**) Superimposition of 50 frames of the 1,600 fps video. (**d**) The same frame of b after denoising by the Chambolle algorithm. (**e**) Three-frame denoising and superimposition by the Chambolle algorithm. Scale bar is 1 nm.

for imaging of C_{60} molecules in the 1D hollow space of a CNT (Fig. 1a).

Materials and Methods

Sample Preparation

Three samples of [60]fullerene (C_{60}) in a CNT were used in this study. C_{60} molecules encapsulated in a CNT (C_{60}@CNT) (Okada et al., 2017) were chosen as the molecular specimen in this study because the hollow and spherical morphology of C_{60} is suitable for the quantitative evaluation of TEM image quality by SNR and edge analyses. The first two datasets, labelled as C60@CNT1 and C60@CNT2, were used to evaluate the effect of electron dose on image quality and evaluate the effect of various denoising algorithms in restoring image quality. The third dataset, labelled as C60@CNT3, was used to determine the effectiveness in restoring low signal images captured at ultra-high frame rates.

These C_{60}@CNT samples were prepared as follows: CNT powder was heated in air in an oven gradually from 296 to 793 K for 12 min, kept at 793 K for 1 min, heated from 793 to 823 K for 20 min, and kept at 823 K for 20 min to remove the terminal caps of CNTs oxidatively. For encapsulation of C_{60} molecules, the opened CNTs (0.2 mg) and C_{60} powder (0.2 mg) were sealed in a glass tube (Pyrex ϕ6 mm) under a pressure of 2×10^{-4} Pa and gradually heated from 296 to 573 K over 1 h, then to 673 K over 1 h, and kept at 673 K for 72 h. The resulting C_{60}-containing CNTs were separated mechanically from remaining C_{60} powder, washed with toluene to remove C_{60} from the surface, and dried

in vacuum. C_{60}@CNT thus obtained was a black solid (0.3 mg). We dispersed the C_{60}@CNT in toluene (0.05 mg/mL) in a vial in a bath sonicator for 1 h to soften it, so that we could secure good contact between the CNTs and the carbon surface of the grid (essential for temperature control). A 10-μL solution of the dispersion was deposited on a copper grid mesh with a lacy carbon (NS-C15, Okenshoji Co., Ltd.) placed on a paper that absorbs excess toluene. The TEM grid was dried in vacuum (60 Pa) to remove solvent for 2 h. To ensure reproducibility, we used the same sample grid of C_{60}@CNTs in a series of experiments.

TEM Imaging

Atomic-resolution TEM observation was carried out on a JEOL JEM-ARM200F TEM equipped with a spherical aberration corrector for imaging, Gatan OneView and K2-IS cameras, and at an acceleration voltage of 80 kV, under 1×10^{-5} Pa in the specimen chamber. Experiments were carried out on a double-tilt holder (JEOL EM-01030RSTH). To remove volatile impurities from the specimen, the holder was heated at 573 K for 30–60 min without electron irradiation before setting a desired temperature. After the stage temperature settled to the target value, we waited for an additional 30 min to minimize thermal drift.

The C60@CNT1 and C60@CNT2 datasets were imaged using the OneView camera under the varying EDR. The raw images were $2,048 \times 2,048$ pixels (binning 2 mode) with a 32-bit depth and a pixel edge length of 0.0213 nm at 1,000,000× magnification. For C60@CNT1, ten images were taken at ten different EDRs ranging from 1×10^5 to 130×10^5 electrons/nm²/s for a total of 100 images. For C60@CNT2, ten images were taken at 18 different EDRs ranging from 1×10^5 to 123×10^5 electrons/nm²/s for a total of 180 images. Each image was taken with an exposure time of 0.5 s (2 fps). The C60@CNT3 dataset was recorded by the Gatan K2-IS direct electron detection camera. The raw images were $414 \times 1,920$ pixels with a pixel edge length of 0.021 nm at 400,000× magnification. Each image was taken with an exposure time of 0.000625 s (1,600 fps). The OneView camera uses a scintillator to convert electrons to photons during the acquisition process, which introduces an intrinsic convolution and blurs the resulting image. The K2-IS uses direct electron detection to produce clearer images.

Image Pre-Processing and SNR Calculation

Before denoising, some preprocessing were performed on the raw images. Images were cropped to the relevant area and rescaled to 8-bit tiff-formatted files from the original 32-bit Digital Micrograph file format. 8-bit images were required by the implementations of some denoising algorithms used here. Rescaling to an 8-bit format resulted in some signal loss, but did not affect the relative amounts of denoising achieved with the different methods nor the calculated SNR values to the significant digits reported here. Some of the images contained illumination differences across the image due to the non-uniform spatial distribution of electron flux from the electron beam. This effect was more pronounced at higher EDRs. These pixel intensity trends in the image were removed, so that they would not affect the results. In order to remove these trends, a second-order 2D polynomial surface was fit to each image's pixel intensity values using a least-squares fitting method (Price-Whelan et al., 2018). The polynomial fit was subtracted by its mean value to normalize it to a mean of zero and then the result was added to the original

image. Sections of these preprocessed images for C60@CNT1 are shown in Figure 3 under the column labelled "Original".

Many methods exist for calculating the SNR in images, and although there is no agreed upon standard, methods have been developed for medical imaging (Dietrich et al., 2007), where the vacuum can be taken as the noise and a region of interest (ROI) is taken as the signal. Here, the region containing the C_{60} molecules represents the signal. The SNR can be calculated using equation (1), where S_{mean} and N_{mean} are the mean pixel intensity in the ROI and vacuum, respectively, and N_{RMS} is the root mean square of pixel intensities in the vacuum.

$$\text{SNR} = \frac{|S_{mean} - N_{mean}|}{N_{RMS}} . \tag{1}$$

Since the signal is calculated from an ROI, which may have been corrupted by denoising, this metric is useful for determining relative amounts of noise in an image, but does not necessarily show that the signal has been preserved. Large amounts of denoising can corrupt the signal by significantly reducing edge contrast or changing the apparent shape and size of features even as the SNR is increased.

Signal Preservation Calculation

Since the SNR does not show whether the original signal is preserved, it is also necessary to quantify the signal preservation after applying denoising. An ideal denoising protocol will remove noise while fully preserving the sample's features. In this work, the signal preservation was quantified by analyzing the preservation of molecule morphology and by analyzing the preservation of feature edges.

The morphology preservation of the C_{60} molecules was quantified with a signal score. For each algorithm, the optimum parameter setting was the setting that produces the highest signal score. The steps to calculate the signal score are illustrated in Figure 2. In this process, a template of the true C_{60} morphology was placed on each molecule, and the signal score was calculated based on the number of matching pixels between the template and the image. A circle detection algorithm using Hough transforms (Yuen et al., 1990) was used to automatically locate the C_{60} molecules in each image. C_{60} fullerene molecules have a diameter of 0.7 nm (Sloan et al., 2000), and the C60@CNT1 and C60@CNT2 images have pixel edge lengths of 0.0213 nm, and thus the circle detection algorithm was set to find circles with a radius of 17 pixels. Circles detected using this algorithm are shown in Figure 2a. A separate template was automatically generated for each image based on the detected circles. An example of such a template is shown in Figure 2b. Wherever a circle was detected, two concentric circles were placed in the template. The outer circle, which represents the molecule ring, had a radius of 17 pixels and was filled with a pixel value of 0 (shown as black). The inner circle represents the interior of the C_{60} molecule, had a radius of 7 pixels, and was filled with a pixel value of 1 (white). An inner radius of 7 pixels was chosen as an appropriate size after many manual measurements on the unprocessed high EDR images. All template pixels that lie outside the circles were set to a value of 2 (shown as gray). Separately, each image was binarized using Otsu's method (Otsu, 1979), such that darker pixels were set to a value of 0 and brighter pixels set to a value of 1 as shown in Figure 2c. Finally, the signal score was the number of matching pixel values when comparing the binary image with
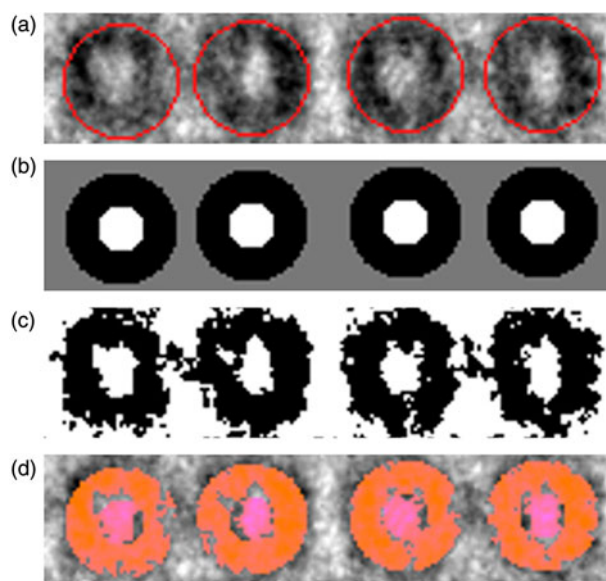
Fig. 2. Calculation of the signal score using circle detection and mask fitting. (**a**) An image from C60@CNT1 taken at the highest EDR with the detected circles outlined in red. (**b**) The mask generated from the circle fits where black, white, and gray pixels have a value of 0, 1, and 2, respectively. (**c**) A binarized version of the original image after thresholding with Otsu's method. Matching pixels in (**b**) and (**c**) are highlighted in orange in (**d**). The final signal score is the number of orange pixels in (**d**) divided by the number of circles detected and scaled between 0 and 1.

the template image. Matching pixels are shown in Figure 2d highlighted in orange. The signal score was normalized by dividing by the number of circles detected. The maximum possible score in this case is 889, which is the number of pixels that can fit inside a circle with a radius of 17 pixels. A score of 445 or lower means that there is no signal since completely random noise is likely to match half the pixels. The final signal score was scaled to values between 0 and 1 with zero representing 445 matching pixels.

Edge preservation was determined from the derivative of pixel intensity profiles. A line was placed parallel to the CNT running along the center of the C_{60} molecules. The pixel intensity profile is the plot of the pixel value along the line. In an image, an edge is the location of a sudden change in contrast, which here represents the edges of the molecules or the CNT. During denoising, these edges can become blurred, such that there is a gradual contrast change instead of a sharp change. If the contrast change at an edge is gradual, then it difficult to precisely locate the edge of a molecule, which greatly reduces the accuracy of size measurements. Edge sharpness was quantified by the magnitude of the derivative of the intensity profile across a molecule edge.

Denoising Algorithms

In total, nine denoising algorithms were applied and evaluated with and without downsampling for a total of 18 tests on each image in each dataset. The tested denoising algorithms were the mean, median, Gaussian, and bilateral (Tomasi & Manduchi, 1998) filters; the Chambolle (Chambolle, 2004) and Bregman (Osher et al., 2005) total variation denoising algorithms; nonlocal means denoising (Buades et al., 2005); Wiener–Hunt deconvolution (Hunt, 1971); and low rank approximation with singular value decomposition (Wold et al., 1987; Lin et al., 2010). Image corruption can include noise and blur where noise is the quasi-random change of individual pixel values, and blur is the
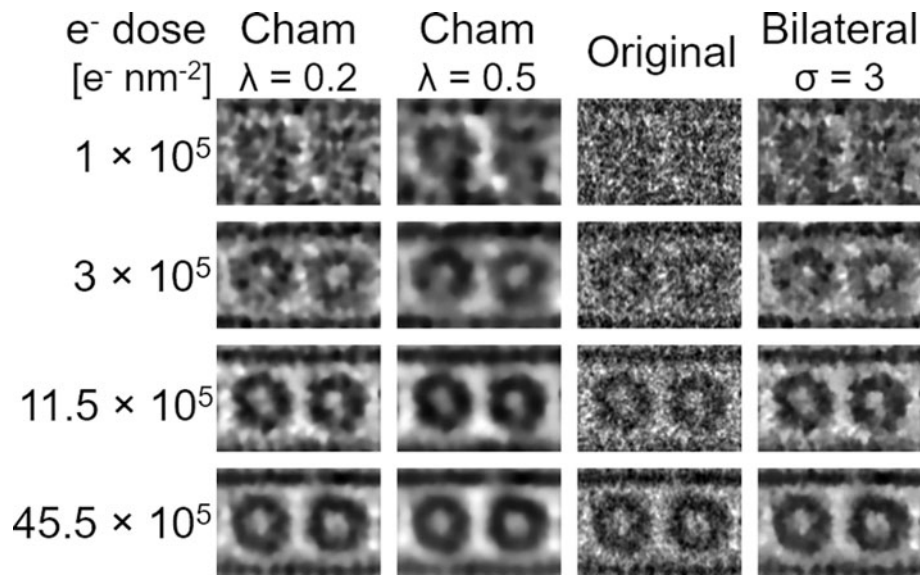
Fig. 3. Cropped images of C60@CNT1 at various EDRs before denoising (original) and after denoising with downsampling and the specified algorithm. Here, Cham specifies the Chambolle total variation denoising algorithm with the weighting parameter λ. The standard deviation of the bilateral kernel size is given by σ.

spreading of a point source of brightness described by the point spread function (PSF) and limits the intrinsic resolution of the microscope. Any denoising algorithm assumes a certain model, often a mathematical equation, which defines the relationship between the captured image, the nature of the corruption, and the ideal image that would have been captured by a perfect microscope system. The algorithm solves for or approximates the solution to the assumed ideal image based on this model.

The mean filter reduces noise by averaging together pixels within a specified neighborhood. The Gaussian filter uses weighted averaging where the weight of neighborhood pixels decreases with increasing distance according to the value of a normalized Gaussian distribution with a specified sigma of radial distance. The median filter replaces each pixel with the median value of pixels within a specified neighborhood, a technique that is less sensitive to extreme value pixels. The bilateral filter uses a weighted averaging method that weights neighborhood pixels based on both their spatial proximity and their intensity similarity. By considering intensity, this method preserves image feature edges such as molecule edges, unlike the mean, Gaussian, and median filters which indiscriminately blur edges. Deconvolution algorithms such as Wiener–Hunt deconvolution seek to reduce blur (Hunt, 1971). These methods, however, require prior knowledge of the PSF and perform poorly on noisy images. Some work has been done to estimate the PSF from microscopy images, but the work is ongoing (Liu et al., 2011; Dalitz et al., 2015; Roels et al., 2016). Here, a numerical Bayesian approach is applied to iteratively estimate the PSF for the Wiener–Hunt filter (Orieux et al., 2010). Without advances in estimating the PSF and in performing deconvolution on noisy images, this method is not expected to perform well, but is included here for completeness and comparison with past work on EM denoising (Kushwaha et al., 2012). The non-local means algorithm works well for denoising images with specific repeated textures (Buades et al., 2005). Instead of averaging together neighboring pixels, pixels are averaged when they are surrounded by similar patches, even if they are in different parts of the image. Low rank approximation with singular value decomposition has also been successful in

denoising (Wold et al., 1987; Lin et al., 2010). A singular value decomposition can be performed on an image matrix $\mathbf{I}$ using $\mathbf{I} = \mathbf{U\Sigma V}^\top$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, and $\mathbf{\Sigma}$ is a diagonal matrix whose entries are called singular values. A denoised image is reconstructed from a low rank matrix which is found by keeping only the top specified n values of $\mathbf{\Sigma}$, setting the rest to zero and solving for a new $\mathbf{I}$.

Total variation techniques assume that image noise takes the form of sharp intensity differences (i.e., variation) between neighboring pixels in the image and seeks to remove this variation while producing a denoised image that is otherwise similar to the input image. By limiting the total variation reduction subject to the similarity between the input and output images, contrast due to the signal, such as molecule edges, is preserved. The total variation solution is formulated as a co-minimization problem as shown in the following equation:

$$u = \min_u \left[D(u, g) + \lambda V(u) \right], \tag{2}$$

where u is the denoised image, g is the noisy image, λ is a weight parameter, D is a difference function, and V is a total variation function. The weight parameter determines the preference for more strongly reducing pixel variation or preserving the original image. The initial total variation formulation (Rudin et al, 1992) used the L_2 norm as the difference function:

$$D(u, g) = \sum_i^n \sum_j^m (u_{i,j} - g_{i,j})^2, \tag{3}$$

where n and m are the number of pixel rows and columns, respectively, and i and j are the pixel indexes. For the variation function they used:

$$V(u) = \sum_i^{n-1} \sum_j^{m-1} \sqrt{|u_{i+1,j} - u_{i,j}|^2 + |u_{i,j+1} - u_{i,j}|^2}. \tag{4}$$

However, this initial formulation is difficult to solve because it is non-differentiable and an infinite-dimensional minimization problem (Duran et al., 2013). Several modifications have been proposed with adjustments to D, V, or the minimization method. But in all cases, the essence captured in equation (2) remains, where the total variation in the denoised image is reduced subject to fidelity to the original image. The Chambolle method uses a projection algorithm based on a dual formulation and is solved with gradient decent, while the Bregman method uses an operator splitting method (Duran et al., 2013). More details can be found in their respective papers (Chambolle, 2004; Osher et al., 2005).

Gaussian downsampling can be applied before denoising to resample the image to a proper sampling frequency where, according to the Nyquist–Shannon Sampling Theorem (Jerri, 1977), the pixel edge length should be 2.3–3 times smaller than the point-to-point resolution of the microscope. Prior to resampling the image to a larger pixel size, a Gaussian filter is applied with an appropriate kernel size to eliminate image frequencies higher than the Nyquist cutoff frequency of the resampled image. The effects of downsampling can be seen by comparing the third and fourth image columns in Figure 3. The results may appear to be insignificant upon visual inspection, but this is often an important first step. All denoising algorithms are applied to both the original and downsampled datasets for comparison.

Open-source code and a standalone application to apply Chambolle denoising to TEM images and video frames are available at https://github.com/JStuckner/smart_preprocess.

Results and Discussion

The C60@CNT1 and C60@CNT2 datasets, recorded at an electron dose per image ranging from 0.5×10^5 to 65×10^5 electrons/nm^2, were used to evaluate the performance of the denoising algorithms in increasing the SNR while preserving the signal. Qualitative results of applying select denoising algorithms are shown in Figure 3. The Chambolle denoising algorithm applied after downsampling increased the clarity of the C$_{60}$ molecules even at an extremely low EDR. A Chambolle weight of 0.5 seemed best at low electron doses and 0.2 seemed best at high electron doses. The bilateral filter also increased clarity compared with the original images, but not as well as the Chambolle algorithm. The Gaussian filter appeared to remove the most noise; however, it also blurred the edges of the molecules.

SNR Comparison

The SNR of each denoising algorithm was compared when applied with the optimum parameter value determined by the signal score metric discussed in the "Signal Preservation Comparison" section. Figure 4 shows the SNR of select methods when applied with the optimum parameter. Under these conditions, the Chambolle algorithm had the highest SNR, which was nearly six times better than the original image. The Chambolle algorithm produces the same SNR at an electron dose of 1.5×10^5 electrons/nm^2 as the unprocessed dataset had at an EDR of 65×10^5 electrons/nm^2. In this case, the Chambolle algorithm yielded an equivalent SNR as an unprocessed image taken at an EDR increased by nearly 1.5 orders of magnitude. The Bregman and Gaussian methods performed nearly as well as the Chambolle by the SNR metric, while the bilateral filter and low rank approximation methods performed moderately well. But even while improving the SNR about half as much as the
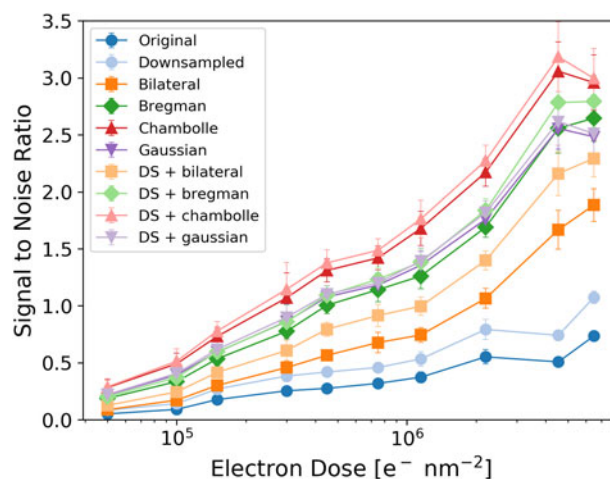

Fig. 4. SNR of original images and with several applied denoising algorithms at varying EDR. Denoising was applied with parameters set to remove the most noise while optimally preserving the signal.

Chambolle algorithm, the bilateral filter with downsampling still performed similarly at an electron dose of 4.5×10^5 electrons/nm^2 as the original dataset at an electron dose of 65×10^5 electrons/nm^2, which was over an order of magnitude reduced electron dose. Downsampling increased the SNR in all cases.

Signal Preservation Comparison

An important metric to judge the performance of a restoration method is how well it preserves edges, because precisely locating the edges is indispensable for size measurements in the SMART-EM molecular imaging. A good edge is characterized by a large pixel intensity gradient, which we calculated by taking the derivative of a pixel intensity profile across an edge. The pixel intensity profiles for select denoising algorithms applied to C60@CNT1 are shown in Figure 5. The line intensity profiles were taken from images denoised under the optimum parameters. The intensity profiles were plotted in blue lines and their derivatives, or intensity gradients, were plotted in dashed red lines. A smooth blue line indicates less noise. In all cases, downsampling reduced edge sharpness. This is because downsampling re-bins the signal into larger pixels, which are spatially less precise.

The mean, median, and Gaussian filters performed very well when measured by their SNR and signal strength. These algorithms removed a significant amount of noise, as seen by the smoothness of the blue curves. The major drawback of these algorithms is that they convolute the image, blurring points and edges. When denoising with these algorithms, the magnitude of the pixel intensity gradient across the edge was very small, as shown by the dashed red lines. Instead of having a sharp edge at a single location, the molecule gradually fades into the background, making it extremely difficult to perform high accuracy measurements of the molecule's size. The size and shape of the molecule was not preserved, and thus an important piece of the signal was corrupted by these denoising algorithms. The rough intensity profile of the low matrix approximation method along with small derivative magnitudes shows that this method was not effective in noise removal or edge preservation. The bilateral filter, having been designed specifically for edge preservation, performed the best in this regard. The derivative magnitudes of bilateral filtering were 2–3 times larger
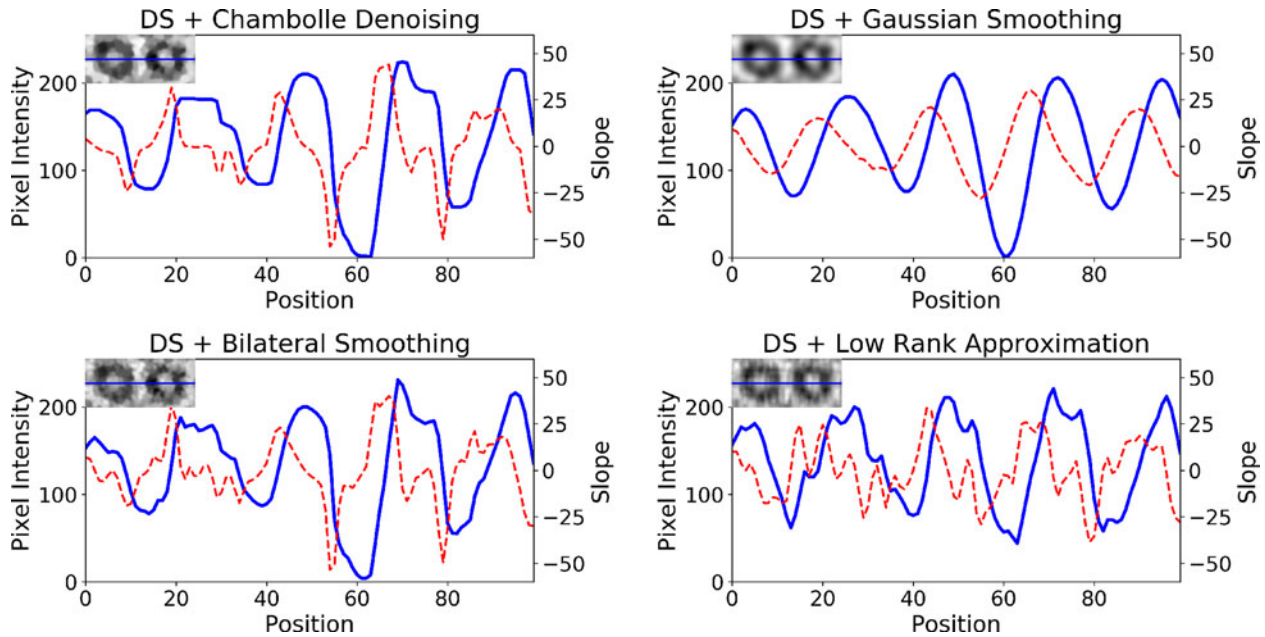
Fig. 5. Select line intensity profiles of a high EDR C60@CNT1 image after denoising with the optimum parameters. Solid blue curves indicate a pixel intensity profile. Dashed red curves indicate intensity gradient. The blue line on inset image indicates the location of the line intensity profile.

than for Gaussian smoothing. The Chambolle algorithm performed almost as well as the bilateral filter for edge preservation.

Chambolle Superimposition

The C60@CNT3 dataset, which was recorded at 1,600 fps at the maximum detector-safe EDR of 213×10^5 electrons/nm^2/s, was used to evaluate the performance of the denoising algorithms on high fps, low-dose images. Images and videos taken at such high framerates had an electron dose per frame of only 0.133×10^5 electrons/nm^2 and a very low SNR due to the short exposure time for each frame. The conventional method of increasing the SNR in this case is to superimpose neighboring frames by averaging each pixel value between frames. Such superimposition reduces the temporal resolution of the dataset by a factor equal to the number of frames superimposed. This type of superimposition gives the appearance of motion blur if an object has moved within the superimposed frames. It has been shown that Chambolle denoising can increase the SNR of each frame, but the algorithm can also be applied to superimposition. When using Chambolle superimposition, only pixels that are sufficiently similar between frames are combined. Through this approach motion blur was limited. Figure 6 compares the SNR after applying superimposition and Chambolle denoising to C60@CNT3. The red curve with small circle markers shows the SNR increase by applying superimposition by the typical averaging method to the dataset with no denoising. After superimposing 100 images, the SNR increased from about 0.1–0.6 with a 100-fold decrease in temporal resolution. Applying Chambolle denoising at a weight of 0.5 and no superimposition produced an SNR of 0.49 with no loss of temporal resolution. Using the average superimposition method, 68 frames were superimposed to generate an SNR of 0.5 representing a nearly two orders of magnitude improvement in temporal resolution at the same SNR by using the Chambolle algorithm. The solid curves with circle markers

show the effects of simultaneous Chambolle denoising and Chambolle superimposition limited to the specified number of frames for superimposition. The dotted curves with circle markers show the results of first superimposition by the average method followed by Chambolle denoising of the superimposed frames. While the SNR is increased by average superimposition followed by Chambolle denoising, Chambolle superimposition is recommended where possible because it will often preserve temporal changes that would be lost by average superimposition. For
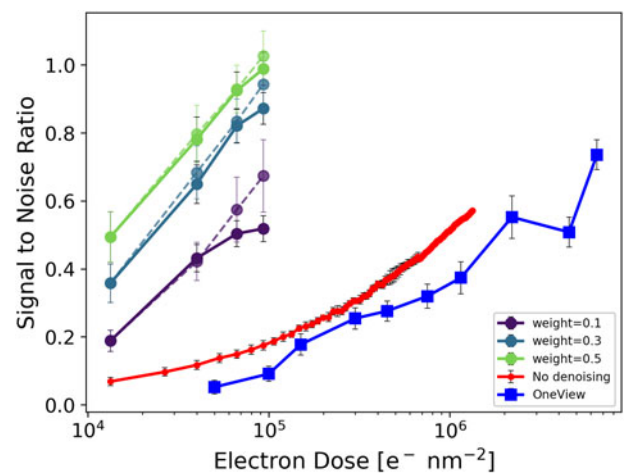
Fig. 6. Plot showing the improvement in the SNR of C60@CNT3 by frame superimposition compared with the effective exposure time. Solid lines with circle markers were downsampled and then simultaneously superimposed and denoised by the Chambolle algorithm. Dashed lines with circle markers were downsampled, then superimposed by the average superimposition method, and finally denoised by the Chambolle algorithm. The red line with small circle markers shows the improvement in the SNR by using average superimposition at the expense of temporal resolution. For comparison, the SNRs of OneView images from C60@CNT1 at varying electron doses are plotted in blue with square markers.

comparison, the SNR of C60@CNT1 is plotted in blue with square markers. Without superimposition, denoising using a Chambolle weight of 0.5 produced an SNR equivalent to the OneView images taken at a nearly equivalent EDR of 44×10^5 electrons/nm^2/s. Superimposing three frames at the same condition produced an SNR of almost 0.8, which is nearly equivalent to the SNR of the OneView images taken at the highest EDR the camera could accept of 130×10^5 electrons/nm^2/s. This represented a 50-fold increase in temporal resolution.

Application Examples

The effectiveness of the Chambolle superimposition and denoising method for improving the quality of high frame rate videos is obvious upon visual inspection as shown in Figure 1. Without denoising or superimposition (Fig. 1b), a single image taken for 0.625 ms appeared to be purely noise. It is only after superimposing 50 frames by a pixel averaging method (Fig. 1c), for an equivalent exposure time of 31.25 ms, that the molecules became clear enough to see (SNR = 0.20), but the details of faster dynamic phenomenon were lost. Figure 1d shows a single frame denoised by the Chambolle algorithm with an SNR of 0.15, which is comparable to that of the 50-frame averaging superimposition. Using simultaneous 3-frame Chambolle superposition and denoising further improved the image (Fig. 1e), with an exposure time of 1.875 ms. With 1/17 of the exposure time, we improved the SNR value from 0.20 in Figure 1c to 0.30 in Figure 1e. While SNR gain by superimposition or longer exposure time is large since the SNR gain is proportional to the square root of the exposure time, the time resolution is inversely proportional to the exposure time. Thus, SNR improvement by longer exposure time or superimposition of many images needs to be avoided in fast molecular imaging to prevent the loss of time resolution (Nakamura, 2017; Okada et al., 2017).

Figure 7 illustrates an application to the imaging of conformation changes of a molecule composed of a biotin end group connected to the tip of a carbon nanohorn (CNH) through a series of flexible organic chains. The biotin molecule (Fig. 7a) continuously
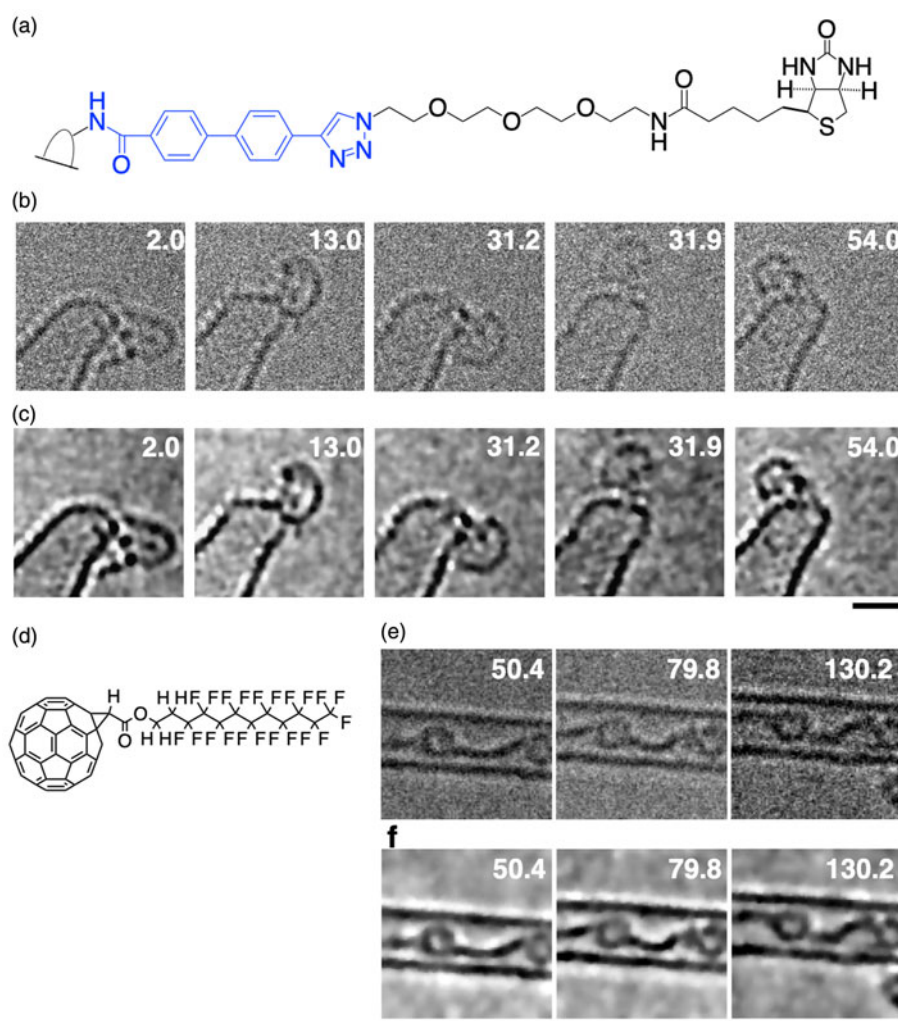
Fig. 7. Denoising of individual frames for single-molecular videos. (**a**) Structure of a biotin molecule attached on CNH. (**b**) Raw images of video frames of the biotin molecule (acceleration voltage 120 kV). A series of TEM images was obtained at intervals of 0.65 s with an exposure time of 0.4 s followed by a readout time of 0.25 s (non-irradiated). Raw images are adapted from Gorgoll et al. (2015). (**c**) Chambolle denoised images of (**b**). (**d**) Structure of a perfluoroalkyl fullerene. (**e**) Raw images of the perfluoroalkyl fullerenes in CNT (acceleration voltage 120 kV). A series of TEM images was obtained at intervals of 2.1 s with an exposure time of 0.5 s followed by a readout time of 1.6 s (non-irradiated). Original images are adapted from Harano et al., (2014). (**f**) Chambolle denoised images of (**e**). Numbers denote time in seconds after starting the video recording. Scale bar is 1 nm.

moves on a CNH during observation and gave different pictures in each frame of the video (Fig. 7b). The images taken on a charge-coupled device (CCD) sensor with a scintillator (Gatan Ultrascan 1000) contained significant noise (SNR of 0.19), and the conformation of the molecule was difficult to identify in some frames, (e.g. frame 31.9 s). The application of Chambolle denoising to the video made the molecular shape stand out from the background (SNR of 1.73), and the conformational changes in the molecule are distinctly recognizable (Fig. 7c). This result indicates that the noise reduction allowed us to obtain more dynamic information from SMART-EM video images recorded on conventional CCD devices. Figures 7e and 7f show another example comparing the original and denoised video frames of molecules in a CNT.

Conclusions

Aberration-corrected TEMs have significantly increased the spatial resolution of SMART-EM imaging, while high frame rate CMOS cameras have reduced the temporal resolution to the 1 ms range. However, low EDR images resulting from a high frame rate (or cryogenic TEM techniques) produce noisy images with a reduced SNR. We examined a variety of methods potentially applicable to denoising SMART-EM images and show that the Chambolle denoising algorithm produced the optimal balance of noise removal and signal preservation. At optimal parameter settings, it produced restored images with the highest SNR, a signal strength statistically equivalent to the best, and was second only to the bilateral filter in edge preservation. The bilateral filter is recommended when precise edge-to-edge measurements are required but produced only a moderate increase in the SNR and signal strength. Gaussian downsampling is an important first step if the data are sampled with a higher frequency than the ideal Nyquist sampling rate. It was shown to increase the signal strength and the SNR when applied before denoising, but it also decreases edge sharpness. The Chambolle algorithm was also effective in denoising high framerate video datasets. The method reduced the need for image superimposition, thus preserving the details of high-speed phenomenon (Shimizu et al., 2020). Additionally, when necessary, Chambolle superimposition allowed frames to be superimposed with less motion blur than pixel averaging superimposition. Our results show that the Chambolle total variation denoising algorithm can produce images with equal or better SNR while preserving morphology features when operating at an EDR reduced by more than an order of magnitude compared with the unprocessed images.

References

Buades A, Coll B & Morel J-M (2005). A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05),* vol. 2, pp. 60–65. San Diego, California, USA: IEEE.

Chambolle A (2004). An algorithm for total variation minimization and applications. *J Math Imaging Vis* **20**, 89–97.

Dalitz C, Pohle-Frohlich R & Michalk T (2015). Point spread functions and deconvolution of ultrasonic images. *IEEE Trans Ultrason Ferroelectr Freq Control* **62**, 531–544.

Dietrich O, Raya JG, Reeder SB, Reiser MF & Schoenberg SO (2007). Measurement of signal-to-noise ratios in MR images: Influence of multi-channel coils, parallel imaging, and reconstruction filters. *J Magn Reson Imaging* **26**, 375–385.

Duran J, Coll B & Sbert C (2013). Chambolle's projection algorithm for total variation denoising. *Image Process On Line* **3**, 311–331.

Gorgoll RM, Yücelen E, Kumamoto A, Shibata N, Harano K & Nakamura E (2015). Electron microscopic observation of selective excitation of conformational change of a single organic molecule. *J Am Chem Soc* **137**, 3474–3477.

Harano K, Homma T, Niimi Y, Koshino M, Suenaga K, Leibler L & Nakamura E (2012). Heterogeneous nucleation of organic crystals mediated by single-molecule templates. *Nat Mater* **11**, 877.

Harano K, Takenaga S, Okada S, Niimi Y, Yoshikai N, Isobe H, Suenaga K, Kataura H, Koshino M & Nakamura E (2014). Conformational analysis of single perfluoroalkyl chains by single-molecule real-time transmission electron microscopic imaging. *J Am Chem Soc* **136**, 466–473.

Hunt B (1971). A matrix theory proof of the discrete convolution theorem. *IEEE Trans Audio Electroacoust* **19**, 285–288.

Jerri AJ (1977). The Shannon sampling theorem—Its various extensions and applications: A tutorial review. *Proc IEEE* **65**, 1565–1596.

Kushwaha HS, Tanwar S, Rathore KS & Srivastava S (2012). De-noising filters for TEM (transmission electron microscopy) image of nanomaterials. In *2012 Second International Conference on Advanced Computing & Communication Technologies*, pp. 276–281. Los Angeles, California, USA: IEEE.

Liao H-G, Zherebetskyy D, Xin H, Czarnik C, Ercius P, Elmlund H, Pan M, Wang L-W & Zheng H (2014). Facet development during platinum nanocube growth. *Science* **345**, 916–919.

Lin Z, Chen M & Ma Y (2010). The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv preprint arXiv:1009.5055.*

Liu G, Yousefi S, Zhi Z & Wang RK (2011). Automatic estimation of point-spread-function for deconvoluting out-of-focus optical coherence tomographic images using information entropy-based approach. *Opt Express* **19**, 18135–18148.

Munezawa T, Goto T, Hirano S & Phung SL (2019). A study on moving image noise removal using 3D and time-domain total variation regularization method. In *International Workshop on Advanced Image Technology (IWAIT) 2019,* vol. 11049, p. 1104913. Singapore: International Society for Optics and Photonics.

Nakamura E (2017). Atomic-resolution transmission electron microscopic movies for study of organic molecules, assemblies, and reactions: The first 10 years of development. *Acc. Chem. Res.* **50**, 1281–1292.

Okada S, Kowashi S, Schweighauser L, Yamanouchi K, Harano K & Nakamura E (2017). Direct microscopic analysis of individual C60 dimerization events: Kinetics and mechanisms. *J Am Chem Soc* **139**, 18281–18287.

Orieux F, Giovannelli J-F & Rodet T (2010). Bayesian estimation of regularization and point spread function parameters for Wiener–Hunt deconvolution. *JOSA A* **27**, 1593–1607.

Osher S, Burger M, Goldfarb D, Xu J & Yin W (2005). An iterative regularization method for total variation-based image restoration. *Multiscale Model Simul* **4**, 460–489.

Otsu N (1979). A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* **9**, 62–66.

Price-Whelan AM, Sipőcz BM, Günther HM, Lim PL, Crawford SM, Conseil S, Shupe DL, Craig MW, Dencheva N & Ginsburg A (2018). The Astropy Project: Building an inclusive, open-science project and status of the v2.0 core package. *arXiv preprint arXiv:1801.02634.*

Roels J, Aelterman J, De Vylder J, Luong H, Saeys Y & Philips W (2016). Bayesian deconvolution of scanning electron microscopy images using point-spread function estimation and non-local regularization. In *38th*

Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 443–447. Orlando, Florida, USA: Institute of Electrical and Electronics Engineers (IEEE).

Rudin LI, Osher S & Fatemi E (1992). Nonlinear total variation based noise removal algorithms. *Physica D* **60**, 259–268.

Shimizu T, Lungerich D, Stuckner J, Murayama M, Harano K & Nakamura E (2020). Real-time video imaging of mechanical motions of a single molecular shuttle with sub-millisecond sub-angstrom precision. *Bull Chem Soc Jpn*. doi:10.1246/bcsj.20200134.

Sloan J, Dunin-Borkowski RE, Hutchison JL, Coleman KS, Williams VC, Claridge JB, York APE, Xu C, Bailey SR & Brown G (2000). The size distribution, imaging and obstructing properties of C60 and higher fullerenes formed within arc-grown single walled carbon nanotubes. *Chem Phys Lett* **316**, 191–198.

Tomasi C & Manduchi R (1998). Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*. Bombay, India: IEEE.

Wold S, Esbensen K & Geladi P (1987). Principal component analysis. *Chemom Intell Lab Syst* **2**, 37–52.

Xing J, Schweighauser L, Okada S, Harano K & Nakamura E (2019). Atomistic structures and dynamics of prenucleation clusters in MOF-2 and MOF-5 syntheses. *Nat Commun* **10**, 1–9.

Yuen HK, Princen J, Illingworth J & Kittler J (1990). Comparative study of Hough transform methods for circle finding. *Image Vision Comput* **8**, 71–77.

Zhu Y, Ciston J, Zheng B, Miao X, Czarnik C, Pan Y, Sougrat R, Lai Z, Hsiung C-E & Yao K (2017). Unravelling surface and interfacial structures of a metal–organic framework by transmission electron microscopy. *Nat Mater* **16**, 532.