Supporting Text

Tomogram Denoising. Direct interpretation of the intracellular mature vaccinia virus (IMV) tomograms was difficult because of the fine-grain noise present in the reconstructions. Therefore, denoising was necessary to discern and interpret structural features in the tomograms. Two different denoising methods were combined: Gaussian filtering and anisotropic nonlinear diffusion (AND). The combination of both methods allowed masking-out of the background while preserving the structural features of the IMVs at good resolution. This procedure did not involve significant modification of the signal in the tomogram; hence, quantitative postprocessing was possible (1).

Gaussian filtering is a linear denoising method (2), which essentially means that any structural feature in the tomogram is low-pass filtered by the same filter. Consequently, Gaussian filtering does not preserve edges in the tomogram, and features are thus blurred. The strength of the Gaussian filtering is defined by the standard deviation of the Gaussian function, σ_g. Different values of σ_g were tested over IMV tomograms, and $\sigma_g = 1$ was the most appropriate value in terms of noise reduction vs. blurring of structural features. The tomograms resulting from the Gaussian filtering kept sharp structural details, being much cleaner than the original ones (see Fig. 1*B*). Nevertheless, a more powerful denoising method was still necessary to remove the considerable residual background.

AND is currently the most powerful method applied in cryo-electron tomography (3, 4). This method is capable of reducing noise while preserving structural features. AND has two main diffusion modes: EED (edge-enhancing diffusion) and CED (coherence enhancement diffusion). The former allows noise reduction with edge enhancement. The latter allows improvement of curvilinear and planar structures. A combination of both methods proved to be an efficient approach for visualizing tomograms (3). In that hybrid approach, a threshold over the local relation between structure and noise is computed from a subvolume of the tomogram containing only noise. Then, all of the voxels whose local structure is higher than that threshold are considered as structural features and hence processed as CED; otherwise, EED is applied.

In this work, the AND strategy described in ref. 4 was applied. This approach tightly combines both EED and CED modes, as well. The important novelty is the use of 2D CED over plane or surface-like structures. This type of CED allows significant enhancement of those features, especially in single-tilt axis tomography, as quantified in that work. Such a diffusion mode has turned out to be especially useful to enhance membranes in IMV tomograms. Moreover, the AND approach used in this work also included Gaussian filtering over all of the voxels whose density measured in its neighborhood was lower than the density measured for the subvolume containing noise. This improvement allowed further smoothing of the background.

Before the application of AND over the tomograms, they were rescaled so that the density was in the range [0,255], where 0 was background and 255 the highest density. Denoising was then performed by using five iterations of the AND according to Fernandez and Li (4), equivalent to 20 iterations of the approach in ref. 3. The values of

the parameters for CED and EED modes were $C = 4$ and $K = 2$, respectively. The local scales used in the AND approach were $\sigma_o = 1$ for initial filter of the input volume, and $\sigma = 2$ for the averaging of the structure tensor used in the computation of CED diffusion.

This AND approach yielded much cleaner tomograms than did the Gaussian filter. The background was almost completely smoothed out, and important structural features were significantly highlighted (see Fig. 1*C*). The results provided by AND were suitable for 3D visualization, with either isosurface or volume rendering. Nevertheless, AND methods have the drawback that the signal in the tomogram is modified (even some artifacts might appear), so interpretation and quantitative postprocessing should be carried out with caution. This disadvantage was already outlined in ref. 1. For that reason, a combination of Gaussian and AND denoising was devised to take advantage of their best properties, as described in the following.

Segmentation of IMVs. The combination of tomograms denoised by Gaussian filtering and by AND was considered more feasible for visualization and interpretation of fine structural details. AND smoothes out the background and highlights features, whereas Gaussian filtering (with a relatively small σ_g) preserves resolution in the important structural features. Consequently, the tomograms denoised with AND were used to define masks to be applied over the ones resulting from Gaussian filtering. The masks allowed the preservation of the areas containing important features, with the background and unimportant structures cancelled out. The tomograms resulting from this combination then contained the important features of the IMVs at sufficient resolution to discern important structural details, without any significant modification of the signal (see Fig. 1*D*).

The creation of a mask from the AND result was carried out with AMIRA (TGS Europe, Merignac, France) according to the following steps:

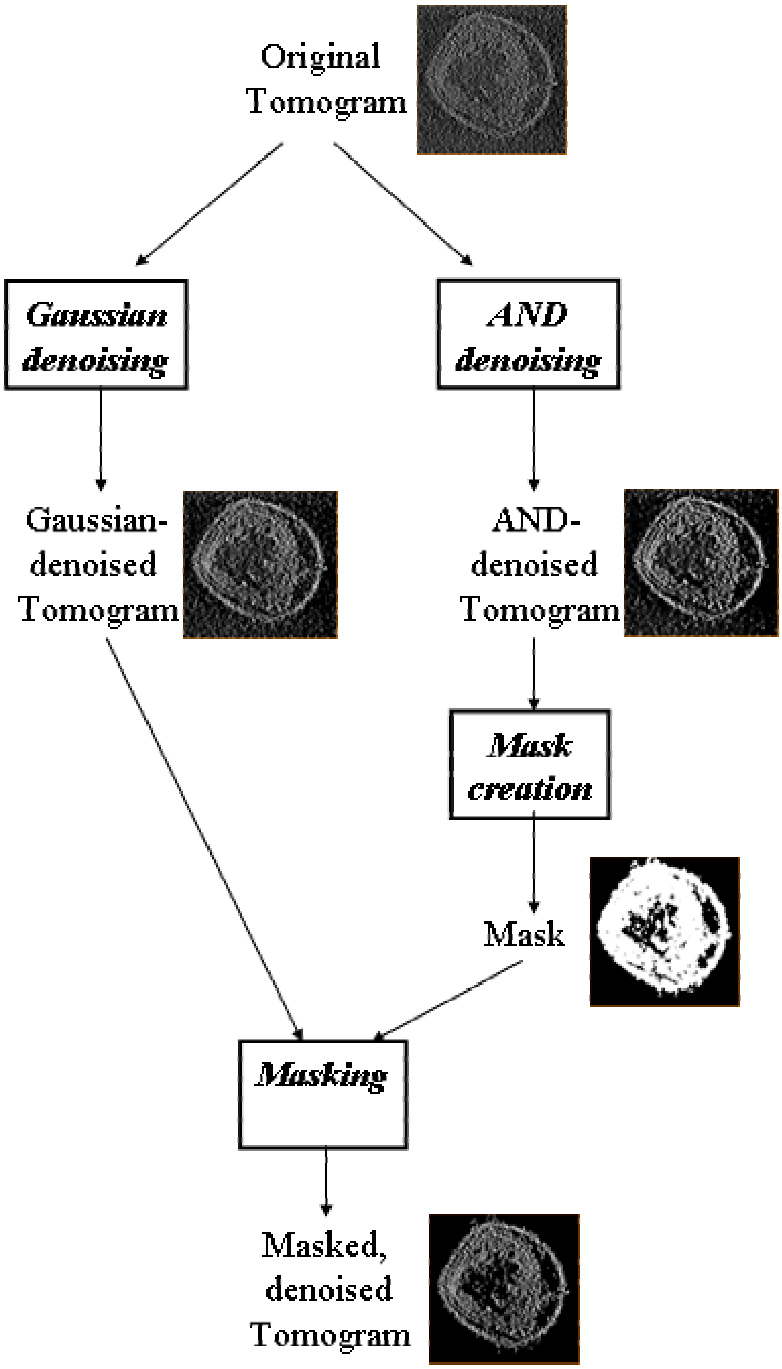
1. Threshold-based binary segmentation of the tomogram resulting from AND denoising. In this step a threshold over the density value (see below a method for objective determination of a density threshold) was defined. Afterward, all of the voxels with density greater than the threshold were considered as part of the mask; otherwise, they were cancelled out. This step also filtered out unconnected small areas spread throughout the background.

2. Interactive monitoring of the binary segmentation. Some noisy areas connected around the extremes of the tomogram, mainly related to the effect of the missing wedge, could be removed semiautomatically.

3. The resulting segmentation was further processed by using a morphological dilation operation (2). This operation from the field of computer vision was used here to expand the segmentation result, preserving its global structure. The aim was the creation of a global connected mask that loosely fit the IMV, more suitable for masking the Gaussian tomogram.

4. The mask just created was then applied over the tomogram resulting from the Gaussian filtering.

Scheme of the procedure for denoising and segmenting IMVs



Optimal thresholding. Objective determination of a density threshold was needed in several stages of the procedure described here. First, the segmentation of IMVs in the tomograms required a threshold for the creation of the mask. Second, a threshold is also required in 3D visualization with an isosurface, because the smooth 3D surface comprises all of the voxels with density lower than the threshold. In 3D electron microscopy of single particles or icosahedral viruses, this threshold is normally found out based on the molecular weight of the complex and the protein density (see, for instance, refs. 5–8). However, such a procedure cannot be applied in electron tomography of complex biological specimens.

In this work, the thresholds for segmentation and visualization were objectively computed by means of optimal thresholding (2), a method derived from the field of computer vision. In essence, the method assumes that the volume is composed by a background and a foreground, and that the density histogram of the volume can then be modeled as the weighted sum of the two probability densities that approach normal distributions (the histograms of the background and foreground). Those assumptions are general enough to be valid under different brightness and contrast conditions and applicable to a wide spectrum of image-processing applications, in particular to cryo-electron tomography. Here, all of the structural components of IMVs made up the foreground.

Under those assumptions, the threshold that allows optimal segmentation is given by the density level corresponding to the minimum probability between the maxima of the two normal distributions (foreground and background). However, determination of such an optimal threshold is far from trivial. The difficulty stems from the fact that histograms are not, in general, bimodal (i.e., two regions are clearly recognizable), and, therefore, the simplest strategy consisting of searching for a minimum between two maxima in the histogram does not work. In practice, the proper strategy searches for the density threshold that maximizes the variance between foreground and background. There are a number of alternative approaches for that goal (2).

In this work, an iterative method inspired by the one shown in ref. 2 was developed for determination of optimal thresholds. In essence, the algorithm iterates by redefining the background and foreground areas in the tomogram until a density threshold is found in the exact middle of the mean densities of the background and foreground. The algorithm is simple and requires a few iterations. In the following, the steps of the algorithm are presented. Let μ_F denote the mean density in the foreground, and μ_B denote the mean density in the background. Let $T^{(t)}$ be threshold computed at the iteration t. Let V denote the tomogram, with $V(i,j,k)$ being the voxel with indices (i,j,k).

1. Initial values for μ_B, μ_F, and $T^{(o)}$:
$\mu_B = 0$; μ_F = mean density of the tomogram; $T^{(o)} = \mu_F$.

2. At step t, segment the tomogram:
If the voxel $V(i,j,k)$ has density greater than or equal to $T^{(t)}$, then the voxel is classified as foreground; otherwise, the voxel is classified as background.

3. At step t, update the mean densities μ_B and μ_F :
μ_B = mean density of the voxels segmented as background; μ_F = mean density of the voxels segmented as foreground.

4. At step t, update the foreground vs. background threshold:
$T^{(t+1)} = (\mu_B + \mu_F) / 2$

5. Iterate?:
If $T^{(t+1)} = T^{(t)}$, then stop, and $T^{(t+1)}$ is the optimal threshold. Otherwise, return to step 2 and iterate.

This algorithm was applied here for determination of the optimal thresholds in several stages: creation of masks in segmentation of IMVs and 3D visualization of tomograms and subareas with isosurface. Furthermore, this technique proved to be very useful to confirm the existence of the pores and measure their aperture.

1. Plitzko, J. M., Frangakis, A. S., Nickell, S., Förster, F., Gross, A. & Baumeister, W. (2002) *Trends Biotech.* **20,** S40-S44.

2. Sonka, M., Hlavac, V. & Boyle, R. (1999) *Image Processing, Analysis, and Machine Vision* (PWS, Pacific Grove, CA).

3. Frangakis, A. S. & Hegerl, R. (2001) *J. Struct. Biol.* **135,** 239-250.

4. Fernandez, J. J. & Li, S. (2003) *J. Struct. Biol.* **144,** 152-161.

5. Paredes, A. M., Heidner, H., Thuman-Commike, P., Venkataram-Prasad, B. V., Johnston, R. E., & Chiu, W. (1998) *J. Virol.* **72,** 1534-1541.

6. Walz, J., Erdmann, A., Kania, M., Typke, D., Koster, A. J. & Baumeister, W. (1998) *J. Struct. Biol.* **121,** 19-29.

7. Barcena, M., Ruiz, T., Donate, L. E., Brown, S. E., Dixon, N. E., Radermacher, M. & Carazo, J. M. *(2001) EMBO J.* **20,** 1462-1468.

8. Heymann, J. B., Cheng, N., Newcomb, W. W., Trus, B. L., Brown, J. C. & Steven, A. C. (2003) *Nat. Struct. Biol.* **10,** 334-341.