A ridge-based framework for segmentation of 3D electron microscopy datasets

Antonio Martinez-Sanchez[a], Inmaculada Garcia[b], Jose-Jesus Fernandez[c,*]

[a]*Supercomputing and algorithms group. Associated Unit CSIC-UAL, University of Almeria, 04120 Almeria, Spain*
[b]*Supercomputing and algorithms group. Dept. Computer architecture. University of Malaga, 29080 Malaga, Spain*
[c]*National Centre for Biotechnology. National Reseach Council (CSIC). Campus UAM. Darwin 3. Cantoblanco. 28049 Madrid. Spain.*

[*]Corresponding author:
National Centre for Biotechnology.
National Reseach Council (CSIC).
Campus UAM. C/ Darwin 3. Cantoblanco.
28049 Madrid. Spain.
Tel: +34 91 585 4619. Fax: +34 91 585 4506
Email: jj.fernandez@csic.es

Preprint submitted to J. Struct. Biol. *September 25, 2012*

Abstract

Three-dimensional (3D) electron microscopy (EM) has become a major player in structural cell biology as it enables the analysis of subcellular architecture at an unprecedented level of detail. Interpretation of the resulting 3D volumes strongly depends on segmentation, which consists in decomposing the volume into their structural components. The computational approaches proposed so far have not turned out to be of general applicability. Thus, manual segmentation still remains a prevalent method. Here, a new computational framework for segmentation of 3D EM datasets is introduced. It relies on detection and characterization of ridges (i.e. local maxima). The detected ridges are modelled as asymmetric Gaussian functions whose parameters constitute ridge descriptors. This local information is then used to cluster the ridges, which leads to the ultimate segmentation. In this work we focus on membranes and locally planar structures in general. The performance of the framework is illustrated with its application to a number of complex 3D datasets and a quantitative analysis.

Key words

Segmentation , Image Processing , Electron Tomography , Serial section 3D Reconstruction , Serial Blockface , Membrane

1. Introduction

Electron tomography (ET) is an essential technique in structural cell biology for visualization of the supramolecular organization of the cellular environment *in situ* at a resolution of a few nanometres (Lucic et al., 2005). It relies on the acquisition of a set of electron microscope (EM) images from the specimen at different views, which are subsequently combined to yield the three-dimensional volume (also known as tomogram) (Fernandez, 2012). Interpretation of the tomogram is complicated due to a number of factors (Volkmann, 2010; Fernandez, 2012). As a consequence, there are a number of post-processing stages devoted to tomogram analysis (e.g. noise reduction). Segmentation is a stage of utmost importance, which aims at decomposing the tomogram into its structural components by identifying the sets of voxels that constitute them. This task is severely hampered by the low signal-to-noise ratio (SNR), the inherent biological complexity and artefacts deriving from the ET imaging conditions. Though numerous automatic or semi-automatic approaches have been proposed (Volkmann, 2010; Fernandez, 2012), none has stood out as a general applicable method yet. In the last few years there is a trend towards methods focused on detection of specific features, such as membranes (Martinez-Sanchez et al., 2011), actin filaments (Rigort et al., 2012) or microtubules (Weber et al., 2012), but the general acceptance has yet to be confirmed. Thus, manual segmentation still remains a prevalent method.

Electron tomography is only applicable to relatively thin samples (up to 0.5–1μm) (Lucic et al., 2005). If pretty thick specimens (several μm thick) are to be studied, there exist other very well-known electron microscopy modalities commonly used. They rely on serial sectioning (SS) of the sample, where the sections are imaged after (classical SS) or before (serial blockface) being cut (Harris, 1999;

Denk and Horstmann, 2004). SS can also be combined with ET (Soto et al., 1994), which yields better resolution along the electron beam direction. In all these SS techniques, segmentation is also a central step to facilitate interpretation of the resulting 3D datasets. Manual segmentation is the standard method, though there is significant progress towards automation (Jain et al., 2010). The methods devised for ET are also directly applicable to SS as long as the interval between the serial sections is short enough.

In this work we introduce a new framework for segmentation of 3D EM datasets in structural cell biology. It relies on detection and characterization of ridges (i.e. local maxima). The local information associated to ridges is extracted and analyzed to characterize and classify the planar structures, which leads to the segmented tomogram. The local properties of ridges are obtained by fitting a Gaussian model. The resulting feature vectors contain attributes that turn out to be visually perceptible by humans, which facilitates further manual supervision or annotation. In this work we focus on membranes and structures that at local scale can be considered planar ridges. Membranes are natural boundaries that encompass compartments within biological specimens. So their detection would involve a good step towards full segmentation of datasets.

2. Ridge-based segmentation framework

Florack et al. (1992) defined the concept of scale and determined how to use differential geometry tools (e.g. gradient, curvature) to describe images. They also introduced the concept of isophote, which allows description of objects as n-manifolds (also known as n-dimensional manifolds). Edges and ridges can be described as 1-manifolds (curves in the plane) in images and 2-manifolds (surfaces in the space) in volumes. Nowadays, these tools are widely used in image

processing (e.g. Arbelaez et al., 2011). Ridges in particular are very useful in Computer Vision, for example to detect and analyse tube-like and plane-like features (e.g. Sato et al., 2003; Kirbas and Francis, 2004; Bauer et al., 2010).

2.1. 1D ridge analysis

This section shows how ridges can be locally described in a one-dimensional function or 1D image. In that case, ridges are 0-manifolds (discrete points in $\mathbb{R}$).

This work assumes that the information of interest is represented by positive curvature structures. Detection of local maxima (i.e. ridges) in a 1D function f allows isolation of the set $\mathcal{D} = \left\{x_j : j = 1, ..., K_{\mathcal{D}}\right\} \in X$, with $X \in \mathbb{R}$ being the domain of f and where x_j is a 0-manifold. The ridges in $\mathcal{D}$ constitute the *details* of interest. A simple Gaussian-based model can be used to locally describe every detail (Figs. 1 and 2). This model is defined as $f_m : \mathbb{R} \rightarrow \mathbb{R}^+$:

$$f_m(x) = o_j + s_j \exp\left(-\frac{(x - x_j)^2}{2t_j^2}\right) \tag{1}$$

where x_j is the location of a local maximum of f, i.e. an element of $\mathcal{D}$, o_j (*offset*) represents its intensity level offset, s_j is the local significance (*sharpness*) and the size is expressed by t_j (*thickness*). The function f_m is a local approximation of f. We must now define what local is for every $x_j \in \mathcal{D}$. Let be $\mathcal{E}_j$ an interval around x_j. The next three premises have been used to establish local sets:

- $x_j \in \mathcal{E}_j \Rightarrow x_j \notin \mathcal{E}_i, \forall i \neq j$ where $i, j \in \{1, 2, \cdots, K_{\mathcal{D}}\}$
- $\mathcal{E}_j \cap \mathcal{E}_i = \emptyset, \forall i \neq j$ where $i, j \in \{1, 2, \cdots, K_{\mathcal{D}}\}$.
- $\bigcup_1^{K_{\mathcal{D}}} \mathcal{E}_j = X - \left\{\bigcup_1^{Ks} \mathcal{S}_i\right\}$, where $\mathcal{S}_i$ is one of the $K_{\mathcal{S}}$ *step* regions that are pure edges, which cannot be approximated by f_m and are not considered in this work.

At each interval $\mathcal{E}_j$ delimited by *antidetails* the function f can be approximated by f_m. In this work the set of antidetails is composed by local minima or saddle points surrounding ridges (see black arrows in Fig. 1). Except for the special cases of $j = 1$ and $j = K_{\mathcal{D}}$, a detail x_j is delimited by the antidetails x_j^l and x_j^r. Consequently $x_j \in \mathcal{E}_j = \left[x_j^l, x_j^r\right]$. For the special cases of $j = 1$ and $j = K_{\mathcal{D}}$, $x_1 \in \mathcal{E}_1 = \left[\inf\{X\}, x_1^r\right]$ and $x_{K_{\mathcal{D}}} \in \mathcal{E}_{K_{\mathcal{D}}} = \left[x_{K_{\mathcal{D}}}^l, \sup\{X\}\right]$, where $\inf\{\cdot\}$ and $\sup\{\cdot\}$ are the infimum and superior of a set, respectively. Using antidetails, the step regions can be isolated as well (Fig. 1). A new parameter can then be introduced, *resolution*, given by $r_j = \varnothing\{\mathcal{E}_j\}$, where $\varnothing\{\cdot\}$ is the diameter of a set.

Fig. 2 shows that data on the left and the right of a detail are not necessarily symmetrical. However, the Gaussian model in Eq. (1) is symmetrical. To accomodate this asymmetry, the model is then divided in two parts to end up with the following piecewise function

$$f_{mp} = \begin{cases} f_l = f_m(x^l) & \forall x^l \in [x_j^l, x_j] = \mathcal{E}_j^l \\ f_r = f_m(x^r) & \forall x^r \in [x_j, x_j^r] = \mathcal{E}_j^r \end{cases} \tag{2}$$

for every detail x_j. This definition of $\mathcal{E}_j = \mathcal{E}_j^l \cup \mathcal{E}_j^r$ makes each parameter require a pair of values ($\bar{o}_j, \bar{s}_j, \bar{t}_j, \bar{r}_j$, with $\bar{o}_j = \{o_j^l, o_j^r\}$, etc.) that are obtained from the information in $\mathcal{E}_j^l$ and $\mathcal{E}_j^r$.

Once every detail and its local neighbourhood has been detected, the model parameters are adjusted to real data f. Model fitting can be expressed as an overdetermined non-linear equations system:

$$\underset{\bar{o}_j, \bar{s}_j, \bar{t}_j, \bar{r}_j}{\arg} \left[\min \left\{ \sum_{\forall x_k \in \mathcal{E}_j} \left(f(x_k) - f_{mp}(x_k; \bar{o}, \bar{s}, \bar{t}, \bar{r}) \right)^2 \right\} \right] \tag{3}$$

An optimization algorithm can solve this problem (in particular, Newton-based methods have been used here) and the result consists of a set of parameter pairs

$\overline{o}_j, \overline{s}_j, \overline{t}_j, \overline{r}_j$ of f_{mp} that minimises the quadratic error (Figs. 1(right), 2). We then characterize ridges with an eight-dimensional vector composed of the maximum value $p_j = \max\{p_j^l, p_j^r\}$ and the asymmetry (Eq. 4) of each parameter $p_j \in \left\{o_j, s_j, t_j, r_j\right\}$.

$$a_j^p = \frac{|p_j^l - p_j^r|}{|p_j^l + p_j^r|} \tag{4}$$

2.2. Extension to three dimensions

For data dimensions $d > 1$ (i.e. 2D for images; 3D for tomograms), an extension of a local maximum can be a ridge described as a $(d-1)$-manifold (*plane ridge*) (Eberly, 1994) (other types of ridges described by others manifolds are not considered in this work). As a result, a point $\mathbf{x}_j \in \mathbb{R}^d$ in a ridge (detail) can be defined by its second order differential structure and the Hessian matrix (Martinez-Sanchez et al., 2011). In the 3D case in particular, an eigenanalysis of the Hessian matrix provides three orthogonal eigenvectors $\mathbf{v_i}$ and the corresponding eigenvalues λ_i, with $i = 1, 2, 3$. For planar ridges, the first eigenvector $\mathbf{v_1}$ is the one whose eigenvalue λ_1 exhibits the largest absolute value and points to the direction of the maximum curvature (second derivative) while the other eigenvalues are null. This kind of ridges can therefore be characterized just analysing the 3D function f in the direction of $\mathbf{v_1}$. It means that for the analysis we are interested in, the function $f : \mathbb{R}^3 \rightarrow \mathbb{R}^+$ can actually be turned into a one-dimensional function $\overline{f} : \mathbb{R} \rightarrow \mathbb{R}^+$ in the neighbourhood of a detail x_j. Consequently, the theory developed in the previous sections can be applied, so $\overline{f}$ is approximated by f_{mp}, whose parameters will describe this detail belonging to a 2-manifold in $\mathbb{R}^3$ (i.e. a surface in a 3D space)(see Suppl. Fig. S1).

These ridges can be identified by means of the detector of locally planar structures introduced by Martinez-Sanchez et al. (2011), M. This local detector was de-

fined as the ratio between the squared second-order and squared first-order derivatives, which in brief can be expressed as:

$$M = \begin{cases} \frac{(|\lambda_1| - \sqrt{\lambda_2 \lambda_3})^2}{|\nabla f|^2} & \lambda_1 < 0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $|\nabla f|$ denotes the gradient. In this work, we also include a threshold t_M over M to discard spurious ridges. Specifically, the ridge detector used in this work is given by:

$$\begin{cases} \Big(f(\mathbf{x}_j) > f(\mathbf{x}_j - \alpha \mathbf{v_1})\Big) \text{ and } \Big(f(\mathbf{x}_j) > f(\mathbf{x}_j + \alpha \mathbf{v_1})\Big) \\ M(\mathbf{x}_j) \geq t_M \end{cases} \tag{6}$$

where the first condition (with α being a small number) denotes that only the local maximum associated to the ridge is to be extracted (Suppl. Fig. S1).

Detection of ridges and the antidetails encompassing the associated local set is sensitive to noise, so a pre-processing step intended to reduce it is required. As thoroughly discussed in our previous work (Martinez-Sanchez et al., 2011), scale-space can allow focusing on structural features at a certain scale σ whereas smaller features and noise are filtered or smeared out. Scale-space is achieved by convolution with a Gaussian function. More aggressive filtering methods are also applicable, such as those based on nonlinear techniques (Fernandez, 2009) or anisotropic nonlinear diffusion (Fernandez and Li, 2003). However, even in those cases scale-space should be applied for a better fit with the Gaussian ridge model in the previous section.

2.3. Ridge classification and tomogram segmentation

The parameters characterizing the ridges constitute eight-dimensional feature vectors that can be used for clustering the ridges. In this work, we have used

several techniques for this. First, thresholding applied to the different parameters turns out to be a simple but effective way to cluster them. The fact that the ridge parameters correspond to features visually perceptible for humans (e.g. offset, thickness, sharpness, resolution) makes them appropriate for interactive user clustering or for further supervision. Second, these feature vectors are suited for unsupervised classification methods as well. In particular, here we have tested the well-known K-means and Self-Organizing Maps (SOMs) (Marslan, 2009), which are methods commonly used in the field of 3D-EM for different clustering purposes (Frank, 2006). Briefly, SOMs (Marabini and Carazo, 1994; Fernandez and Carazo, 1996; Kohonen, 2001) are neural networks that are trained by unsupervised learning. Given an input vector, the neurons of the SOM compete by means of mutual lateral interactions and only one is thus activated. The winner neuron and its neighbours are then tuned to the input vector. As the algorithm iteratively proceeds with all input vectors, the locations of the neural activity tend to become ordered. The trained SOM approximates the probability density function of the input data. Each neuron of the SOM may be considered as a class, and it contains a so-called code vector that is the representative of the input data mapped onto that neuron. However, the fact that the SOM reflects the probability density function results in the ability to better discriminate among similar input vectors that occur more frequently. As a consequence, it is common a post-processing stage consisting of clustering the SOM neurons. The goal is to yield the definite set of classes that are as different as possible from each other (Fernandez and Carazo, 1996; Yu and Frangakis, 2011).

The clustering partitions the set of ridges $\mathcal{D} = \{\mathbf{x}_j\}$ into K_C classes or regions $C = \{C_k : k = 1, \ldots, K_C\}$. This classification is, however, limited to the discrete

set of points making up the 2-manifolds (i.e. surfaces). It is thus necessary to extend it to the voxels of the tomogram. To this end, a region growing procedure is applied whereby the ridges are considered initial 'seed' points and neighbour voxels are progressively added to the regions. In this growing process, every voxel under consideration is assigned to the region of the nearest ridge $\mathbf{x}_j$. On the other hand, those voxels unlikely belonging to locally planar structures are ignored (i.e. those with $M < t_M$, see Eqs. 5 and 6). This process can be expressed as:

$$\mathbf{x}_j \in C_k \;\Rightarrow\; \mathbf{x} \in C_k \;\; \forall \mathbf{x} \in X_M : \; j = \arg\min_j \{\|\mathbf{x} - \mathbf{x}_j\|\} \tag{7}$$

with $X_M = \{\mathbf{x} \in X : M(\mathbf{x}) \geq t_M\}$ and X denoting the whole set of voxels of the tomogram.

Finally, a region merging procedure is applied with the aim of eliminating spurious segmented areas. An analysis based on the size of the segmented regions is performed first. Then, any region with a volume size lower than a given threshold t_v is merged to the neighbour region that contains more voxels apposed to it. For illustration purposes, Suppl. Fig. S2 shows an example of the region growing and merging procedures.

The ridge-based framework for 3D segmentation is thus summarized as:

1. Pre-processing: noise reduction and scale-space.
2. Ridge detection according to the local plane detector M (Eqs. 5 and 6).
3. Ridge characterization:
    - Extraction of the 1D ridge and the associated local set along the direction given by the first eigenvector of the Hessian matrix $\mathbf{v_1}$.
    - Parameterization based on Gaussian-model fitting: offset, sharpness, thickness, resolution and their corresponding asymmetries.

4. Ridge classification.
5. Extension of ridge clusters to tomogram segmentation.
6. Segmentation refinement by region merging.

3. Validation

Validation of segmentation algorithms is a difficult topic in the field mainly because of the lack of reliable and unquestionable 'ground truths' (Volkmann, 2010). For this reason, the quantitative assessment of the algorithm has been based on a synthetic phantom resembling representative experimental datasets. This allows us to directly measure the sensitivity, i.e. fraction of true positives (TPF, points correctly segmented and classified) and the specificity, i.e. true negatives (TNF, points that have been correctly left out of the segmented objects), as commonly defined in the field (Garduno et al., 2008; Martinez-Sanchez et al., 2011; Langlois and Frank, 2011). As complementary metrics, we have also used those previously proposed for segmentation of relatively thin structures (Martinez-Sanchez et al., 2011). They were specially focused on analysis of outlined shapes rather than voxel-based comparisons:

- **Centroid**: centre of mass.
- **Bounding box**: *Width* and *Height* of the smaller rectangular box containing the shape.
- **Axes**: Length of the *Major* and *Minor axes* of the ellipse with the same normalized second central moment as the shape.

These metrics are calculated planewise along the X, Y and Z axes and the final figures are obtained by weighted averaging, as described previously (Martinez-

Sanchez et al., 2011).

There are not many existing methods to detect and classify ridge structures. A recent, related tool (TomoSegMem) for segmentation of membranes (i.e. locally planar structures) based on differential structure (Martinez-Sanchez et al., 2011) has also been tested for comparison. That tool manages to detect membranous structures, but it has limited capabilities to classify and distinguish them.

4. Results

The ridge-based segmentation framework has been tested on different 3D volumes obtained by electron tomography and serial blockface EM. The datasets were subjected to anisotropic nonlinear diffusion first (Fernandez and Li, 2005) to reduce noise, preserve the features of interest and flatten the background. For this, the software package TomoAND and its capability of automated parameter tuning were used (Fernandez and Li, 2003, 2005). Afterwards, a scale-space operation was applied (with scale σ in the range $[1, 3]$) in order to ensure good ridge detection with the Gaussian model fitting. The density in tomograms was normalized so as to be in the range $[0, 1]$, with high values representing electron dense objects. The threshold t_M was set in the range $[0.2, 0.5]$ and t_v in $[500, 1000]$. Most of the datasets tested were taken from the CCDB database (ccdb.ucsd.edu), where detailed information about the preparation techniques are available (mostly using chemical fixation, high pressure freezing and freeze substitution). For ridge clustering, we used parameter thresholding and also the automated clustering methods K-means and SOMs with a number of initial clusters in the range 16–20. Similar results were obtained for the three clustering techniques, and for that reason only one will be shown in the following illustrative examples.

With regard to parameter setting, the best results were obtained with values in the ranges specified in the previous paragraph. The most critical parameter turns out to be σ, which is used in the scale-space operation. However, it can be easily tuned according to the scale (i.e. thickness) of the membranes or structures that are targeted. The threshold t_M used in the ridge detector also has an important influence because it contributes to reduce spurious ridges. It should be set by trial-and-error, though the range given above is expected to be broadly applicable. The threshold on the volume size t_v used for region merging strongly depends on the dataset and the structures under study. However, we found that optimal setting of this parameter is not crucial, in general. Finally, we also found that the number of clusters involved in the automated clustering methods is not critical either. A moderate number of clusters should be used to accomodate the variability of ridges present in the dataset.

4.1. Segmentation of experimental datasets

The first dataset was an electron tomogram (CCDB, ID: 8154) containing an axonal mitochondrion in a Schwann cell of the peripheral nerve of adult rat (Perkins and Ellisman, 2011). Fig. 3 illustrates the different steps of this segmentation framework on a small piece of the tomogram. Clustering of ridges was performed by thresholding the ridge parameters (Fig. 4), and the actual threshold values are presented in Suppl. Table S1. Fig. 5 shows the complete segmentation of the tomogram, where separation of the different structural components is apparent. The precise delineation of the fine details in the myelin sheath is particularly remarkable. This example clearly exhibits the potential of the method, as the components of this complex tomogram are extracted and labelled (i.e. coloured) with almost no user intervention (except threshold settings).

The second dataset was a tomogram focused on another type of neuronal mitochondrion, prepared as those taken from the CCDB. Fig. 6 shows a slice of the tomogram and the resulting segmentation. Ridge classification was also based on parameter thresholding, as shown in Table 1. The interest here was to confirm the ability of this ridge framework to identify and separate the membranes and cristae of the mitochondrion. Previous works failed to discriminate perfectly these structures (Martinez-Sanchez et al., 2011). However, the methodology presented here succeeds. It is important to note from Fig. 6 that the density within the mitochondrion is, in general, higher than the surrounding background. Therefore, the sharpness parameter is key in this segmentation (Table 1). Moreover, the asymmetries (in particular of offset and sharpness) are other important parameters as well. At the ridges associated to the mitochondrion membranes, these asymmetries are high because at one side the background is found whereas at the other side there is the denser content within the mitochondrion. However, the cristae are embedded in a relatively homogeneous content (hence, those asymmetries are not significant). Therefore, it was readily simple to set the values for parameter thresholding.

An electron tomogram of a cerebellar synapse (CCDB, ID: 3684) (Sosinsky et al., 2008) was then tested. Fig. 7 shows the resulting segmentation. In this particular example, ridge classification was performed using SOMs composed of 20 neurons. A user-guided post-processing stage to cluster the neurons was then applied, as usual in the field (Fernandez and Carazo, 1996), to end up with the 5 classes shown in Fig. 7. Note how well the pre- and post-synaptic membranes are extracted out based on the ridge parameters, as highlighted in yellow.

A dataset from adult mouse myocardium imaged by electron tomography was

also taken from the CCDB (ID: 3603) (Hayashi et al., 2009). It presented complex, densely packed structures that proved to be difficult to be separated even by manual delineation. The framework introduced here managed to segment the main structures in the tomogram (Fig. 8). Note that myosin fibres are well extracted as they are considered planes at a local level. This is partially due to one of the well-known artefacts introduced by electron tomography, which is the blurring along the beam direction (Fernandez and Li, 2003).

A tomogram of Vaccinia virus was used to test the robustness of the method on datasets with low SNR and low contrast typically found in the modality known as electron cryo-tomography (Cyrklaff et al., 2005). The method succeeded in extracting the membranes of the virion (outer and core's). As shown in Fig. 9, the difference of the characteristics of those membranes in terms of asymmetries and offset and sharpness is apparent, which was key for the segmentation.

The last example of application of this segmentation framework is a serial blockface EM volume from mouse retina (CCDB, ID: 7742) (Nguyen et al., 2011). SOMs were used for ridge classification. Twenty neurons were initially used, which were then clustered into seven groups to end up with the structures actually segmented (Fig. 10). The method succeeds in separating the different major layers within the retina. As illustrated in the figure, the user-guided post-processing stage is key to fuse clusters and yield the definite set of seven classes. The need for this stage arises from the fact that a given structural feature may be composed by ridges with different parameters (e.g. offset, thickness, etc.), as clearly happened with the cells at the Inner or at the Outer Nuclear Layer (Fig. 10, centre).

4.2. Validation results

For quantitative assessment, a synthetic phantom resembling the axonal mitochondrion in a Schwann cell (see Fig. 5) was designed. Six different biological structures were included: axon membrane, myelin sheath, a vesicle, and an axonal mitochondrion including outer membrane, crista and crista junction (Suppl. Fig. S3). Scattered blobs intended to resemble macromolecular complexes were also spread throughout. To simulate noise conditions, Gaussian white noise was added to produce versions at signal-to-noise ratio (SNR) of 1 and 6.

The phantom was subjected to segmentation by TomoSegMem (Martinez-Sanchez et al., 2011) and by the ridge-based procedure using scale-space with σ=2. The segmentation results with both methods can be seen in Suppl. Fig. S4. The results with SNR of 6 and 1 turned out to be very similar, thus showing robustness against noise. The limitation of TomoSegMem to discern different structures is clearly seen: it can only classify isolated structures by their size. Structures that are connected or that have the same size are thus considered as a whole (for instance, the mitochondrion: membrane, crista and junction; or axon membrane and myelin sheath). Another limitation is that it is unable to get rid of spurious locally planar features apposed to the true membranes (for instance, the inner side of the vesicle's membrane). Ridge-based segmentation overcomes these limitations (Suppl. Fig. S4). Ridge clustering by any of the strategies (SOMs, K-means, parameter thresholding) yielded very similar results.

The phantom represents the 'ground truth', which allows calculation of the fraction of true positives (TPF) and true negatives (TNF) based on voxel-based comparisons. In addition, complementary metrics about shape analysis (Martinez-Sanchez et al., 2011) were also measured, obtaining values higher than or around

90%. Suppl. Tables S2–S9 contain a detailed report of the quantitative results. For brevity, only a representative subset of TPF and TNF as well as average shape values are presented here (Table 2). As expected, the results show that, in global terms, both algorithms (TomoSegMem vs. Ridge-based) behave similarly if just detection of the structures is considered (i.e. row labelled as Global in Table 2), with high sensitivity (TPF > 91%) and higher specificity (TNF > 98%). However, ridge-based segmentation stands out by its ability to extract the six different structures present in the phantom with extremely high specificity. It exhibits high sensitivity as well, yet it is still susceptible to false positives due to other structures apposed to the local planes (e.g. at the crista junction, which is a fine detail, or at the vesicle due to the dense content). The three strategies for ridge clustering yielded similar performance in global terms, as demonstrated by Suppl. Tables S3–S5,S7–S9. To further study the results with these classification strategies, we computed the centroids of the clusters derived from SOM and K-means (Suppl. Tables S10 and S11) and then calculated the Euclidean distance between them for the six classes (Axon, Myelin, Mem.Mito, Crista, Junction and Vesicle). The distance turned out to be 0.15, 0.16, 0.24, 0.08, 0.39, 0.37, respectively, which are relatively low values for a maximum of $\sqrt{8}$ in the eight-dimensional space where the components are normalized to be in $[0, 1]$. There are two classes with poorer agreement: Junction (distance of 0.39) and Vesicle (0.37). As discussed a few lines above, these classes are precisely characterized by a slightly larger rate of false positives (i.e. poorer TPF). A comparison of the parameter thresholds with these centroids (Suppl. Tables S10–S12) shows that they all are, in general, in good agreement. Therefore, the similar behaviour of the different clustering strategies is confirmed.

5. Discussion and Conclusion

A new framework for segmentation of 3D EM datasets has been introduced. The methodology relies on the detection of planar ridges and their characterisation according to asymmetrical Gaussian model fitting. The wealth of local information obtained this way is then exploited for ridge clustering and the subsequent extension to the tomogram. The application to representative complex experimental datasets has clearly shown the good performance of the technique. A quantitative analysis has also proved its high specificity and sensitivity.

The framework consists of several consecutive stages. The first is the scale-space representation of the tomogram, which has a two-fold aim. It helps to reduce noise and smear out all structures lower than the scale of interest. Secondly, it ensures that local maxima in the resulting tomogram approach a Gaussian profile. More aggressive noise reduction techniques can be used before scale-space so as to substantially remove noise, flatten background and preserve and enhance features of interest. Here we have systematically used anisotropic nonlinear diffusion. However, though in all cases it has been useful, the actual utility of this previous strong filtering depends on the dataset under consideration. In particular, for datasets with very low signal-to-noise ratios, noise filtering may turn out to be an essential pre-processing step, as happens with other segmentation approaches (Rigort et al., 2012; Rusu et al., 2012).

Extraction of ridge descriptors is then performed based on fitting with a Gaussian function. The ridges are thus characterized by eight-dimensional feature vectors, which can be used for their unsupervised classification. Here we have tested several techniques for this clustering, which have provided similar results. The simplest one is based on parameter thresholding carried out by the user in a

straightforward manner. The fact that the ridge descriptors correspond to features visually perceptible (e.g. offset, thickness, sharpness) makes them appropriate for this interactive process. Two automatic clustering methods have also been tested: K-means and Self-Organizing Maps. Though they work nicely in an automated fashion, we have observed that the user still needs to post-process the classification results. In particular, they have to be further clustered so as to yield the definite classes that actually represent the data. Our future interests include exploration of other techniques for clustering and dimensionality reduction, which could facilitate and accelerate the process.

This framework is specially focused on membranes and other structures that, at a local scale, can be considered planes. Membranes constitute natural boundaries of biological compartments, so their extraction and classification facilitate interpretation of the whole volume. There are, however, some membranes that may not comply with the Gaussian model used here (e.g. those associated to filled vesicles would essentially be steps). These should be segmented using a different approach. On the other hand, in electron tomography the structural features are blurred along the electron beam direction because of the imaging conditions found in this discipline. This fact actually turns any structural detail into a plane at local scale. For that reason, this framework may also be helpful for segmentation with general applicability in that field.

Acknowledgements

Dr. GA Perkins kindly provided the mitochondrion dataset (Fig. 6). This work has been partially supported by the Spanish Ministry of Science (TIN2008-01117, TIN2012-37483), J. Andalucia (P10-TIC-6002, P11-TIC-7176), in part

thanks to European Reg. Dev. Funds (ERDF). AMS is a fellow of the Spanish FPI programme.

References

Arbelaez, P., Maire, M., Fowlkes, C., Malik, J., 2011. Contour detection and hierarchical image segmentation. IEEE Tr. Pattern Anal. Mach. Intell. 33, 898–916.

Bauer, C., Pock, T., Sorantin, E., Bischof, H., Beichel, R., 2010. Segmentation of interwoven 3D tubular tree structures utilizing shape priors and graph cuts. Med. Image Anal. 14, 172–184.

Cyrklaff, M., Risco, C., Fernandez, J. J., Jimenez, M. V., Esteban, M., et al., 2005. Cryo-electron tomography of vaccinia virus. Proc. Natl. Acad. Sci. USA 102, 2772–2777.

Denk, W., Horstmann, H., 2004. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. PLoS Biology 2, 1900–1909.

Eberly, D., 1994. Fast algorithms for ridge construction. Proceedings SPIE Photonics 2356, 231–242.

Fernandez, J. J., 2009. Tomobflow: Feature-preserving noise filtering for electron tomography. BMC Bioinformatics 10:178.

Fernandez, J. J., 2012. Computational methods for electron tomography. Micron 43, 1010–1030.

Fernandez, J. J., Carazo, J. M., 1996. Analysis of structural variability within two-dimensional biological crystals by a combination of patch averaging techniques and self organizing maps. Ultramicroscopy 65, 81–93.

Fernandez, J. J., Li, S., 2003. An improved algorithm for anisotropic nonlinear diffusion for denoising cryo-tomograms. J. Struct. Biol. 144, 152–161.

Fernandez, J. J., Li, S., 2005. Anisotropic nonlinear filtering of cellular structures in cryo-electron tomography. Comput. Sci. Eng. 7 (5), 54–61.

Florack, L. M. J., ter Haar Romeny, B. M., Koenderink, J. J., Viergever, M. A., 1992. Scale and differential structure of images. Image Vis. Comput. 10, 376–388.

Frank, J., 2006. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Visualization of Biological Molecules in Their Native State. Oxford University Press.

Garduno, E., Wong-Barnum, M., Volkmann, N., Ellisman, M. H., 2008. Segmentation of electron tomographic data sets using fuzzy set theory principles. J. Struct. Biol. 162, 368–379.

Harris, K. M., 1999. Structure, development, and plasticity of dendritic spines. Curr. Opin. Neurobiol 9, 343–348.

Hayashi, T., Martone, M. E., Yu, Z., Thor, A., Doi, M., et al., 2009. Three-dimensional electron microscopy reveals new details of membrane systems for calcium signaling in the heart. J. Cell Sci. 122, 1005–1013.

Jain, V., Seung, H. S., Turaga, S. C., 2010. Machines that learn to segment images: a crucial technology for connectomics. Curr. Opin. Neurobiol. 20, 653–666.

Kirbas, C., Francis, Q., 2004. A review of vessel extraction techniques and algorithms. ACM Comput. Surv. 36, 81–121.

Kohonen, T., 2001. Self-Organizing Maps. Springer, Berlin.

Langlois, R., Frank, J., 2011. A clarification of the terms used in comparing semi-automated particle selection algorithms in cryo-em. J. Struct. Biol. 175, 348–352.

Lucic, V., Forster, F., Baumeister, W., 2005. Structural studies by electron tomography: From cells to molecules. Annu. Rev. Biochem. 74, 833–865.

Marabini, R., Carazo, J. M., 1994. Pattern recognition and classification of images of biological macromolecules using artificial neural networks. Biophys. J. 66, 1804–1814.

Marslan, S., 2009. Machine Learning: An Algorithmic Perspective. CRC Press, Boca Raton, FL.

Martinez-Sanchez, A., Garcia, I., Fernandez, J. J., 2011. A differential structure approach to membrane segmentation in electron tomography. J. Struct. Biol. 175, 372–383.

Nguyen, J. V., Soto, I., Kim, K. Y., Bushong, E. A., Oglesby, E., et al., 2011. Myelination transition zone astrocytes are constitutively phagocytic and have synuclein dependent reactivity in glaucoma. Proc. Natl. Acad. Sci. USA 108, 1176–1181.

Perkins, G. A., Ellisman, M. H., 2011. Mitochondrial configurations in peripheral nerve suggest differential atp production. J. Struct. Biol. 173, 117–127.

Rigort, A., Gunther, D., Hegerl, R., Baum, D., Weber, B., et al., 2012. Automated

segmentation of electron tomograms for a quantitative description of actin filament networks. J. Struct. Biol. 177, 135–144.

Rusu, M., Starosolski, Z., Wahle, M., Rigort, A., Wriggers, W., 2012. Automated tracing of filaments in 3D electron tomography reconstructions using sculptor and situs. J. Struct. Biol. 178, 121–128.

Sato, Y., Tanaka, H., Nishii, T., Nakanishi, K., Sugano, N., et al., 2003. Limits on the accuracy of 3d thickness measurement in MR images-effects of voxel anisotropy. IEEE Trans. Med. Imaging 22, 1076–1088.

Sosinsky, G. E., Crum, J., Jones, Y. Z., Lanman, J., Smarr, B., et al., 2008. The combination of chemical fixation procedures with high pressure freezing and freeze substitution preserves highly labile tissue ultrastructure for electron tomography applications. J. Struct. Biol. 161, 359–371.

Soto, G. E., Young, S. J., Martone, M. E., Deerinck, T. J., Lamont, S., et al., 1994. Serial section electron tomography: A method for three-dimensional reconstruction of large structures. Neuroimage 1, 230–243.

Volkmann, N., 2010. Methods for segmentation and interpretation of electron tomographic reconstructions. Methods Enzymol. 483, 31–46.

Weber, B., Greenan, G., Prohaska, S., Baum, D., Hege, H. C., et al., 2012. Automated tracing of microtubules in electron tomograms of plastic embedded samples of Caenorhabditis elegans embryos. J. Struct. Biol. 178, 129–138.

Yu, Z., Frangakis, A. S., 2011. Classification of electron sub-tomograms with neural networks and its application to template-matching. J. Struct. Biol. 174, 494–504.

Figure Legends

Figure 1. Ridge characterization in 1D. (left) Six ridges (A–F) with different properties that constitute the details of interest in a 1D function. Black arrows denote antidetails, which represent the boundaries of the local neighbourhood around each detail of interest. (right) Description of ridges in terms of parameters based on a Gaussian model (see main text). Here, only asymmetry associated to the offset is shown. This description allows classification of ridges based on offset ({A,B,C,D} and {E,F}), sharpness ({A,E}, {D} and {B,C,F}), thicknesss ({A,B,D,E,F} and {C}), offset asymmetry ({A} and {B,D,C,E,F}), resolution ({A,D,E}, {B}, {C}, and {F}) or any other far more complex criterion by combining parameters.

Figure 2. Sketch of ridge parameters over a zoomed view of ridge A in Fig. 1.

Figure 3. Steps of ridge-based segmentation. (top-left) Slice of the original tomogram of axonal mitochondrion. (top-right) Pre-processed data (contrast inversion, noise reduction and scale-space). (bottom-left) Classification of detected ridges based on parameter thresholding (see actual ridge parameters in Fig. 4) using the thresholds in Suppl. Table S1. (bottom-right) Extension of ridge classification to voxel segmentation (color code as in Fig. 5).

Figure 4. Parameters describing the ridges in the tomogram of axonal mitochondrion. top-left: offset. top-right: sharpness. bottom-left: thickness. bottom-right: asymmetry of sharpness. Values are according to the colormap on the right.

Figure 5. Segmentation of axonal mitochondrion tomogram. (left) Slice of the original tomogram. The rectangle encloses the data shown in Figs. 3 and 4. (centre and right) Different views of the segmented tomogram. Visualization at a higher level of detail of the segmentation result of the enclosed area is available in Figs. 3 and 4.

Color code: yellow – mitochondrion membrane; pink – mitochondrion cristae; green – axon membrane and other axoplasmic plane-like structures; violet – myelin sheath of the Schwann cell; red – Schwann cell's mitochondrion; light green – Schwann cell's membranous structure. In transparency, other sharp axonal structures (mainly microtubules and neurofilaments).

Figure 6. Neuronal mitochondrion. (left) Slice of the pre-processed tomogram. (right) Segmentation with the ridge-based framework. Arrows indicate areas where this framework behaves particularly well and overcomes the failures and misclassification of other membrane segmentation approach (Martinez-Sanchez et al. (2011), Fig. 9). Dataset courtesy of Dr. G.A. Perkins.

Color code: yellow – mitochondrion membrane; pink – cristae.

Figure 7. Cerebellar synapse. (left) Slice of the pre-processed data. (right) Segmented tomogram with the proposed method using SOMs for ridge classification. Color code: yellow – pre- and post-synaptic membranes; pink – vesicles; green – mitochondrion membrane; violet – mitochondrion crista; red – other membranous structures.

Figure 8. Mouse myocardium. (top-left) Slice of the pre-processed data. (right) Segmented tomogram with the proposed method using SOMs for ridge classification. (bottom-left) The area dashed in the right panel is shown at a higher level of detail and overlying the density data.
Color code: violet – Z-bands; pink – myosin fibres; yellow – T-tubules, junctional sarcoplasmic reticulum and neighbour mitochondria.

Figure 9. Vaccinia virus. (left) Slice of the original cryo-tomogram. (right) Segmentation with the ridge framework using SOM for ridge classification.
Color code: yellow – outer membrane; pink – membrane of the core; transparent blue – lateral bodies.

Figure 10. Mouse retina serial blockface EM. (left) Slice of the original data. (centre) Ridge classification (only the area boxed at left panel is presented) as comes directly from SOM (i.e. 20 classes as shown in different colors and tones) and after clustering the neurons into 7 groups. (right) 3D view of the segmented volume with color code: yellow and red – cells at the Inner Nuclear Layer; pink – cells at the Outer Nuclear Layer; light blue and light green – Outer Plexiform Layer; dark green – Inner Plexiform Layer; dark blue – Inner segments.

Figures

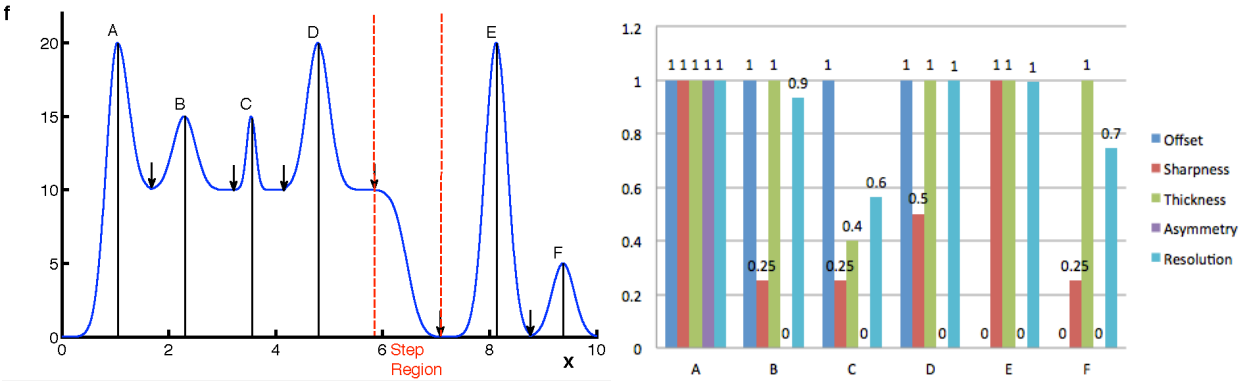

Figure 1: Ridge characterization in 1D. (left) Six ridges (A–F) with different properties that constitute the details of interest in a 1D function. Black arrows denote antidetails, which represent the boundaries of the local neighbourhood around each detail of interest. (right) Description of ridges in terms of parameters based on a Gaussian model (see main text). Here, only asymmetry associated to the offset is shown. This description allows classification of ridges based on offset ({A,B,C,D} and {E,F}), sharpness ({A,E}, {D} and {B,C,F}), thicknesss ({A,B,D,E,F} and {C}), offset asymmetry ({A} and {B,D,C,E,F}), resolution ({A,D,E}, {B}, {C}, and {F}) or any other far more complex criterion by combining parameters.

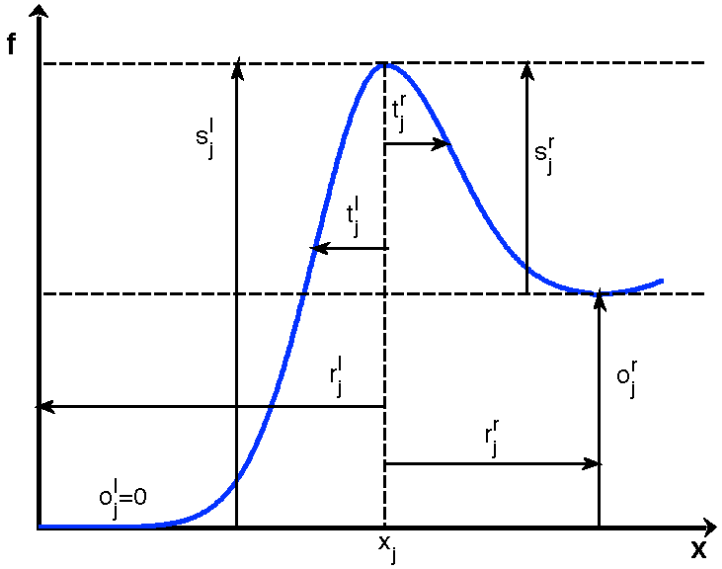

Figure 2: Sketch of ridge parameters over a zoomed view of ridge A in Fig. 1.

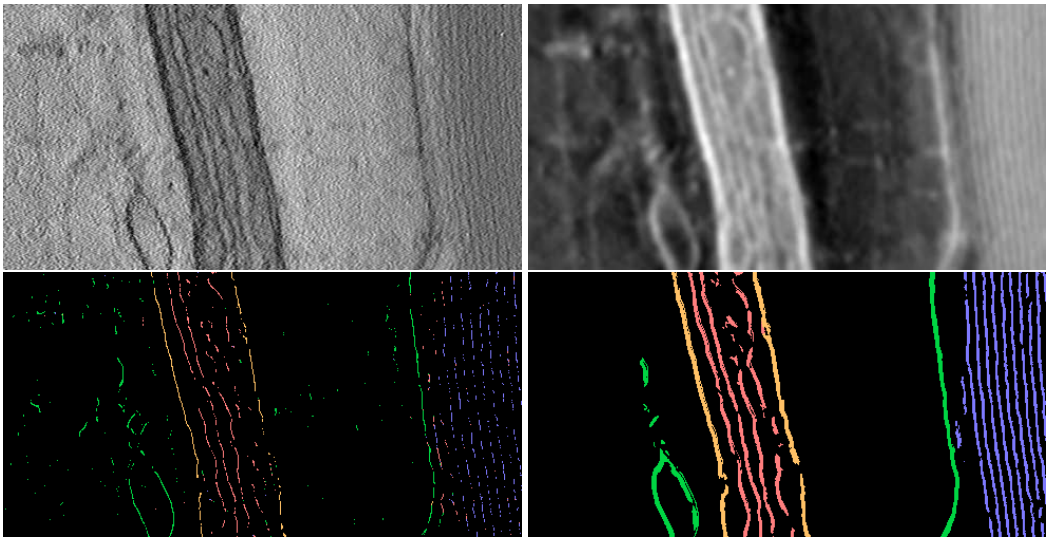

Figure 3: Steps of ridge-based segmentation. (top-left) Slice of the original tomogram of axonal mitochondrion. (top-right) Pre-processed data (contrast inversion, noise reduction and scale-space). (bottom-left) Classification of detected ridges based on parameter thresholding (see actual ridge parameters in Fig. 4) using the thresholds in Suppl. Table S1. (bottom-right) Extension of ridge classification to voxel segmentation (color code as in Fig. 5).

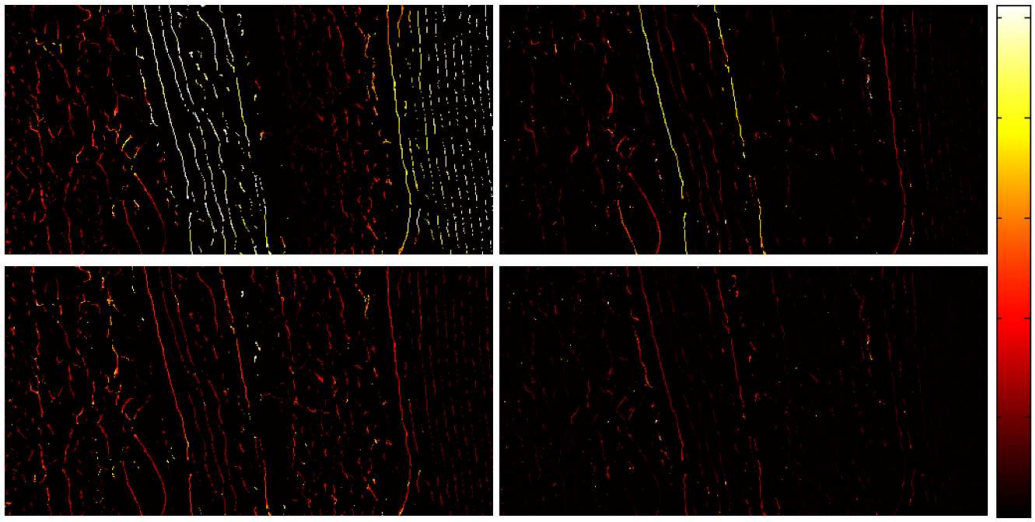

Figure 4: Parameters describing the ridges in the tomogram of axonal mitochondrion. top-left: offset. top-right: sharpness. bottom-left: thickness. bottom-right: asymmetry of sharpness. Values are according to the colormap on the right.

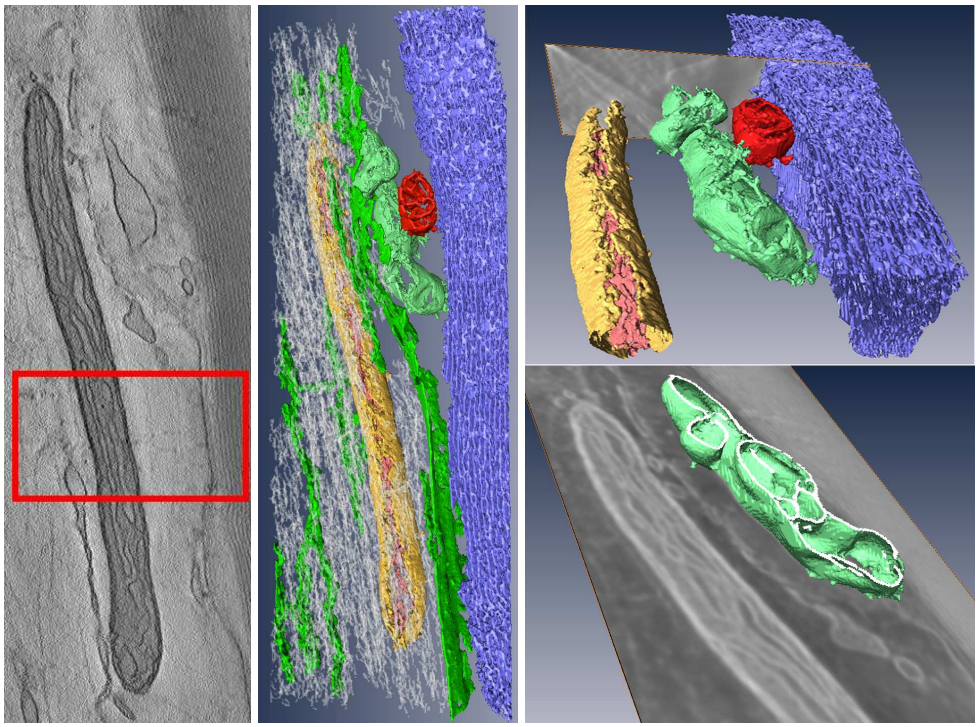

Figure 5: Segmentation of axonal mitochondrion tomogram. (left) Slice of the original tomogram. The rectangle encloses the data shown in Figs. 3 and 4. (centre and right) Different views of the segmented tomogram. Visualization at a higher level of detail of the segmentation result of the enclosed area is available in Figs. 3 and 4.
Color code: yellow – mitochondrion membrane; pink – mitochondrion cristae; green – axon membrane and other axoplasmic plane-like structures; violet – myelin sheath of the Schwann cell; red – Schwann cell's mitochondrion; light green – Schwann cell's membranous structure. In transparency, other sharp axonal structures (mainly microtubules and neurofilaments).

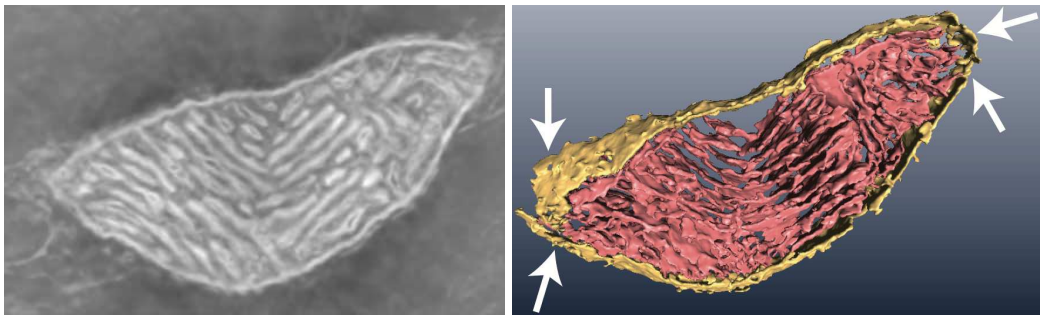

Figure 6: Neuronal mitochondrion. (left) Slice of the pre-processed tomogram. (right) Segmentation with the ridge-based framework. Arrows indicate areas where this framework behaves particularly well and overcomes the failures and misclassification of other membrane segmentation approach (Martinez-Sanchez et al. (2011), Fig. 9). Dataset courtesy of Dr. G.A. Perkins.
Color code: yellow – mitochondrion membrane; pink – cristae.

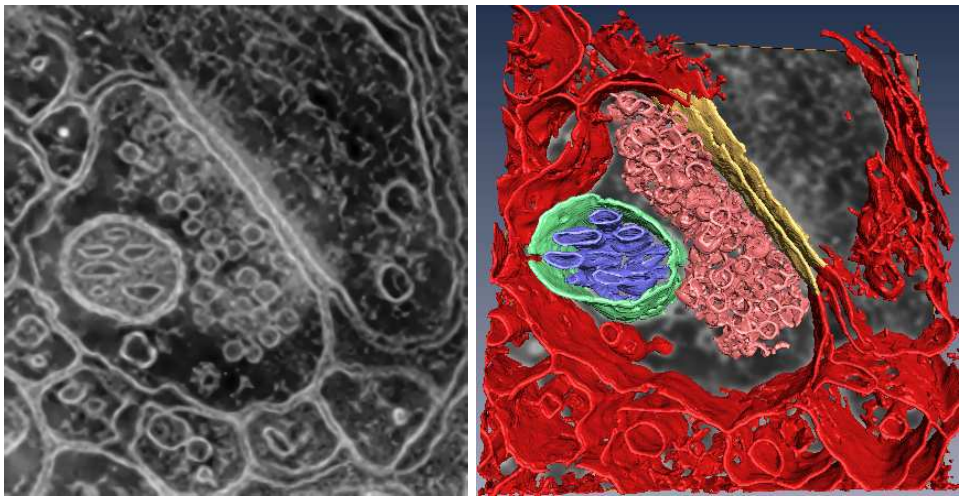

Figure 7: Cerebellar synapse. (left) Slice of the pre-processed data. (right) Segmented tomogram with the proposed method using SOMs for ridge classification.
Color code: yellow – pre- and post-synaptic membranes; pink – vesicles; green – mitochondrion membrane; violet – mitochondrion crista; red – other membranous structures.

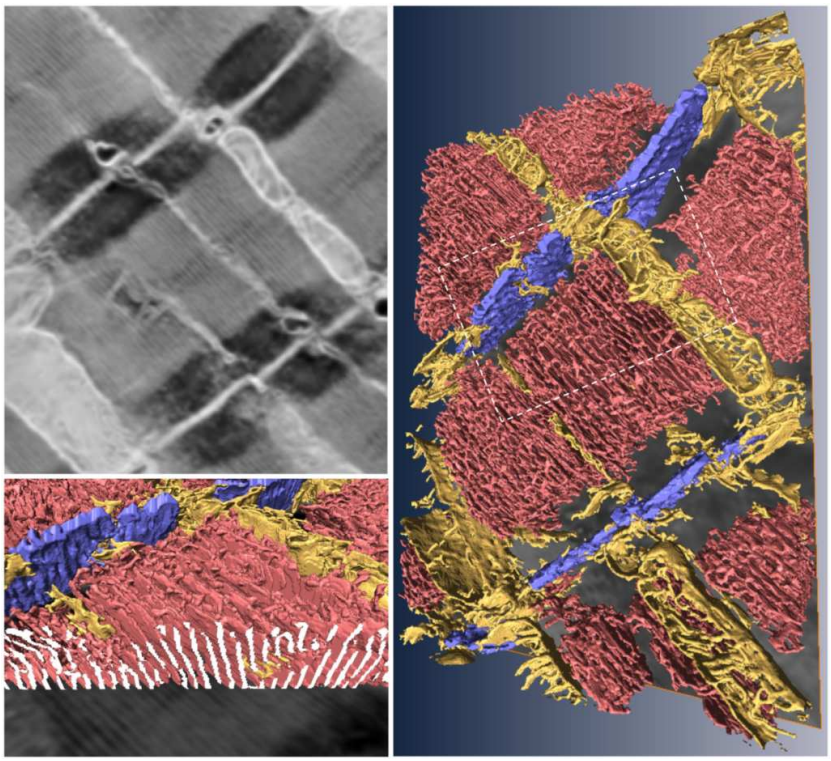

Figure 8: Mouse myocardium. (top-left) Slice of the pre-processed data. (right) Segmented tomogram with the proposed method using SOMs for ridge classification. (bottom-left) The area dashed in the right panel is shown at a higher level of detail and overlying the density data.
Color code: violet – Z-bands; pink – myosin fibres; yellow – T-tubules, junctional sarcoplasmic reticulum and neighbour mitochondria.

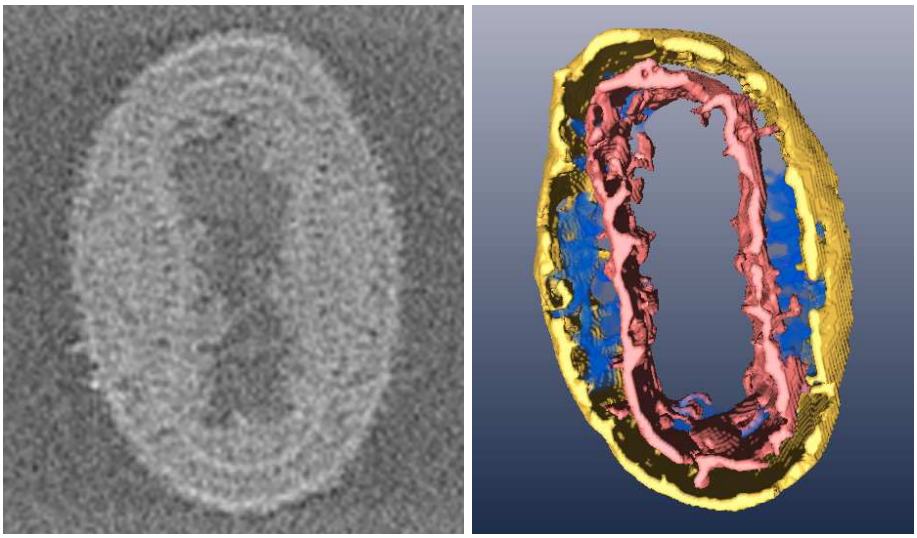

Figure 9: Vaccinia virus. (left) Slice of the original cryo-tomogram. (right) Segmentation with the ridge framework using SOM for ridge classification.
Color code: yellow – outer membrane; pink – membrane of the core; transparent blue – lateral bodies.

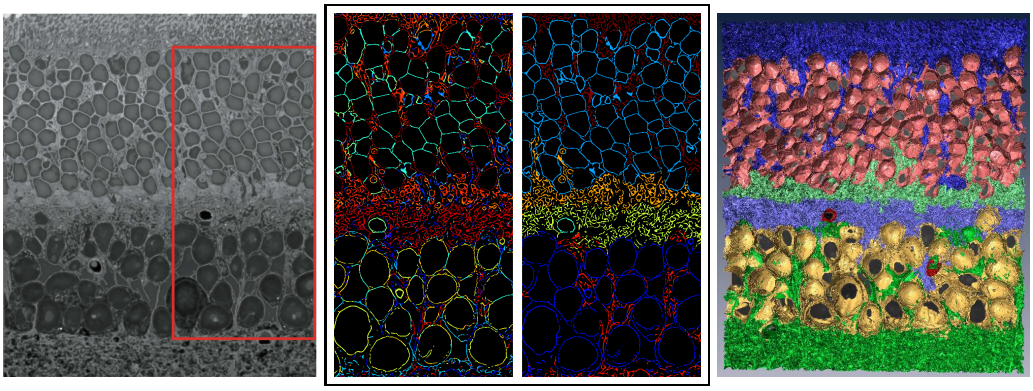

Figure 10: Mouse retina serial blockface EM. (left) Slice of the original data. (centre) Ridge classification (only the area boxed at left panel is presented) as comes directly from SOM (i.e. 20 classes as shown in different colors and tones) and after clustering the neurons into 7 groups. (right) 3D view of the segmented volume with color code: yellow and red – cells at the Inner Nuclear Layer; pink – cells at the Outer Nuclear Layer; light blue and light green – Outer Plexiform Layer; dark green – Inner Plexiform Layer; dark blue – Inner segments.

Tables

Table 1: Parameter thresholding for segmentation in Fig. 6. [1]

Class	o	a^o	s	a^s	t	a^t	r	a^r
1	↑ .4	↑ .15	↑ .2	↑ .2	–	–	↑ 30	–
2	↑ .4	–	↓ .2	↓ .2	↑ 2	–	↓ 20	↓ .3

[1] The eight ridge parameters represent offset (o), sharpness (s), thickness (t), resolution (r) and their corresponding assymetries (a^o, a^s, a^t, a^r).
Class 1 and 2 represent membranes and cristae, respectively. The values indicate the actual thresholds used for the segmentation. '↑' indicates upthresholding (higher than), '↓' downthresholding (lower than) and '–' no thresholding applied for this parameter.

Table 2: Quantitative analysis based on phantom (%). [2]

Class	SNR=1 TPF	SNR=1 TNF	SNR=1 Shape	SNR=6 TPF	SNR=6 TNF	SNR=6 Shape
			TomoSegMem			
Myelin	95.54	99.85	99.13	97.50	99.99	99.27
Mito.	85.03	98.93	96.87	86.79	99.78	98.79
Vesicle	71.28	99.46	93.80	70.68	99.44	94.04
Global	91.33	98.01	97.15	93.12	99.09	98.16
			Ridge-based Segmentation			
Axon	90.32	99.97	98.36	88.09	99.98	98.38
Myelin	97.99	99.71	98.56	98.45	99.75	98.77
Mem.Mito.	85.69	99.76	98.82	85.61	99.79	98.84
Crista	96.81	99.61	96.01	96.61	99.62	96.55
Junction	43.98	99.99	95.36	74.20	99.99	94.92
Vesicle	67.59	99.69	97.00	67.45	99.75	96.72
Global	92.76	98.65	98.26	92.87	98.77	98.46

[2] Global denotes that all segmented structures are treated as a whole, i.e. belonging to only one class.

Shape denotes the average value of the five metrics used for the shape analysis (bounding box, centroid and axes).

The results shown for the ridge-based segmentation were obtained using SOMs for classification. The results with other classification methods turned out to be similar (see Suppl. Tables S2–S9)