

Model-Based Automated Extraction of Microtubules from Electron Tomography Volume

Ming Jiang, Qiang Ji*, *Senior Member, IEEE*, and Bruce F. McEwen

I. INTRODUCTION

Microtubules form the skeletons of cells and serve as the intracellular network for protein transportation. More importantly, the microtubules attach to chromosomes via kinetochore to drive the movement of chromosomes during mitosis. Recently, the application of electron microscopy by structural biologists has opened up new possibilities in investigating kinetochore microtubule functions by observing their high resolution structures [1]. Unfortunately, in the low SNR and low contrast tomography volume, the extraction of microtubules is particularly challenging since they are in close contact with the cellular environment and are densely surrounded by proteins of similar appearance, as shown in Figure 1. As a result, practical segmentation of the tomography volume is exclusively a manual operation, which is very subjective and time consuming. While interpolation from a few contours can produce an approximation [2], it is not accurate enough for quantitative biological study related with this work [1]. Therefore, improving the accuracy and efficiency of volume segmentation and reducing the subjectivity will have enormous impact on the study of microtubules.

The segmentation of biological structures from electron tomography has already received some attention. Shen et al. [3] reported the segmentation of filamentous structures by propagating manually segmented features through successive slices. The segmentation is not robust since in each slice it uses 2-D features only. Bartesaghi et al. [4] developed a geodesic active contour for the semi-automatic segmentation of features in cellular electron tomography. A point inside cell features is required to initialize the algorithm. Bajaj and Yu [5] proposed a fast marching based method for boundary segmentation and skeletonization of cell tomographic imaging. Since the speed function of the fast marching is defined using image gradient only, this method cannot handle image noise and contrast variations. Recently, normalized cut [6] and watershed [7] have also been applied to the segmentation of electron tomography data. While normalized cut has difficulty in incorporating high level domain-specific constraints, the watershed method has the drawback of over-segmentation and the results are not suitable for quantitative analysis.

Outside the field of electron tomography, parametric de-

formable models [8] [9] and geometric deformable models [10][11][12][13] have been proposed for the segmentation of blood vessels and thin tubular features respectively. As these pure data-driven methods are unable to converge correctly when image quality is poor, statistical shape constraints are imposed to make them robust [14][15][16]. However, it is difficult to build 3D statistical shape models for microtubules due to their large length variations and the limited number of available training samples.

In this paper, we propose a model-based automated approach for the extraction of microtubules in three major phases: volume enhancement, microtubule localization and microtubule boundary segmentation. In the first phase, the volume is first processed with an anisotropic invariant wavelet transform to accentuate the microtubule globally, followed by a 3D tube-enhancing filter to enhance the microtubules locally. A modified coherence-enhancing diffusion (CED) then completes the interruptions along the microtubules. In the second phase, the microtubules are localized with a center line extraction algorithm adapted for tubular structures. In the third phase, the localized microtubules are finally segmented with an extended active shape model (ASM) method [17] that allows robust segmentation and certain deformation.

Fully taking advantage of the geometric and photometric properties of the tubular objects, our contributions can be summarized as follows. First, we adapt wavelets using anisotropic basis and invariant transform to globally enhance elongated features in low SNR images. Second, we develop the spatial domain filters based on the shape information of the local image structures to achieve superior enhancement performance against low contrast and noisy images. Third, we realize robust segmentation with an improved statistical shape model to characterize the cross-sectional contours of the tubular objects. Fourth, statistical shape model is then combined with Kalman filtering to impose smoothness constraint along the longitudinal direction of the microtubules. The segmentation improved in this way is reliable against missing data and outliers in the tomography volume.

The automated approach was experimented on kinetochore tomography volumes acquired with electron microscopy. Quantitative study shows that the microtubules segmented with automated approach were in close agreement with those produced by manual segmentation and using only a fraction of the time of the latter.

II. PROPOSED METHOD

The proposed method consists of three major steps: volume enhancement, microtubule localization and microtubule

Manuscript received August 9, 2004. This work was supported in part by NIH grants GM066270 and RR01219 and the Wadsworth Core Facility for Electron Microscopy

M. Jiang and *Q. Ji (qji@ecse.rpi.edu) are with the Department of Electrical, Computer and System Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180. Bruce F. McEwen is with Wadsworth Center, New York State Dept. Health, Albany, NY 12201

boundary segmentation. Volume enhancement accentuates microtubules, while removing noise and irrelevant structures. The localization determines the spatial locations of the microtubules. The boundary segmentation reconstructs the microtubule's surface morphology. In the sections to follow, we outline our algorithms for each step.

A. Volume Enhancement

The low SNR and the presence of irrelevant but intensity-similar structures near the microtubules necessitates a preprocessing step to enhance the microtubules. To this end, we propose a hybrid method by combining transform domain technique and spatial domain techniques: anisotropic invariant wavelet filtering, 3D tube-enhancing filter and a modified CED, as detailed below.

1) *Anisotropic Invariant Wavelet Filtering*: In the tomography volume, the microtubule features are so weak locally that any general averaging de-noising technique will easily smooth out the microtubules completely. Considering the elongated shape of the microtubules, we propose to enhance them using wavelet transform with anisotropic basis, which is advantageous in enhancing anisotropic image features [18][19][20].

Given a 1-D orthogonal wavelet basis

$$\psi_{j,k}(t) = 2^{-(j/2)}\psi(2^{-j}t - k), \quad j, k, t \in Z \quad (1)$$

where Z is the integer set, j denotes scale and k denotes the translation, the anisotropic basis functions for higher dimension can be constructed by the tensor product of the 1-D wavelet basis functions. For example, the basis functions for 2-D wavelet transform can be obtained as

$$\psi_{j_1, k_1; j_2, k_2}(t_x, t_y) = \psi_{j_1, k_1}(t_x) \cdot \psi_{j_2, k_2}(t_y) \quad (2)$$

where $j_i, k_i \in Z, (t_x, t_y) \in Z^2$. Therefore, the anisotropic basis has more combinations of different scales/frequencies along horizontal and vertical directions. For example, the scales of the isotropic basis combine as 1 : 1, 2 : 2, 3 : 3, whereas the scales of the anisotropic basis combine as 1 : 1, 1 : 2, 1 : 3, 2 : 1, 2 : 2, 2 : 3, 3 : 1, 3 : 2, 3 : 3. This makes anisotropic basis suitable for capturing elongated objects.

To make the transform shift invariant, we use averaging over shifts: applying wavelet transform to shifted version of the image, thresholding the wavelet transform coefficients, shifting back the de-noised image and averaging the results from all shifted versions of the image. Mathematically, this shift invariant de-noising process can be expressed as

$$\bar{T}(x) = Ave[S_{-h}(T(S_h(x)))] \quad (3)$$

where \bar{T} is the shift invariant de-noising process, x is the noisy image, h is a set of shifts, Ave is the average operation, S is the shift operation, T is the shift sensitive de-noising operation. According to Beylkin's theory [21], the computation operation of this invariant transform can be reduced from $O(N^2)$ to $O(N \log N)$.

To handle the microtubules located along arbitrary orientation in the image, the orientation invariance is realized by averaging over rotation [22], which can be expressed as

$$\hat{T}(x) = Ave[R_{-\theta}(\bar{T}(R_\theta(x)))] \quad (4)$$

where \hat{T} is the rotation invariant de-noising operator, θ is a set of angles, R is a rotation operator.

Figure 2 (b) shows the enhancement of the tomography image with the anisotropic invariant wavelet transform. The microtubules, as elongated features, have their overall features well preserved for further processing. To produce the results in this work, Symmlet 8 basis is used. The images are normalized to 256×256 and 4-level decomposition is used. The threshold value was chosen as two percent of the maximum magnitude of the transform coefficients. In practice, the 2D wavelet transform is applied to each slice of the 3D volume. True 3D invariant transform is not used out of concern of computation cost and memory requirement.

2) *3D tube-enhancing Filter*: After wavelet filtering, the microtubules are still embedded in the low contrast volume. In this situation, we propose to enhance the volume with a 3D tube-enhancing filter by exploiting the local geometric properties of tubular structures.

Using the eigenvalues of the Hessian matrix, Frangi [23] developed a filter to enhance tubular structures. The problem of that filter is that gradient information is not used. The output of the filter typically decays smoothly towards the boundary of the tubular object. In our application, a filter producing sharp response will be more effective for feature localization. To address this issue, we improve the tube-enhancing filter with the eigenvalues of the Weingarten matrix derived using both first order and second order image derivatives [24]. Consider the hyper-surface $S(x, y, z)$ associated with the 3D volume intensity $I(x, y, z)$ described by

$$S(x, y, z) = (x, y, z, I(x, y, z)) \quad (5)$$

then the normal to the hyper-surface at (x, y, z) is given by

$$\vec{n} = \vec{S}_x \wedge \vec{S}_y \wedge \vec{S}_z = \frac{1}{\sqrt{1 + I_x^2 + I_y^2 + I_z^2}}(I_x, I_y, I_z, -1)^T \quad (6)$$

The first fundamental form of the hyper-surface is thus

$$\begin{aligned} F_1 &= \begin{pmatrix} \vec{S}_x \cdot \vec{S}_x & \vec{S}_x \cdot \vec{S}_y & \vec{S}_x \cdot \vec{S}_z \\ \vec{S}_x \cdot \vec{S}_y & \vec{S}_y \cdot \vec{S}_y & \vec{S}_y \cdot \vec{S}_z \\ \vec{S}_x \cdot \vec{S}_z & \vec{S}_y \cdot \vec{S}_z & \vec{S}_z \cdot \vec{S}_z \end{pmatrix} \\ &= \begin{pmatrix} 1 + I_x^2 & I_x I_y & I_x I_z \\ I_x I_y & 1 + I_y^2 & I_y I_z \\ I_x I_z & I_y I_z & 1 + I_z^2 \end{pmatrix} \end{aligned} \quad (7)$$

The second fundamental form is

$$\begin{aligned} F_2 &= - \begin{pmatrix} \vec{S}_x \cdot \frac{\partial \vec{n}}{\partial x} & \vec{S}_x \cdot \frac{\partial \vec{n}}{\partial y} & \vec{S}_x \cdot \frac{\partial \vec{n}}{\partial z} \\ \vec{S}_y \cdot \frac{\partial \vec{n}}{\partial x} & \vec{S}_y \cdot \frac{\partial \vec{n}}{\partial y} & \vec{S}_y \cdot \frac{\partial \vec{n}}{\partial z} \\ \vec{S}_z \cdot \frac{\partial \vec{n}}{\partial x} & \vec{S}_z \cdot \frac{\partial \vec{n}}{\partial y} & \vec{S}_z \cdot \frac{\partial \vec{n}}{\partial z} \end{pmatrix} \\ &= - \frac{1}{\sqrt{1 + I_x^2 + I_y^2 + I_z^2}} \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \end{aligned} \quad (8)$$

The Weingarten matrix is then obtained as [24]

$$W = F_2 \cdot F_1^{-1} \quad (9)$$

The image derivatives involved are calculated reliably using a cubic fitting model, the details of which can be found in our previous work [25].

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of the matrix W with $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$, and v_1, v_2, v_3 be the corresponding eigenvectors. The eigenvalues are the curvatures of the hyper-surface, with the corresponding eigenvectors representing the principal directions. The relations among these curvatures provide insight into the shape of the local image structures.

For bright tubular structures against dark background in 3D image, it can be assumed that λ_1 is small, while both λ_2 and λ_3 have large magnitude and negative signs. The tubular criteria can thus be summarized as $|\lambda_1| \approx 0$, $|\lambda_1| \ll |\lambda_2|$ and $\lambda_2 \approx \lambda_3$. We then propose the following filter to estimate the local tubular shape

$$L = \begin{cases} 0 & \text{if } \lambda_2 > 0 \text{ or } \lambda_3 > 0 \\ (1 - \exp(-\frac{K_a^2}{2\alpha^2}))\exp(-\frac{K_b^2}{2\beta^2})(1 - \exp(-\frac{h^2}{2\gamma^2})) & \text{else} \end{cases} \quad (10)$$

where $K_a = \frac{|\lambda_2|}{|\lambda_3|}$ indicates cross-sectional circularity and $K_b = \frac{|\lambda_1|}{\sqrt{|\lambda_2\lambda_3|}}$ indicates elongation. $h = \|\mathbf{W}\|_F = \sqrt{\sum_i \lambda_i^2}$ is used to suppress weak background features. α, β and γ are coefficients that control the sensitivity of the filter to the shape measures K_a, K_b and h . α, β and γ are selected experimentally: α and β are set to 0.5, γ is set to half of the maximum norm of matrix W in the image to suppress background voxels.

Our filter is effective at enhancing elongated features and attenuating irrelevant objects. Figure 2 (c) shows that our geometric filter enhances the microtubules successfully based on the wavelet enhanced image.

For the experiments in this work, we use a fitting window size comparable to the diameter of the microtubule, which is about 16 pixels in the image. This allows effective capture of the local tubular properties of the microtubules. Window sizes from $15 \times 15 \times 15$ to $19 \times 19 \times 19$ all work well. Since the tube size does not vary much, we do not need to adjust the fitting scale.

In practice, we found that the microtubules can always be enhanced to a satisfying level by applying the filter in two sweeps. The first sweep enhances the overall microtubule and lifts the intensity value in the inner region, while the second sweep produces sharper response.

3) *Modified Coherence-Enhancing Diffusion*: The last step of the enhancement is to complete the interruptions along the microtubules enhanced in previous steps. The key point is to smooth along the prominent direction but not across them. This leads us to CED [26] governed by the following equation:

$$I_t = \text{div}(\mathbf{D}\nabla I) \quad (11)$$

where div is the divergence operator, ∇ is the gradient operator. The diffusion tensor \mathbf{D} is usually defined to have the same eigenvectors as those of the gradient structure tensor

[27], but with adaptive eigenvalues that represent the diffusion strength in each principal directions. The anisotropic diffusion (11) is solved using the finite difference method as described in [26].

A weakness of the above CED and their variations [19][28] is that their coherence measurement is not independent of intensity scalings. This will result in inappropriate diffusion of features with strong contrast but not necessarily anisotropic. We therefore modify the coherence-enhancing diffusion with the local shape measurement described in Section II-A.2. Specifically, we use the response of the filter defined in (10) to adapt the first eigenvalue in the diffusion tensor

$$d_1 = \begin{cases} a & \text{if } \lambda_1 = \lambda_2 = \lambda_3 = 0 \\ a + (1 - a)L & \text{else} \end{cases} \quad (12)$$

where a is a small constant value, chosen as 0.01 in our application, to ensure the positive definiteness of the diffusion tensor. λ_i s are the eigenvalues of the matrix W and L is defined by (10). The last two eigenvalues of the diffusion tensor are designed in the same way as in (13)

$$d_2 = d_3 = a \quad (13)$$

The new diffusion tensor is then constructed by adapting to the principal direction defined by the eigenvectors v_1, v_2, v_3 of the Weingarten matrix W

$$\mathbf{D} = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3) \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \end{pmatrix} \quad (14)$$

Figure 3 compares our modified CED and the conventional CED. Our modified CED is shown to be more suitable for enhancing anisotropic features in images with large contrast variations, since the shape measurement L is insensitive to intensity scaling. Figure 2 (d) demonstrates the completion of minor interruptions along the microtubules with our modified CED method.

B. Microtubule Localization

Since most of the fine microtubule boundary features can be distinguished meaningfully only in a constrained volume, localization is indispensable for the subsequent segmentation. We localize the microtubule in the enhanced volume with a centerline extraction algorithm adapted for tubular structures. Our method is similar to [29], but it requires neither seeding points nor given tracing directions.

Specifically, the centerlines are extracted in the enhanced volume as follows:

1) Evaluate each foreground voxel against its eight neighbors in the local normal plane defined in 3D by the eigenvectors v_2 and v_3 . If the voxel is a local maxima in the normal plane, then it is recorded as a candidate centerline voxel.

2) Form the centerlines by connecting the neighboring local maxima voxles. Only long centerlines are used to localize the microtubules and centerlines too short are discarded. In practice, we choose the discarding threshold as half of the length of the longest centerline in the volume.

Figure 2 (e) illustrates the extracted centerlines that are finally used to determine the sub-volume containing each individual microtubule.

C. Segmentation with Active Shape Model

Segmentation is performed for each microtubule in the localized sub-volumes around the extracted centerlines. To deal with the fuzzy boundaries and overlapping materials, we adapt ASM method for robust segmentation. Because the construction of 3D statistical shape model is difficult, we decompose a microtubule's shape into cross-sectional contour shape and longitudinal shape. The cross-sectional contour is modeled with ASM. The longitudinal smoothness constraint is imposed using Kalman filtering.

In this section, we first introduce the image feature enhancement to enable ASM shape searching. We then describe the statistical modeling of microtubule cross sectional contours, followed with our other improvements on ASM method.

1) *Contour Feature Enhancement*: The conventional gray-level appearance model [17] of ASM is not suitable for our application since the microtubule cross-sectional boundaries are highly blurred and are often corrupted by clutters. Instead, we use 3D local surface enhancement to exploit the saliency of microtubule wall along the longitudinal direction.

With the Weingarten matrix and its eigenvalues calculated in the same way as in Section II-A.2, for a voxel on the microtubule wall, ideally it can be assumed that $|\lambda_1|$ is small, $|\lambda_2|$ is slightly larger than $|\lambda_1|$ while $|\lambda_3|$ is significantly larger than $|\lambda_1|$ and $|\lambda_2|$. However, since the wall surface can be locally flat and thus λ_1 and λ_2 are approximately the same, assuming $\lambda_2 \geq \lambda_1$ will produce bias towards surface of high curvature. Therefore, we design the following equation to estimate a voxel's likelihood of being on a bright surface morphology

$$F = \begin{cases} 0 & \text{if } \lambda_2 > 0 \text{ or } \lambda_3 > 0 \\ \exp(-\frac{K_c^2}{2\alpha^2})(1 - \exp(-\frac{h^2}{2\gamma^2})) & \text{else} \end{cases} \quad (15)$$

where $K_c = \frac{|\lambda_1|}{|\lambda_3|}$. α , γ and h have the same meaning as in (10).

Figure 4 right column shows that the surface-enhancing filter (15) enhances the microtubule contour significantly, while the original boundaries can barely be seen.

2) *Modeling Microtubule Contour with Point Distribution Model*: Even given the enhanced image, a shape model is still necessary to overcome difficulties due to clutters and missing data. The point distribution model is built from a set of training shapes, each represented by a set of n labeled landmark points. The training shapes must be first aligned as follows

- 1) Uniformly sample each training shape using a B-Spline to produce the same number of boundary points for all shapes.
- 2) Choose the first shape as the initial target shape.
- 3) For each shape do
 - a) Choose an initial set of correspondences between the current shape and the target shape. Use the

Procrustes method [30] to align them and calculate the sum of the squared distances.

- b) Repeat (a) for all possible correspondences and choose the one with the minimum sum of the squared distances as the alignment result.
- 4) Update the target shape with the mean of the current aligned shapes, repeat Step 3 until the alignment error over the whole training set converges.

Once a set of N aligned shapes are available, the mean shape and variability can be found as follows:

Let \mathbf{x}_i be a vector describing the n points of the sample shape i in the training set:

$$\mathbf{x}_i = (x_{i0}, y_{i0}, x_{i1}, y_{i1}, \dots, x_{in-1}, y_{in-1})^T$$

where (x_{ij}, y_{ij}) is the point j in sample shape i . The mean shape $\bar{\mathbf{x}}$ is calculated using

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad (16)$$

The modes of shape variations can be found by applying principal component analysis to the deviations from the mean as follows. For each shape in the training set, its deviation from the mean is calculated:

$$d\mathbf{x}_i = \mathbf{x}_i - \bar{\mathbf{x}} \quad (17)$$

The covariance matrix \mathbf{C} is then calculated using:

$$\mathbf{C} = \frac{1}{N-1} \sum_{i=1}^N d\mathbf{x}_i d\mathbf{x}_i^T \quad (18)$$

The modes of variations of the points of the shape are described by \mathbf{p}_k , the unit eigenvectors of \mathbf{C} such that

$$\mathbf{C}\mathbf{p}_k = e_k\mathbf{p}_k \quad (19)$$

where e_k is the k 'th eigenvalue of \mathbf{C} and $e_k \geq e_{k+1}$

$$\mathbf{p}_k^T \mathbf{p}_k = 1 \quad (20)$$

Then any shape in the training set can be approximated using the mean shape and a weighted sum of these deviations obtained from the first m modes

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \quad (21)$$

where \mathbf{P} is a $2n \times m$ matrix whose columns are the unit vectors \mathbf{p}_k , and \mathbf{b} is a vector of weights for the eigenvectors. This equations allow the generations of new shape examples by varying the parameters in \mathbf{b} within suitable limits so that the new shape will be similar to those in the training set.

In our application, the statistic shape model was trained with 12 sample microtubules manually segmented from three large data sets, which are shown in Section III. 50 contours are drawn from each microtubule, each with about 40 boundary points re-sampled up to the same number as previously described. 90 percent shape variation is captured by the statistical model in the first 6 modes.

3) *Shape Searching with Adaptive Weight*: The ASM search involves finding the pose and shape parameters that cause the model to coincide with the structures of interest in the image. In our case, finding the microtubule boundary \mathbf{X} in the image is to search along the normals of the model in the cross sectional slices for the points with maximum boundary strength as defined by (15).

The ASM searching matches a shape in model space to a shape in image space by rotating, translation and scaling

$$\mathbf{z} = T(\mathbf{x}; \Omega) = \tau R(\theta)\mathbf{x} + \mathbf{M}\mathbf{t} \quad (22)$$

where T is the transformation, \mathbf{x} is the shape in the model space, \mathbf{z} is the transformed shape, $\Omega = (\tau, \theta, \mathbf{t})$ is the transformation parameter set, τ is the scale factor, $\mathbf{t} = [t_x, t_y]^t$ is the translation vector, $R(\theta)$ is the matrix rotating \mathbf{x} by angle θ and

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 1 & \dots & 0 & 1 & 0 \\ 0 & 1 & 0 & \dots & 1 & 0 & 1 \end{bmatrix}^t$$

that maps the translation to all the points of the model.

In order to obtain reliable segmentation in presence of missing or corrupted features along the microtubule boundary, we improve the ASM searching procedure by relating the shape matching error with a weight matrix defined with the boundary strength

$$E(\mathbf{X}, \mathbf{z}) = E(\mathbf{X}, \mathbf{x}; \Omega) = (\mathbf{X} - T(\mathbf{x}; \Omega))^T \mathbf{W}(\mathbf{X} - T(\mathbf{x}; \Omega)) \quad (23)$$

where E is the matching error, \mathbf{W} is a diagonal matrix with diagonal element i as

$$w_i = \frac{F_i}{\sum_{i=1}^n F_i} \quad (24)$$

where F_i is the boundary strength (15) at boundary point i .

This adjusts the weights adaptively based on the feature saliency at the boundary points. This has practical significance in our problem, since the microtubule boundaries are subject to defect or missing data due to the limitation of the electron tomography imaging process.

The transformation with appropriate parameters is then found by minimizing the matching error

$$\Omega^* = \arg \min_{\Omega} E(\mathbf{X}, \mathbf{x}; \Omega) \quad (25)$$

Once the appropriate transformation is found, the shape in image space is projected to model space and parameterized as

$$\mathbf{b} = \mathbf{P}^T (T^{-1}(\mathbf{z}; \Omega) - \bar{\mathbf{x}}) \quad (26)$$

This data driven vector \mathbf{b} is then constrained to the limits derived from the parameter distributions at training stage, for instance, $-3\sqrt{e_i} \leq b_i \leq +3\sqrt{e_i}$. The regularized \mathbf{b} is then substituted into (21) and (22) to obtain the constrained shape as the segmentation result. In this way, the model is allowed to search for valid shape only.

Initialized with the mean shape and the mean scale, the above procedure is performed iteratively to refine the shape

and pose parameters to find a better match between the model and the structures in the image.

4) *Integrating ASM and Kalman Filtering*: Since the microtubule boundary can be easily corrupted by noise or clutters, segmentation in a single slice may not be accurate. We therefore utilize the longitudinal smoothness along the microtubules by improving ASM with Kalman filtering, assuming that the shape variation between neighboring slices is small. The Kalman filtering has been combined with ASM for object tracking [31][32] and has been used for centerline detection to aid the segmentation of tubular object [33]. However, the Kalman filtering has not been combined with ASM for the segmentation of tubular object. In the following paragraphs, we formulate our combination of the Kalman filtering and ASM.

In relating microtubule contours across longitudinal slices, we assume that no rotation is involved between neighboring contours. Therefore, we can simplify (22) by assuming an identity rotation and substituting shape x by the sum of mean shape and variations

$$\mathbf{z} = \tau(\bar{\mathbf{x}} + \mathbf{P}\mathbf{b}) + \mathbf{M}\mathbf{t} \quad (27)$$

which can be further written as

$$\mathbf{z} = \begin{bmatrix} \bar{\mathbf{x}} & \mathbf{P} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \tau \\ \tau\mathbf{b} \\ \mathbf{t} \end{bmatrix} \quad (28)$$

The measurement model for the Kalman filter is thus

$$\mathbf{z}_k = \mathbf{H}\mathbf{s}_k + \mathbf{v}_k \quad (29)$$

where $\mathbf{H} = [\bar{\mathbf{x}}\mathbf{P}\mathbf{M}]$, $\mathbf{s}_k = [\tau \ \tau\mathbf{b}^t \ \mathbf{t}^t]^t$ is the state vector consisting of the shape parameters, and \mathbf{v}_k represents measurement uncertainty normally distributed as $\mathbf{v}_k \sim N(0, \mathbf{R})$.

The state model then relates the shape parameter obtained in image slice $k+1$ with that in the current slice k :

$$\mathbf{s}_{k+1} = \Phi\mathbf{s}_k + \mathbf{w}_k \quad (30)$$

where the state transition matrix Φ is an identity matrix in this case, and \mathbf{w}_k represents system perturbation normally distributed as $\mathbf{w}_k \sim N(0, \mathbf{Q})$.

With the derived state model and measurement model, the Kalman filtering consists of state prediction and state updating. The state prediction involves state projection and error covariance estimation as

$$\mathbf{s}_{k+1}^- = \Phi\mathbf{s}_k \quad (31)$$

$$\Sigma_{k+1}^- = \Phi\Sigma_k\Phi^T + \mathbf{Q} \quad (32)$$

where \mathbf{s}_{k+1}^- denotes the predicted state consisting of the shape parameters, Σ_k denotes the covariance matrix at iteration k , and Σ_{k+1} denotes the estimated covariance for iteration $k+1$.

The Kalman gain is computed as

$$\mathbf{G}_{k+1} = \Sigma_{k+1}^- \mathbf{H}^T (\mathbf{H}\Sigma_{k+1}^- \mathbf{H}^T + \mathbf{R})^{-1} \quad (33)$$

which is a weighting factor to determine the contribution of measurement z_{k+1} and prediction s_{k+1}^- to the posterior state estimate s_{k+1} .

In state updating, the final state estimate is then obtained by combining the predicted state and the new measurement z_{k+1}

$$\mathbf{s}_{k+1} = \mathbf{s}_{k+1}^- + \mathbf{G}_{k+1}(\mathbf{z}_{k+1} - \mathbf{H}\mathbf{s}_{k+1}^-) \quad (34)$$

The posterior error covariance estimate Σ_{k+1} is computed as well

$$\Sigma_{k+1} = (\mathbf{I} - \mathbf{G}_{k+1}\mathbf{H})\Sigma_{k+1}^- \quad (35)$$

The Kalman filtering from (31) to (35) is an recursive process across the cross-sectional slices. At each iteration, the shape parameters can be solved from the posterior state estimation \mathbf{s}_{k+1} and the segmentation boundary in spatial domain can be obtained easily.

The effect of Kalman filtering on segmentation is demonstrated in Figure 5. The results indicate that the segmentation with Kalman filtering is less sensitive to missing boundary or clutters in individual slice. The possible limitation of the ASM with Kalman filtering is that it cannot capture large local shape variations, which are unlikely to exist for our application.

5) *Initialization of ASM*: To initialize the ASM close to the desired object, we use Hough transform to detect the rough microtubule contours. The Hough transform is applied to the slices perpendicular to the estimated microtubule axis. The detected circles are correlated across the slices. The positions where circles are detected consistently are chosen to initialize the ASM. In the same manner, the Hough transform also decides the stopping locations for ASM by detecting the ends of the microtubule, beyond which the detected circles are randomly scattered. Figure 6 shows the rough detection of the microtubule cross-sectional contours with Hough transform.

III. RESULTS AND EVALUATION

A. Experimental Results

In this work, we have applied the above methods to more than 50 volume data sets containing multiple microtubules. In this section, we present sample enhancement and segmentation results. The original kinetochore tomography data sets are $300 \times 300 \times 60$ 3D volumes as shown in Figure 7 (a) and Figure 9. In Figure 7 (a), the microtubules are not easily visible in the projection of the raw volume since they are buried in noise and surrounded by numerous cellular materials. Given the specific tomography imaging process, we can assume that the axes of the microtubules are parallel to the $x - y$ plane. Therefore, we process the 3D volume in a slice-by-slice manner using 2D wavelet transform. To achieve averaging over rotation, each slice is rotated over eight angles only in the $x - y$ plane. The processed slices are stacked together to form 3D volume, which is then further processed with local shape filtering and the modified CED. Figure 7 (b) shows that the microtubules stand out obviously after the enhancement. Figure 8 shows the selected 2D slices of the enhanced 3D volume. The enhancement results on the other two data sets in Figure 9 demonstrate that our method can enhance the

tomography volume with various image quality and contrast variations.

Segmentation is then performed around each localized microtubule with the improved ASM method. The automatic segmentation results and the manual results are compared in Figure 10. For illustration purpose, the images were magnified. Obviously, the automatic results have much more smooth morphologies, due to the consistency and robustness of the automated method. In contrast, the manual method produces rough surface morphology, which is less plausible biologically.

Figure 11 presents more 3D segmentation results for microtubules with various surface morphologies. It demonstrates that the automatic method can capture the morphological characteristics of the microtubules in spite of the cluttered image.

B. Quantitative Evaluation

To quantitatively evaluate the performance of our automated approach, we first use 2D contour overlap ratio to compare the manual results and the automatic results, which is defined as

$$OL = \frac{A \cap B}{A \cup B} \quad (36)$$

with A and B as the areas enclosed by two contours. A large contour overlap ratio means that the manual segmentation and automatic segmentation are in good agreement with each other.

The other important measures are the centerline smoothness and the narrowing and widening of microtubules. The centerline smoothness is approximated by the contour centroid change. To describe the narrowing and widening, we normalize the contour areas of each microtubule by its mean contour area and then approximate the contours as circular shapes.

Figure 12 is the quantitative evaluation of typical segmentation results based on a relatively high quality data set as in Figure 10. In Figure 12 (a), the overlap ratio, with mean value as 0.85, indicates that the automatic result and the manual result are in good agreement with each other. In Figure 12 (b) and (c), the centroid change trends for automatic result and manual result appear to be similar to each other, with the mean value of the centroid deviation as 0.5 pixel between them. In Figure 12 (d), both results agree with each other in describing the narrowing and widening trend, showing four obvious local maximums.

The same evaluation has also been done on low contrast data set. In this case, the mean overlap ratio decreases to 0.7 and the mean centroid deviation increases to 3.5 pixels between the automatic result and the manual result. While the automatic segmentation results have smooth centroid change across the slices, the manual results have sudden centroid change across the slices, which is not plausible biologically.

The above analysis indicates that the automatic segmentation is consistent and reliable for both low quality data and good quality data, while performance of manual segmentation becomes worse on low quality data. With the current limited number of biological samples, the automatic segmentation has so far been able to characterize the gross microtubule structural features. In addition, for all the microtubules used in the

experiments, the segmentation performance under leave-one-out scheme coincides with those under leave-all-in scheme. This is because the average shape of the contours of one microtubule is similar to that of another microtubule so that adding even more contours does not proportionally introduce new shape information. This indicates that the training data is sufficient to represent all 2D contour shapes.

IV. SUMMARY AND CONCLUSIONS

We have presented an automatic approach to extracting microtubules from electron tomography volume. Our approach consists of volume enhancement at coarse scale, microtubule localization and microtubule boundary segmentation as fine features. In volume enhancement, an anisotropic invariant wavelet transform is first applied to effectively enhance the elongated features, followed by a 3D tube-enhancing filter to capture the local geometric properties of the tubular structure. The enhancement ends with a modified CED to complete the interruptions along the microtubules without affecting other regions in the images. Our innovations in volume enhancement are that we systematically combine transform domain technique with spatial domain techniques and improve each of them as follows for enhancing microtubule features. We make the wavelet sensitive to elongated features in noise by using anisotropic basis and invariant transform. We design the tube-enhancing filter based on curvatures of hyper-surface to produce more distinguishing response for tubular structures. We integrate the CED with the shape measurement provided by the tube-enhancing filter so that the CED can complete interruptions along microtubules despite contrast variations. Experiments show that the proposed enhancement procedure can effectively accentuate the microtubules while simultaneously de-emphasizing irrelevant structures. After enhancement, the microtubules are localized by applying a centerline extraction algorithm to the enhanced volume. The microtubule boundary segmentation is finally performed around each localized microtubule with an ASM method, which we have novelly modified and extended in three aspects. We first replace the conventional gray-level appearance model in ASM with new contour features obtained using 3-D local surface enhancement, which more reliably characterizes the microtubule boundary. We also improve the ASM shape searching by relating the boundary strength model with the weight matrix of the ASM searching error. We then integrate the ASM with Kalman filtering to utilize the smoothness constraint along the longitudinal direction of the microtubules, which makes the segmentation robust against missing boundaries and outliers. Experimental results demonstrate that the automatically segmented 3D microtubules are of comparable quality to those by manual process but use only a fraction of the time of the latter.

REFERENCES

- [1] B. F. McEwen, R. M. Barnard, T. Portuese, and C.-E. Hsieh, "Using electron tomography to determine the interaction between mammalian kinetochores and microtubules," *Proc. Int. Congress Elect. Microsc. 15*, pp. 179–180, 2002.
- [2] J. R. Kremer, D. N. Mastrorarde, and J. R. McIntosh, "Computer visualization of three-dimensional image data using imod," *Journal of Structural Biology*, no. 116, pp. 71–76, 1996.
- [3] W. Shen, P. Zhang, D. Germain, T. Rouault, and S. Subramaniam, "A template-propagation method for segmentation of filamentous structures in electron tomograms," in *Proceedings of the 2nd IEEE International Symposium on Biomedical Imaging*, 2004, pp. 1–4.
- [4] A. Bartesaghi, G. Sapiro, S. Lee, J. Lefman, S. Wahl, S. Subramaniam, and J. Orenstein, "A new approach for 3D segmentation of cellular tomograms obtained using three-dimensional electron microscopy," in *Proceedings of the 2nd IEEE International Symposium on Biomedical Imaging*, 2004, pp. 5–8.
- [5] C. Bajaj, Z. Yu, and M. Auer, "Volumetric feature extraction and visualization of tomographic molecular imaging," *Journal of Structural Biology*, vol. 144, no. 1-2, pp. 132–143, 2003.
- [6] A. S. Frangakis and R. Hegrel, "Segmentation of two and three dimensional data from electron microscopy using eigenvector analysis," *Journal of Structural Biology*, vol. 138, no. 1, pp. 105–113, 2002.
- [7] N. Volkmann, "A novel three dimensional variant of the watershed transform for segmentation of electron display maps," *Journal of Structural Biology*, vol. 138, no. 1, pp. 123–129, 2002.
- [8] A. Frangi, W. Niessen, R. Hoogeveen, T. van Walsum, and M. Viergever, "Model-based quantitation of 3D magnetic resonance angiographic images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 946–956, 1999.
- [9] P. J. Yim, J. J. Cebral, R. Mullick, H. B. Marcos, and P. L. Choyke, "Vessel surface reconstruction with a tubular deformable model," *IEEE Transactions on Medical Imaging*, vol. 20, no. 12, pp. 1411–1421, 2001.
- [10] T. Deschamps and L. Cohen, "Fast extraction of tubular and tree 3d surfaces with front propagation methods," *Proceedings. 16th International Conference on Pattern Recognition*, vol. 1, pp. 731–734, 2002.
- [11] L. Lorigo, O. Faugeras, W. Grimson, R. Keriven, R. Kikinis, A. Nabavi, and C.-F. Westin, "Codimension-two geodesic active contours for the segmentation of tubular structures," *Proceedings 2000 IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 444–451, 2000.
- [12] C. M. van Bommel, L. Spreuwers, M. Viergever, and W. Niessen, "Level-set-based artery-vein separation in blood pool agent CE-MR angiograms," *IEEE Transactions on Medical Imaging*, vol. 22, no. 10, pp. 1224–1234, 2003.
- [13] A. Vasilevskiy and K. Siddiqi, "Flux maximizing geometric flows," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1565–1578, 2002.
- [14] M. Leventon, W. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2000, pp. 316–323.
- [15] A. Tsai, A. Yezzi, W. Wells, C. Tempny, D. Tucker, A. Fan, A. Grimson, and A. Willsky, "Model-based curve evolution technique for image segmentation," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2001, pp. 363–468.
- [16] M. Rousson and N. Paragios, "Shape priors for level set representations," in *Proceedings of European Conference in Computer Vision*, vol. 2, 2002, pp. 78–93.
- [17] T. F. Cootes, C. J. Taylor, and D. H. Cooper, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.
- [18] A. Stoschek, T. P. Y. Yu, and R. Hegerl, "Denoising of electron tomographic reconstructions from biological specimen using multidimensional multiscale transforms," *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2793–2796, 1997.
- [19] A. S. Frangakis, A. Stoschek, and R. Hegerl, "Wavelet transform filtering and nonlinear anisotropic diffusion assessed for signal reconstruction performance on multidimensional biomedical data," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 2, pp. 213–222, 2001.
- [20] M. Harlow, D. Ress, A. Stoschek, R. Marshall, and U. McMahan, "The architecture of active zone material at the frog's neuromuscular junction," *Nature*, vol. 409, no. 6819, pp. 479–84, 2001.
- [21] G. Beylkin, "On the representation of operators in bases of compactly supported wavelets," *SIAM J. Numer. Anal.*, vol. 29, pp. 1716–1740, 1992.
- [22] T. P. Y. Yu, A. Stoschek, and D. L. Donoho, "Translation- and direction-invariant denoising of 2d and 3d images: experience and algorithms," *Proceedings of the SPIE, Wavelet Applications in Signal and Image Processing IV*, vol. 2825, pp. 608–619, 1996.
- [23] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," *Lecture Notes in Computer Science*, vol. 1496, pp. 130–137, 1998.

- [24] O. Monga and S. Benayoun, "Using partial derivatives of 3d images to extract typical surface features," *Proceedings of the Third Annual Conference of AI, Simulation, and Planning in High Autonomy Systems*, pp. 225–236, 1992.
- [25] M. Jiang, Q. Ji, and B. F. McEwen, "Enhancement of microtubules in em tomography," in *Proceedings of the 2nd IEEE International Symposium on Biomedical Imaging*, 2004, pp. 1123–1126.
- [26] J. Weickert, "Coherence-enhancing diffusion filtering," *International Journal of Computer Vision*, vol. 31, no. 2, pp. 111–127, 1999.
- [27] J. Bigun and G. Granlund, "Optimal orientation detection of linear symmetry," *Proceedings of the First Intl. Conf. on Comp. Vision*, pp. 433–438, 1987.
- [28] Y.-P. Wang, Q. Wu, K. Castleman, and Z. Xiong, "Flux-based anisotropic diffusion applied to enhancement of 3-D angiogram," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1440–1442, 2002.
- [29] S. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction," *IEEE Transactions on Medical Imaging*, vol. 21, no. 2, pp. 61–75, 2002.
- [30] I. L. Dryden and K.V.Mardia, *Statistical shape analysis*. John Wiley, Chichester, 1998.
- [31] A. Baumberg and D. Hogg, "An efficient method for contour tracking using active shape models," *Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, pp. 194–199, 1994.
- [32] C. J. T. G. J. Edwards and T. F. Cootes, "Learning to identify and track faces in image sequences," in *Proceedings of the 8th British Machine Vision Conference*, vol. 1, 197.
- [33] T. Behrens, K. Rohr, and H. Stiehl, "Robust segmentation of tubular structures in 3-D medical images by parametric object detection and tracking," *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 33, no. 4, pp. 554–561, 2003.