DEFORMABLE TRELLISES ON FACTOR GRAPHS FOR ROBUST MICROTUBULE TRACKING IN CLUTTER

Rahul Kidambi , Min-Chi Shih and Kenneth Rose

{*rkidambi,minchi_shih,rose*}@ece.ucsb.edu
Department of Electrical and Computer Engineering
University of California Santa Barbara
California 93106 USA

ABSTRACT

A main challenge in microtubule tracking is due to clutter, or the presence of many similar intersecting structures. This paper proposes a two-layered probabilistic formulation which has at its foundation a factor graph serving as a multi-label inference engine designed to provide distinction between open contours of interest and other microtubules or noise. The second layer is a deformable trellis defined over the resulting label probability map, where a *Hidden Markov Model* (HMM) is employed to determine the most probable current location of the microtubule body. The overall framework enjoys the "best of both worlds" - the factor graph is effective in discriminating between contours of interest and others that exhibit similar statistical properties, while the deformable trellis with its HMM offer accurate modeling of microtubule dynamics in terms of growth and shortening, as well as precise body tracing, accounting for prior information, all within a principled Bayesian framework. Simulation results provide evidence that the proposed approach outperforms existing techniques.

***Index Terms*—** Microtubules, Factor Graphs, Deformable Trellis, Belief Propagation, Hidden Markov Models.

1. INTRODUCTION

Object tracking is one of the fundamental problems in vision, and is important in bio-image analysis. Microtubule (MT) tracking is necessary for understanding MT dynamics, which is the focus of extensive biological research, has been conjectured to have a significant role in critical cell functions and processes, and correspondingly have potential implication in various diseases [1], [2]. MT tracking poses major challenges that go beyond those typically encountered in object tracking in video sequences, including: (a) low signal to noise ratio of MT microscopy images, (b) considerable clutter due to frequent intersections with other MTs. The focus of this paper is on powerful new approaches to effectively track MTs despite the adverse conditions (see, example, Figure 1). The tracking results we present are for MTs imaged by fluorescence microscopy.

2. PRIOR WORK

Earlier approaches to the problem of tracking curvilinear structures and their dynamic analysis include approaches based on active contours [4], and Kalman filtering [5], [6]. In [7], significant tracking performance improvement was obtained by employing an HMM approach, which has natural capability to probabilistically account for growth and shortening of the MT, and is computationally efficient due to the applicability of the Viterbi algorithm. However, while the method is highly effective in tracking an MT in noise, its shortcomings emerge when we encounter clutter due to intersection with other MTs that exhibit similar statistical attributes.

This work was supported by NSF-III-0808772. The authors thank Prof. Stu Fienstein, Prof. Les Wilson and Elmer Guzman for the datasets provided for the experiments in the paper.

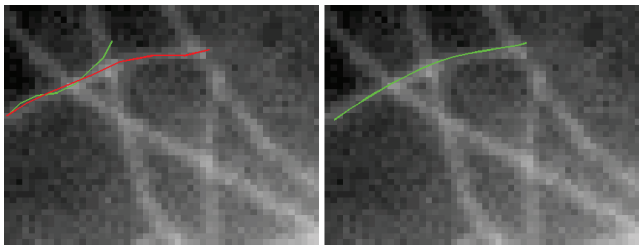


Fig. 1. A specific example comparing output of [3] (on the left) versus the proposed approach (on the right). The ground truth is annotated in red, while green represents the tracker output

An alternative approach involves belief propagation on a factor graph [3]. It formulates tracking in clutter as a multi-label inference problem, which may be resolved by belief propagation. The outcome is a label probability map in which MT tracing is performed using a constrained maximization method from [8]. This method, while providing the ability to discriminate between the contour of interest from other intersecting structures, does not have an intrinsic ability to model growth and shortening and hence to effectively model the MT dynamics. The motivation for the method we propose herein stems from the realization of the complementary capabilities inherent to the HMM and factor graph approaches.

3. THE PROPOSED FRAMEWORK

The proposed framework consists of two layers. A multi-label factor graph defined as in [3] forms the first layer. Posterior probabilities associating each node of the factor graph with the possible labels are obtained using belief propagation [9]. The second layer involves a deformable trellis that is overlaid on the *probability map* obtained in the first layer. The trellis is positioned about the MT contour of the previous frame, which allows for deformation of the contour in the current frame, as well as length changes. An HMM is defined over the trellis to compute the maximum a posteriori estimate of the current contour location via the Viterbi algorithm.

3.1. Factor Graph and Belief Propagation

A factor graph is defined over the image, where an image pixel is associated with a 'variable' node of the factor graph. Each variable node can take value in a set of three labels - contour of interest l_{CoI}, other contour l_{Oc} and background l_{Bg} [3]. There are two types of factor nodes in this factor graph, which are described next.

A factor node that is defined on a single variable node i, is commonly referred to as a unary potential $\beta(z_i)$, which measure the prior probability that the node belongs to the class or label z_i. These probabilities are a diffused version of posteriors from the previous frame (and initialized by the user in the first frame of the sequence). Diffusion is carried out in consecutive frames using distance transforms.

The second type of factor node is the binary potential $\alpha(z_j, z_k)$, associated with two connected variable nodes j, k, i.e., neighboring nodes. The binary potentials quantify two aspects of the model: (i) Label transition probabilities - the probability of a node assuming a particular label conditioned on the label of its neighbor. Such dependence can be extracted from a training set of images where the foreground and background classification is known. (ii) The data likelihood term - which quantifies the probability of the observation measured between two adjacent nodes, conditioned on the node labels. This non-standard aspect is important in the context of tracking curvilinear structures.

Observation features used to compute likelihood in the factor graph are the ridge features, obtained by filtering the image with a second order derivative gaussian filter. Observations computed between node pairs are obtained by averaging ridge features of the two nodes, along the normal to the line joining them.

To obtain statistics for computing the binary potential, a training image set is used, where MTs and background had been labeled. In a training image, ridge features are measured for different pairs of labels. Observations for each class (each possible pair of labels) are modeled as Gaussian densities, whose means and variances are evaluated. Any observation in the test image can then be evaluated as evidence via the likelihood associating it with different classes.

With the above definition of potentials, the expression for the joint label probability $p(\underline{\mathbf{z}})$ of the set of M pixels in the image is:

$$p(\underline{\mathbf{z}}) = \frac{\mathbf{1}}{\mathbf{Z}} \prod_{\mathbf{k=1}}^{\mathbf{M}} \beta(\mathbf{z_k}) \prod_{\mathbf{j=1, j \in \mathcal{N}(k)}}^{\mathbf{N'}} \alpha(\mathbf{z_k}, \mathbf{z_j}) \qquad (1)$$

Where, $N^{'}$ = 4 is the number of neighbors for each pixel, and Z provides normalization to obtain a valid probability distribution.

Once we evaluate the potentials for a given frame, belief propagation is performed on the factor graph, to obtain posterior probabilities for labeling each node of the graph.

3.2. Deformable Trellis and Hidden Markov Model

A deformable trellis is overlaid on the probability map corresponding to the label l_{CoI}. More specifically, the deformable trellis is positioned about the contour estimated for the MT in the previous frame. This trellis enables employing an HMM to optimize the MT contour.

A first important distinction with the deformable trellis of [10] and [7], is that here the trellis (and HMM) is overlaid on, and processes the probability map, rather than the original image itself. The intuition behind this paradigm is that it allows us to leverage the factor graph capability to separate the contours of interest from the clutter of other contours. The probability map l_{CoI} is effectively a highly processed version of the current frame which consists of probabilities associating each node of the factor graph with the contour(s) of interest. We basically propose to use this probability map as the 'observations' for the HMM. Hence, our problem now reduces to finding the best sequence of states along the trellis (positioned about the MT contour of the previous frame), given the observations in the current frame. The deformable trellis with the HMM offer powerful statistical modeling capabilities, including the ability to model MT motion, growth and shortening, in a principled fashion.

In summary, we position a trellis about the MT contour from the previous frame (see Fig. 2). The trellis is constructed by forming equally spaced normals to the initial contour, on which we define states where the new contour may pass. (Note that to avoid graphical clutter Fig. 2 does not show the actual trellis edges, which connect every state on a given normal to adjacent states in the next normal along the trellis). The HMM identifies the most probable sequence of states, i.e., the most probable contour in the current frame. This corresponds to the maximum a posteriori estimate of the MT contour given its location in the previous frame, the model for MT dynamics, and current observations in the form of the probability map produced by the factor graph layer.

3.2.1. Details of the Hidden Markov Model

A Hidden Markov Model is parameterized by (A,B,π), where 'A' is the transition probability matrix, 'B' is the observation or emission matrix, and 'π' is the initial state occupancy vector. π is modeled as a Gibbs distribution, centered around the mid-state $\frac{N+1}{2}$ in a trellis composed on N states, where, N is odd:

$$\pi_i \propto e^{\gamma(i-\frac{N+1}{2})^2} \forall i \in \{1, 2, ...N\}, \qquad (2)$$

where γ is extracted from training data.

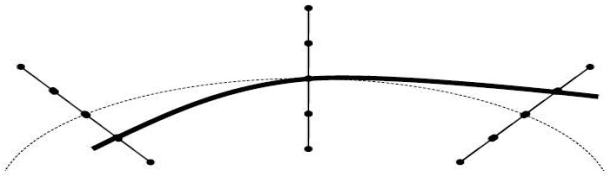


Fig. 2. Trellis overlaid about the MT contour of the previous frame (dotted line), where the most probable sequence of states estimates the actual MT contour of the current frame (solid line)

We incorporate in the HMM auxiliary states defined similar to [7], namely, the begin state, S_b, and end state, S_e. These states enable modeling length contractions in either end of the MT. The end state is modeled as an absorbing state (it is impossible to transition out of it). The begin state is correspondingly an expelling state. So, the trellis is defined with the states $\{s_i : i \in \{1, 2, ...N\}, s_b, s_e\}$.

To obtain a smooth contour post inference in the HMM, the Arc-emission HMM is used, which differs from a conventional state-emission HMM in that its observations depend on state pairs, the current and previous states (Figure 3), rather than on the current single state only as in the state-emission HMM. Decoding the state sequence of an Arc Emission HMM offers more flexibility at the cost of some additional model complexity.

To model the observations which are drawn from the posterior probability map produced by the underlying factor graph, we average the posterior values at the two locations corresponding to the two consecutive states on the trellis, i.e., the two states forming the state

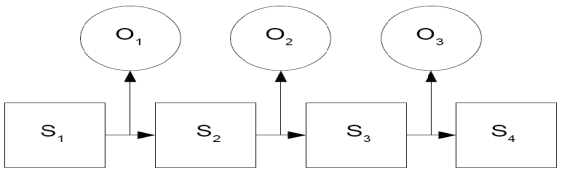


Fig. 3. Graphical Model representation of an Arc Emission HMM. S_i denote states and O_i observations.

pair. So, the observation for the arc connecting the pair of states i,j on adjacent normals of the trellis is:

$$o_{ij} = \frac{p_i + p_j}{2}, \forall i \in \phi_k, \forall j \in \phi_{k+1}, i, j \in \mathcal{E}, \quad (3)$$

where, $\mathcal{E}$ represents the set of connected edges in the trellis, p_i and p_j are the posteriors at positions i and j, ϕ_k are the M normals along the trellis overlaid about the contour inherited from the previous frame.

In order to model the observation emission matrix B, deformable trellises are overlaid on training images with available ground truth (i.e., the true contour is known and the probability map after belief propagation is available). Here the correct contour is simply the sequence of mid-states in the trellis. Features measured along the mid-state sequence represent observations that support the foreground, and features measured between all other state pairs are observations that support the background. For simplicity, in simulations we modeled the emission probabilities for observations corresponding to the foreground and background as two Gaussian densities, whose mean and variance were evaluated from the training set (despite the obvious fact that these features are probabilities themselves and hence vary between 0 and 1). Given these parameters it is straightforward to generate the emission matrix B.

The transition probability matrix A assigns probabilities (or inversely, costs), that govern transitions from any given state to any other state. Such a matrix is composed of elements modeled using a tilted Gibbs distribution, which can be learnt using the fixed point iterative method described in [7].

Once the parameters of the HMM are estimated, the MAP estimate of the state sequence that represents the contour in the given frame is obtained using the Viterbi Algorithm [11]. Thus, with the above construction, lateral deformations of the contour, and length contractions are modeled. Thus, in summary:

$$q^* = \arg\max_{q \in Q} P(q/O_t, \lambda) \quad (4)$$

$$C_t = I(q^*, C_{t-1}), \quad (5)$$

where the optimal state sequence q^* is obtained by Viterbi algorithm on the trellis given the probability map of the current frame, and C_t is the position of the contour in the current frame obtained from the mapping $I(.)$, calculated from q^* which is a sequence of states in the trellis constructed about the previous frame's contour C_{t-1}.

To account for potential growth into new areas of the image, a funnel shaped graph is employed on the output of Viterbi's algorithm, as shown in Figure 4. We choose a set of points along the funnel graph that have features that possess maximal statistical evidence about the presence of the contour, (which must exceed a specified threshold, as evidence for real MT growth). Then, an 'augmented' trellis is constructed over the set of points obtained from Viterbi's decoding, along with the points obtained from the construction of the funnel graph. Inference is performed on this augmented trellis, to obtain the MAP estimate of the state sequence that enable us to estimate the position and length of the contour in the current frame.

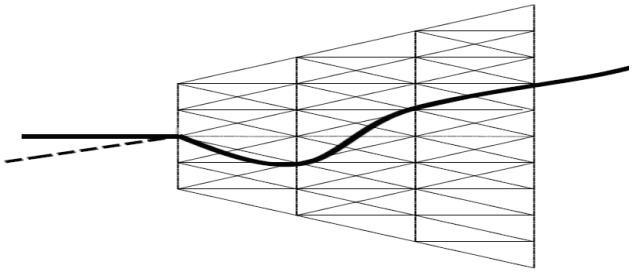


Fig. 4. Overlaying the Funnel Graph in the probability map obtained as the output of *Belief Propagation*, with the contour in the given frame (solid line), the contour inherited from the previous frame (dotted line)

4. RESULTS

In this section, we compare the proposed two-layered method with the HMM approach of [7] and the factor graph approach of [3]. We elaborate on multiple metrics for comparison, in order to provide a measure of precision in the contour estimate, but also account for the fact that major tracking errors may overwhelm contour distance measures. The metrics are: (i) Body distance between the contour estimate and ground truth, averaged on the entire experiment, denoted $d_{overall}$. The body distance is a constrained distance calculation technique requiring dynamic programming for its evaluation, see [7]. (ii) Body distance averaged over a set that excludes "tracking errors", denoted $d_{success}$. We declare tracking error if the body distance exceeds a threshold, $\theta \in \{3, 4, 5, 6\}$ pixels. (iii) Percentage of tracking errors at a given threshold θ. Figures 5 and 6 display some visual results comparing our approach with other approaches.

Table 1 displays the comparison in terms of the 3 measures, where $\theta = 5$, while figure 7 depicts how the percentage of tracking errors varies with θ for the three methods (values are evaluated using 150 annotated frames of highly cluttered data belonging to particularly challenging image sequences, indicating that the method outperforms other approaches in these scenarios, while offering similar performance in other relatively uncluttered sequences). All results show favorable performance gains for the proposed method.

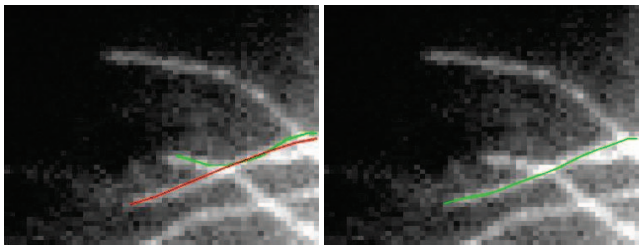


Fig. 5. A specific example comparing output of [3] (on the left) versus the proposed approach (on the right). The ground truth is annotated in red, while green represents the tracker output

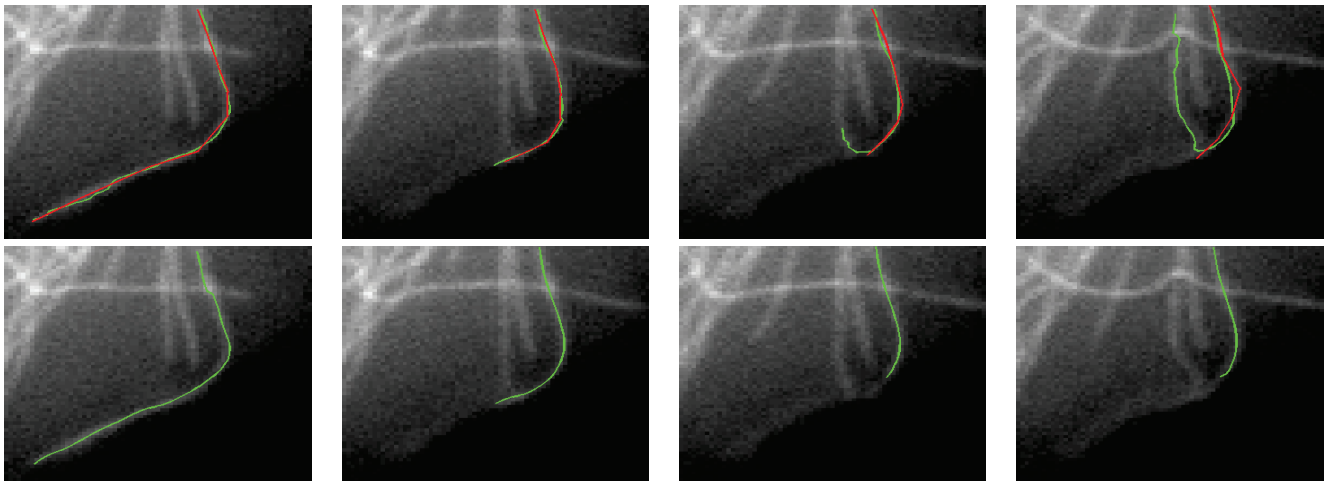


Fig. 6. Comparing results of [7] (top row) with the proposed approach (bottom row) in growth and shortening phases of the Microtubule: ground truth in red, tracker output in green

Table 1. Body-Distance Comparison, error threshold $\theta = 5$

Method	$d_{overall}$	$d_{success}$	% of tracking errors
HMM	7.9628	1.7877	57.89
Factor Graph	3.3624	1.4182	24.34
Proposed	2.2346	1.3560	9.21

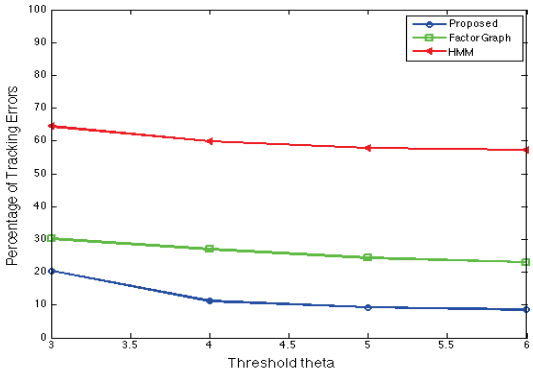


Fig. 7. Tracking error rate versus the threshold θ

5. CONCLUSION

A two-layer probabilistic approach to track cluttered Microtubules is presented. By overlaying a deformable trellis and a corresponding HMM on the probability map generated from inference in an underlying factor graph, an effective tracking mechanism is achieved. The framework enjoys the complementary advantages of two principled subsystems - the factor graph approach, which lends the discriminatory power to distinguish between the contour of interest from other contours, and the deformable trellis enables seamless tracking of length changes of the contour in successive frames. Results, in terms of several criteria, provide evidence that the proposed approach outperforms other techniques in tracking cluttered Microtubules.

6. REFERENCES

[1] B.Alberts et al., "Molecular biology of the cell," *Garland Scientific*, p. 4^{th} Ed., 2002.

[2] S.C.Feinstein and L.Wilson, "Inability of tau to properly regulate neuronal microtubule dynamics: a loss of function mechanism by which tau might mediate neuronal cell death," *Biochimica et Biophysica Acta*, vol. 1739, pp. 268 – 279, 2005.

[3] P. Koulgi, M.E. Sargin, K. Rose, and B.S. Manjunath, "Graphical model-based tracking of curvilinear structures in bio-image sequences," *Proc. IEEE ICPR*, pp. 2596 –2599, 2010.

[4] A. Altinok, M. El-Saban, A.J. Peck, L. Wilson, S.C. Feinstein, B.S. Manjunath, and K. Rose, "Activity analysis in microtubule videos by mixture of hidden markov models," *Proc. IEEE CVPR*, vol. 2, pp. 1662 – 1669, 2006.

[5] R. Toledo, X. Orriols, X. Binefa, P. Radeva, J. Vitria, and J.J. Villanueva, "Tracking elongated structures using statistical snakes," *Proc. IEEE CVPR*, vol. 1, pp. 157–162, 2000.

[6] N. Peterfreund, "Robust tracking of position and velocity with kalman snakes," *IEEE Pattern Analysis Mach. Intell.*, vol. 21, no. 6, pp. 564 –569, 1999.

[7] M.E. Sargin, A. Altinok, B.S. Manjunath, and K. Rose, "Variable length open contour tracking using a deformable trellis," *IEEE Trans. Image Processing*, vol. 20, no. 4, pp. 1023 –1035, 2011.

[8] M.E. Sargin, A. Alttnok, K. Rose, and B.S. Manjunath, "Tracing curvilinear structures in live cell images," *Proc. IEEE ICIP*, pp. 285 –288, 2007.

[9] J.M.Mooij, "libDAI:a free and open source C++ library for discrete approximate inference in graphical models," *JMLR*, pp. 2169–2173, 2010.

[10] M.E. Sargin, A. Altinok, K. Rose, and B.S. Manjunath, "Deformable trellis: open contour tracking in bio-image sequences," *Proc. IEEE ICASSP*, pp. 561 –564, 2008.

[11] L.R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proc. IEEE*, pp. 257 –286, 1989.