

Robust Biological Image Sequence Analysis Using Graph Based Approaches

Diana Delibaltov, S. Karthikeyan, Vignesh Jagadeesh, B.S. Manjunath

Center for Bioimage Informatics, Department of Electrical and Computer Engineering, University of California, Santa Barbara

Abstract—Robust methods for segmentation and tracking are critical for quantitative biology. We give an overview of our recent work on graph based methods for various microscopy image analysis, including tracing over 3-D electron microscopy image stacks, tracking in time-lapse confocal image sequences, and 3D segmentation. We present results on a variety of datasets such as 3-D confocal membrane volumes of the ascidian *Ciona*, electron micrograph stacks from the rabbit retina, and bright field microscopy time sequence data from mouse retina.

I. INTRODUCTION

Bioimage informatics refers to managing, processing, analysing, annotating, navigating and understanding biological image collections. Progress in biology research is hindered by the need for manual annotation, segmentation and tracking, which are often intractable.

For example, the study of morphogenetics at the Smith Lab at UCSB uses 3-D confocal microscopy membrane volumes of the ascidian *Ciona*, for which the individual segmentation of a couple of thousands of cells in 3-D is required [1]. Morphogenetics is the study of cell and tissue development and the mechanisms that drive these changes (cell motility, division and shape changes). The ascidian *Ciona* is a very good candidate for the study of morphogenetics, since it has a small, compact embryo. This makes it possible to obtain a global image of the embryo at sufficient detail. The notochord tissue of the ascidian, highlighted in blue in Fig. 1, has a total of 40 cells. The equivalent tissue in a vertebrate contains hundreds of thousands of cells. Nevertheless, the entire ascidian embryo contains six tissues and a couple of thousands of cells, thus making 3-D manual segmentation an intractable problem. Furthermore, this dataset is challenging for automated algorithms because of the varying membrane intensity and the faint presence of other organelles such as nuclei (Fig. 2(a)).

The second dataset we consider is the time-lapse image sequences of melanosomes obtained using bright field microscopy from Williams Lab at UCLA [2]. Melanosomes are organelles present in the retinal pigment epithelium in the retina layer, and provide the tissue with color and photoprotection. Understanding melanosome dynamics provides insights to the cause of genetic diseases. The melanosomes are generally densely packed (Fig. 2(b)) and the coarse time sampling of their motion, imaging noise and out of focus moving structures present challenges to their detection and tracking.

The third dataset we consider is from the Marc lab at

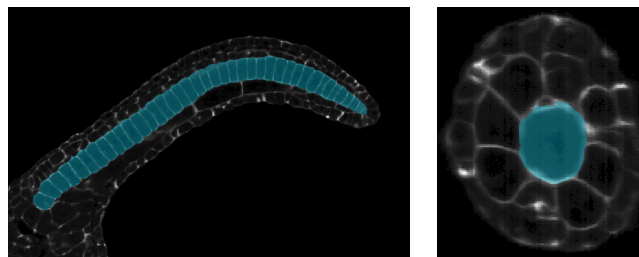


Fig. 1. Left: Original confocal section of cell membrane microscopy image of the ascidian *Ciona*. Right: Reconstructed cross-section. The notochord tissue is marked in blue.

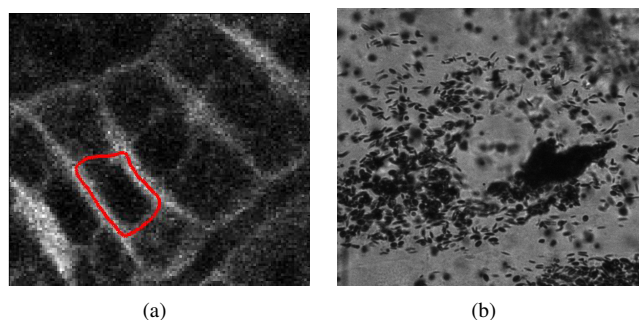


Fig. 2. (a) The challenges of the confocal microscopy membrane images of the ascidian *Ciona* include varying boundary intensity and the faint presence of other organelles. In this example, the method in [3] incorrectly attached to the faint presence of the nucleus boundary instead of the cell membrane (boundary in red). (b) Bright field microscopy image from timelapse of melanosomes in mouse retina. The dark oval structures are melanosomes for which automatic detection and tracking is necessary.

Utah. This data consists of nanometer resolution electron micrograph stacks of the rabbit retina for the study of neuronal interconnectivity (Fig. 10(a)) [4]. The aim is to trace neuronal structures through the stack, detect the synapses between the neurons and construct a neuronal circuit. The primary problem is the scale of the dataset (terabytes of data), where the annotation has been an ongoing effort over the past few years. Tracing these cellular and sub-cellular structures requires robustness to abrupt topological deformations in the z-direction, the inherent noisy texture of the structures in the electron micrographs, and be able to scale to large datasets.

The above examples demonstrate the need for robust segmentation, tracking and tracing methods that can minimize the need for human intervention in various biological imaging applications. In this paper we address these problems by using graph-based techniques. We extract low-level cues (superpixels, detections and tracklets) and model them as nodes in a graph. The interaction between these low-level cues is

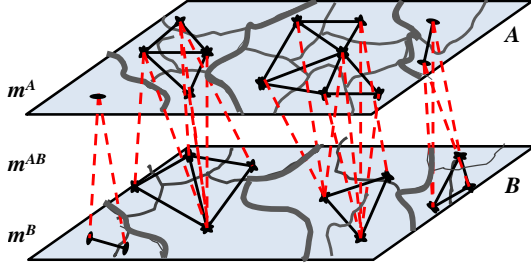


Fig. 3. Conceptual example of a possible resulting segmentation, where m^A , m^B and m^{AB} are in agreement with each other. Superpixels are modelled as nodes in the graph. The black edges in the graph mark pairs of neighboring super-pixels which should be merged inside each segmentation as a result of the linear program. The dotted red edges mark pairs of super-pixels across the two segmentation which segment out the same object in the original image, and thus result in a reward for agreement across the two segmentations. The total cost function that the linear program accumulates the total cost for merging inside label map A , the total cost for merging inside label map B , and the total reward for merging across the two label maps.

modelled as edges in the graph. We solve a grouping problem on the graph, where the resulting groups yield segments, traces and tracks.

II. APPLICATION 1: A LINEAR PROGRAM FORMULATION FOR JOINT SEGMENTATION

Approach: We introduce a linear programming framework for correcting two given over-segmentations (label maps) in a principled manner in order to achieve consensus among the boundary pixels. Each of the two input label maps is an over-segmentation of the desired result. The methods generating these segmentations could differ in their segmentation scheme or in the choice of parameters for a single algorithm. The intuition here is that the boundaries of interest are present in both over-segmentations, while many other spurious edges are method- or parameter-dependent. Our algorithm simultaneously corrects the two over-segmentations by maximizing the agreement between the two. We apply our method for the cell segmentation of a 3-D confocal microscopy membrane volume of the ascidian *Ciona* (Fig. 1).

Representation: Given two over-segmentations (label-maps) of an image, we intend to correct these two over-segmentations by merging pairs of segments inside each of the two label-maps, such that the resulting segmentations are in agreement with each other. The two over-segmentations are modelled as two interconnected planar graphs. Each segment is modelled as a node in the graph, where the edges represent pairs of neighboring segments. Two segments from the two different segmentations are considered neighbors if they share at least one pixel. A cost for connecting segments within each label map is assigned to each pair of neighboring segments inside each label map. A reward for connecting segments across the two label maps is assigned to each pair of segments from the two different over-segmentations that share at least one pixel. The goal is to merge (connect) segments within each over-segmentation while minimizing the total cost, and to merge (connect) segments across the two over-segmentations while maximizing the total reward.

Consider two label maps (over-segmentations), A and B in Fig. 3. We introduce the parameter C^A , an $N_1 \times N_1$ dimensional matrix whose entries C_{ij}^A represent the penalty for merging segments i and j within label-map A . Similarly, for label-map B we introduce parameter C^B of dimensions $N_2 \times N_2$ which models the cost for merging two segments in label map B . Consequently, the reward parameter is represented by the $N_1 \times N_2$ dimensional matrix R^{AB} , whose entries R_{ij}^{AB} quantify the agreement between segment i in label-map A and j in label-map B . The cost and reward values range between 0 and 1.

Next, we introduce the decision variables. Variable m^A (m^B) is an $N_1 \times N_1$ ($N_2 \times N_2$) dimensional binary matrix whose entries m_{ij}^A (m_{ij}^B) are non-zero if the two segments i and j from label-map A (B) should be connected (merged). Next, we introduce variable m^{AB} as a $N_1 \times N_2$ dimensional binary matrix whose entries m_{ij}^{AB} , indicate whether segment i from label-map A and segment j from label-map B are likely to be part of the same object from the original image, meaning that segment i is connected to segment j across the two label maps.

Three objectives can be identified: minimizing the total cost of merging within label-map A , minimizing the total cost of merging within label-map B , and maximizing the total reward for agreement across label map A and label map B .

$$f = \left(\sum_{i=1}^{N_1} \sum_{j=1}^{N_1} m_{ij}^A C_{ij}^A + \sum_{i=1}^{N_2} \sum_{j=1}^{N_2} m_{ij}^B C_{ij}^B \right) - \lambda \cdot \left(\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} m_{ij}^{AB} R_{ij}^{AB} \right) \quad (1)$$

The objective function of the proposed optimization program is represented by equation (1). The first two summations indicate the total cost of merging within each label-map. The second summation represents the reward for agreement (merging) across the two label-maps. Note that minimizing the negative total reward is equivalent to maximizing the total positive reward. The real valued parameter λ is used to balance the total cost and total reward, and bias the final result towards more or less merging.

In order to make sure that the resulting solutions for the three sub-problems are in agreement with each other, we identify a few constraints associated with this program. The range constraint restricts the decision variables to the set $\{0,1\}$. The symmetry constraints ensure that if segment i is merged to segment j , then segment j is also merged to segment i . The connectivity constraint ensures that only segments which are neighbors to each other can be merged. Finally, the transitivity constraint makes sure that if segment i is merged with k , which in turn is merged with j , then as a result segment i is also merged with j (i, j, k from either label-map).

The current optimization program is non-linear, but can be relaxed to a linear program. The decision variables are relaxed to take values within $[0,1]$, thus allowing a convex set.

Experiments: The formulation of our method is agnostic to data dimensionality. We apply our algorithm for the segmen-

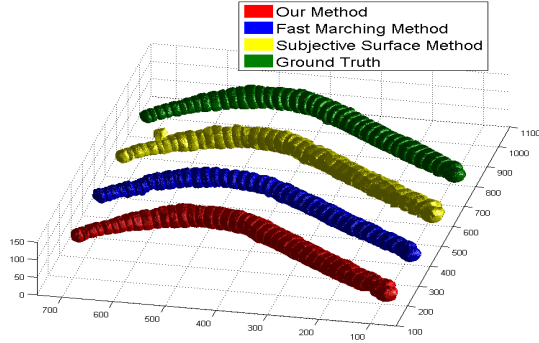


Fig. 4. 3-D rendering of the segmented notochord cells.

TABLE I
F-MEASURE FOR THE NOTOCHORD SEGMENTATION OF THE 3-D
CONFOCAL MICROSCOPY MEMBRANE DATASET OF THE ASCIDIAN *Ciona*.

Rank	Method	Avg. F-Measure
1	Proposed Method	89.36%
2	Method in [5]	86.25%
3	Fast Marching [7]	85.58%
4	Initial S^1 (Watershed based)	80.00%
5	Subjective Surf [3]	77.87%
6	Initial S^2 (Watershed based)	61.62%

tation of 3-D confocal microscopy volumes of the ascidian *Ciona*. We use two over-segmentations obtained as follows (1) Watershed method directly applied to the image data, and (2) Watershed with a different threshold level applied the filtered image (3-D box-filter). The cost for merging two segments within each label map is an edge-based cost function which compares the intensity on the border of the two segments as opposed to the intensity within the segments. The reward function for two segments across the two label maps consists of the percent area of overlap.

The following methods are relevant in the context of confocal membrane image segmentation. The Subjective Surface variant in [3] is a level-set method specifically designed for membrane image segmentation, which requires the initialization with a seed point in every cell. The method in [5], also corrects an over-segmentation by merging segments, however this method requires the user to create training data. Lastly, we compare our method against the Fast Marching Method [6], which computes the geodesic distance in discrete image domain, from a seed-point manually placed in every cell. The comparison metric is the F -measure, which is a volume based error metric and is defined as: $F = \frac{2PR}{P+R}$, where P and R are the precision and the recall for a given ground truth volume. Fig. 5 and Table I show that the proposed method achieves the best performance.

III. APPLICATION 2: SIMULTANEOUS DETECTION AND TRACKING OF MULTIPLE STRUCTURES

Approach: In order to track the melanosomes in bright field microscopy time lapse images, we utilize a tracking by detection approach where we detect the melanosomes in all the frames and tracking is solved by associating these detections.

Representation: The tracking problem is divided into two stages: tracklet building and track construction. The tracklet building stage is a conservative association scheme where

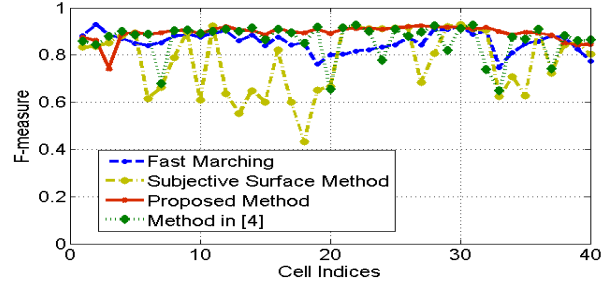


Fig. 5. F measures for each of the 40 notochord cells.

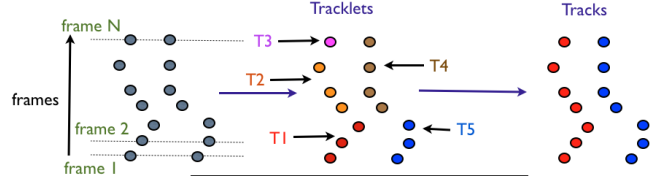


Fig. 6. Tracking by detection: First melanosomes are detected in each frame of the time sequence. Next, tracklets are formed by making confident associations of the detections in one frame to the detections in the next frame. Finally, tracklets are connected to form tracks. This process repeats, such that tracklet association can help identify missing detections. These improved detections give improved tracks in the following iterations.

every detection is associated to another detection in the previous frame if the association probability is high and there is no other conflicting alternate association. The tracks are further obtained by a bipartite graph matching approach where the association probability for every pair of tracklets is computed and optimal track associations are obtained by solving the graph matching problem. We note that the tracklet association probabilities can either be learnt from a training set or evaluated empirically using simple probabilistic models for appearance, motion and spatial statistics. An example scenario for the two step process is shown in Fig. 6

We detect the melanosomes using an elliptical template based detector, but obtaining perfect detections is extremely difficult. Therefore, we propose a simultaneous detection and tracking approach which jointly recovers missing detections and estimates associations in a unified probabilistic framework. Tracking methods typically work with detections which yield low false positives and low false negatives. In our work we augment this set of detections with another set consisting of high false positives without any missing detections as shown in Fig.7. This can be obtained by moving along the precision recall curve of the detector by adjusting suitable thresholds. We simultaneously use both these sets of detections in the following manner. The first set of detections provides the tracklet association cost. The second set of detections (no missing detections) provide a missing detection cost and both of these are summed to obtain a new association matrix which accounts for both tracklet association and missing detections. The bipartite graph matching problem is solved on this joint matrix and tracklet associations and missing detections are jointly recovered. For a detailed description of our approach, the reader is referred to [8]

Experiments: In our experiments we compare our proposed

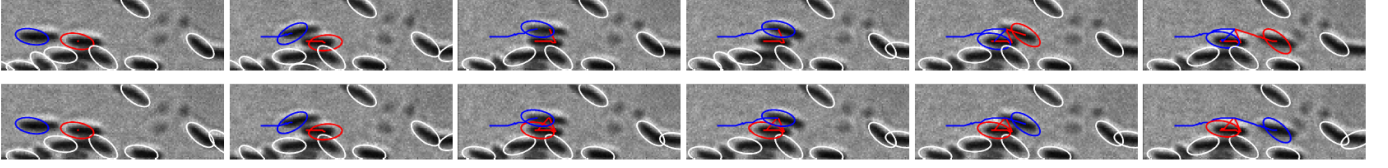


Fig. 8. This scenario shows a moving melanosome (blue) and a stationary melanosome (red). The baseline approach (top row) misses a few detections of the stationary melanosomes and hence wrongly switches the moving melanosome and the stationary melanosome when tracking. Our approach (bottom row) is robust to missed detections and tracks the two melanosomes correctly.

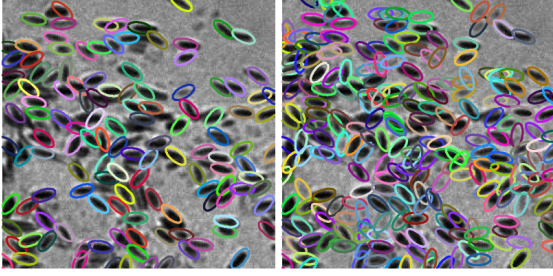


Fig. 7. Example of two sets of detections. (Left) Detections with low false positive and false negative (Right) Detections with high false positive, no missing detections.

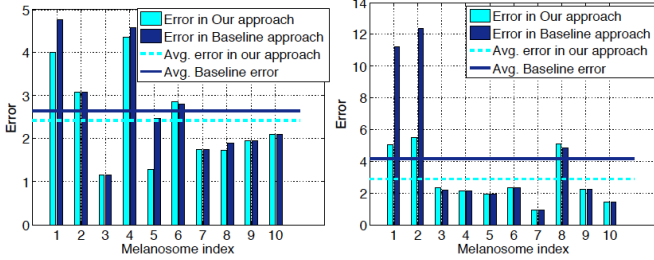


Fig. 9. Comparison of our approach with the baseline approach [9] on two different image sequences.

approach to the baseline approach in [9] which does not model the cost for missing detections. Fig.8 shows an example scenario where our approach outperforms the baseline approach [9] in tracking melanosomes. Our approach recovers missing detections while creating associations which lead to robust tracking.

We evaluated our tracking results by creating ground truth for 20 melanosomes in two different image sequences of 50 frames each. The first sequence consists of approximately 120 melanosomes and the second consists of approximately 150 melanosomes. Fig. 9 shows that our algorithm outperforms the baseline approach [9] in these datasets.

IV. APPLICATION 3: SCALABLE TRACING IN ELECTRON MICROGRAPHS IMAGES

Approach: We now turn to the problem of scalable tracing in Electron Micrograph (EM) stacks by fusing top down and bottom up cues in a unified framework. The top down cues can be thought of as trajectories or 3D reconstructions obtained using an auxiliary algorithm that serves as a prior to the main tracing procedure.

Representation of Bottom Up Cues: We adopt superpixels, or oversegmented image regions as the base representation in our formulation. The EM stacks comprise multiple 2D

image slices that are independently segmented using the mean shift algorithm to generate a set of superpixels per slice. The resulting superpixels are connected to their neighbors in the same slice and the slices immediately above and below the current slice, see Fig. 10(a). The above method for constructing a graph by connecting immediate neighbors constitutes of pairwise interactions or the bottom up cues. The edge weight between graph nodes i and j is computed by a combination of the KL-divergence between the gray scale h_{gray} and Local Binary Pattern (LBP) h_{lbp} histograms:

$$KL[h(i), h(j)] = \sum_k h(i, k) \ln \frac{h(i, k)}{h(j, k)} + h(j, k) \ln \frac{h(j, k)}{h(i, k)}$$

$$w(i, j) = \exp(-b(KL[h_{gray}(i), h_{gray}(j)] + KL[h_{lbp}(i), h_{lbp}(j)])) \quad (2)$$

Representation of Top Down Cues: While pairwise edges are useful in modelling local image structure, they are not robust to clutter and other artifacts that distract tracing algorithms. In order to address this problem, we model higher order interactions between superpixels by generalizing pairwise edges to hyperedges, see Fig. 10(b). Hyperedges are initially constructed by linking a superpixel with its K-nearest neighbor superpixels. Further, if a complementary algorithm yields grouping constraints between superpixels, those constraints are introduced as new hyperedges for inducing higher order label consistency. The auxiliary algorithm used in our work employs boosted edge learning [10] / watersheds for generating homogeneous 2D image regions followed by shortest path computations on the volume, yielding higher order grouping constraints on superpixels. The reader is referred to [11], [12] for details pertaining to the cost function unifying top down and bottom up cues, and related inference issues.

Experiments: We present experimental results on a dataset sampled from a connectome volume imaged from a rabbits retinal tissue, where the challenge is to trace 95 neuronal structures. A visual illustration of the input images and results obtained are presented in Fig. 11. Observe the severe clutter, object deformations and imaging artifacts that make the problem very challenging. We further benchmark the proposed approach with state of art tracing techniques, see [11] for more details. An important finding from our experiments is a marked rise in performance when hyperedges are utilized (F-Measure: 0.78), as opposed to using pairwise edges alone (F-Measure: 0.23). Further, hypergraph tracing performs significantly better without sacrificing computation time pointing to the promise of the proposed technique.

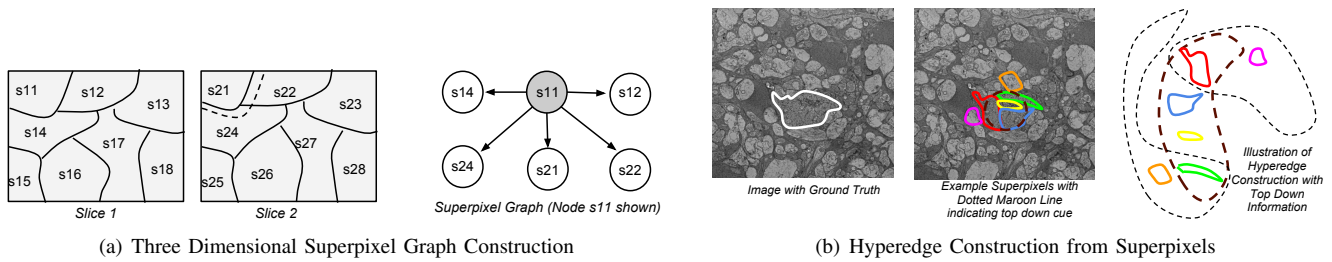


Fig. 10. Construction of the 3D superpixel hypergraph

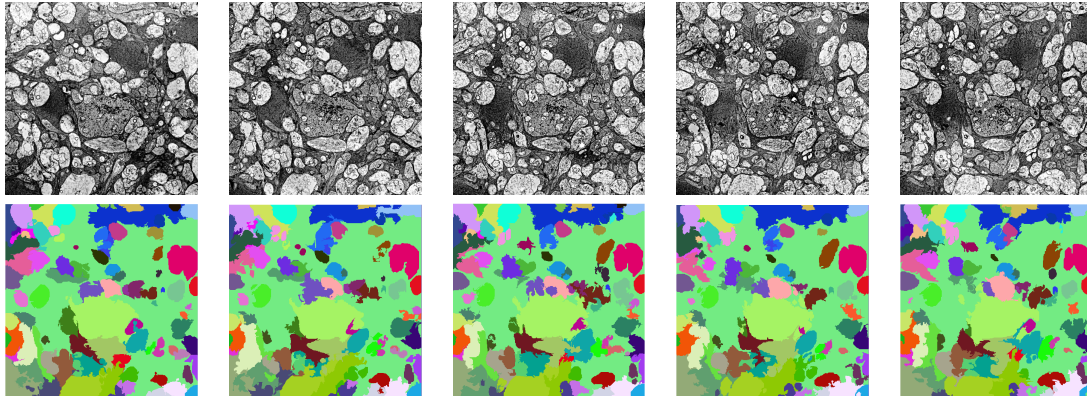


Fig. 11. Top row: example of slices from the electron micrograph stack. Bottom row: Result of hypergraph diffusion on Dataset-I.

V. CONCLUSION

Robust image analysis methods are essential for quantitative biology. In this paper, we discussed three graph-based methods. First, 3-D image segmentation by simultaneous correction of two over-segmentations, based on a linear program formulation was discussed and applied to notochord segmentation of confocal microscopy membrane images of the ascidian *Ciona*. Next we addressed the problem of simultaneous tracking of multiple melanosomes in a time sequence of bright field microscopy images of the rabbit retina, by associating confident tracklets. The final application discussed in this paper addresses connectome tracing in stacks of electron microgram images of the rabbit retina, based on top-down and bottom-up cues using hypergraph diffusion. All these problems are tackled as graph based association algorithms from low-level tracklets and superpixels.

ACKNOWLEDGMENT

This work was supported by award 2R01HD059217 from the National Institutes of Health, and by awards III-0808772 and OIA-0941717 from National Science Foundation.

The authors thank Michael Veeman from Prof. William Smith's lab at the University of California, Santa Barbara for collecting the confocal microscopy membrane images. *Ciona* embryos were fixed, stained with Bodipy-FL phalloidin to label the cortical actin cytoskeleton, cleared in Murrays Clear (BABB), and imaged on an Olympus FV1000 LSCM using a 40x 1.3NA oil immersion objective.

The authors thank Mei Jiang and Dr. David Williams, from the Williams Lab at UCLA for contributing the bright field microscopy time-lapse images of melanosomes in mouse retina.

The authors thank Dr. Robert Marc, Dr. Bryan Jones and Dr. James Anderson from the Univ. of Utah for providing data used in experiments and for useful discussions. The imaging was done using serial section transmission electron microscopy.

The authors thank Dr. Volkan Rodoplu for the very helpful discussions regarding the linear program formulation for joint over-segmentation correction.

REFERENCES

- [1] M. Veeman, S. Chiba, W. Smith *et al.*, "Ciona genetics," *Methods Mol. Biol.*, vol. 770, p. 401, 2011.
- [2] D. Gibbs, A. Cideciyan, S. Jacobson, and D. Williams, "Retinal pigment epithelium defects in humans and mice with mutations in myo7a: imaging melanosome-specific autofluorescence," *Investigative ophthalmology & visual science*, vol. 50, no. 9, pp. 4386–4393, 2009.
- [3] C. Zanella, M. Campana, B. Rizzi, C. Melani, G. Sanguinetti, P. Bourguin, K. Mikula, N. Peyrieras, and A. Sarti, "Cells segmentation from 3-D confocal images Of early zebrafish embryogenesis," *IEEE Transactions on Image Processing (TIP '10)*.
- [4] J. Anderson, S. Mohammed, B. Grimm, B. Jones, P. Koshevoy, T. Tasdizen, R. Whitaker, and R. Marc, "The viking viewer for connectomics: scalable multi-user annotation and summarization of large volume data sets," *Journal of microscopy*, 2010.
- [5] D. D. et al., "An automatic feature based model for cell segmentation from confocal microscopy volumes," in *IEEE ISBI 2011*.
- [6] L. Cohen and R. Kimmel, "Fast marching the global minimum of active contours," in *Proceedings of International Conference on Image Processing*, vol. 96.
- [7] J. Sethian, "A fast marching level set method for monotonically advancing fronts," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 4, p. 1591, 1996.
- [8] S. Karthikeyan, D. Delibaltov, U. Gaur, M. Jiang, D. Williams, and B. Manjunath, "Unified probabilistic framework for simultaneous detection and tracking of multiple objects with application to bio-image sequences," in *ICIP 2012*.
- [9] C. Huang, B. Wu, and R. Nevatia, "Robust object tracking by hierarchical association of detection responses," *Computer Vision—ECCV 2008*, pp. 788–801, 2008.
- [10] P. Dollar, Z. Tu, and S. Belongie, "Supervised learning of edges and object boundaries," in *CVPR06*.
- [11] V. Jagadeesh, M. Shih, B. Manjunath, and K. Rose, "Scalable tracing of electron micrographs by fusing top down and bottom up cues using hypergraph diffusion."
- [12] D. Zhou, J. Huang, and B. Scholkopf, "Learning with hypergraphs: Clustering, classification, and embedding," *NIPS07*.