

Tunable Tensor Voting Improves Grouping of Membrane-Bound Macromolecules

Leandro A. Loss
Lawrence Berkeley Nat. Lab.
Berkeley, CA
LALoss@lbl.gov

George Bebis
University of Nevada, Reno
Reno, NV
Bebis@cse.unr.edu

Bahram Parvin
Lawrence Berkeley Nat. Lab.
Berkeley, CA
B.Parvin@lbl.gov

Abstract

Membrane-bound macromolecules are responsible for structural support and mediation of cell-cell adhesion in tissues. Quantitative analysis of these macromolecules provides morphological indices for damage or loss of tissue, for example as a result of exogenous stimuli. From an optical point of view, a membrane signal may have nonuniform intensity around the cell boundary, be punctate or diffused, and may even be perceptual at certain locations along the boundary. In this paper, a method for the detection and grouping of punctate, diffuse curvilinear signals is proposed. Our work builds upon the tensor voting and the iterative voting frameworks to propose an efficient method to detect and refine perceptually interesting curvilinear structures in images. The novelty of our method lies on the idea of iteratively tuning the tensor voting fields, which allows the concentration of the votes only over areas of interest. We validate the utility of our system with synthetic and annotated real data. The effectiveness of the tunable tensor voting is demonstrated on complex phenotypic signals that are representative of membrane-bound macromolecular structures.

1. Introduction

Cell membrane macromolecules mediate the cell-cell adhesion of epithelial cells allowing the composition of monolayers in culture. One such macromolecule, E-cadherin, is pathoneumonic for normal epithelia and its down regulation is associated with motility, epithelial-mesenchymal transition (EMT) and cancer initiation [2]. Research in the area of quantitative analysis of cell-based assay has spanned learning techniques using texture-based features for characterizing patterns of macromolecule expression [8], geometric techniques using nonlinear filtering and curve evolution [3], and shape regularization for segmentation of subcellular compartments [11]. While segmentation of nuclear regions

provides context for localization studies, membrane-bound macromolecules also need to be delineated [10]. Such delineation enables, for example, membrane-bound macromolecules to be quantified on a cell-by-cell basis. At optical resolution, however, membrane signals may have nonuniform intensity around the cell boundary, be punctate (e.g., connexin) or diffused (e.g., E-cadherin), and may even be perceptual at certain locations (Fig. 1). Furthermore, there is a significant amount of heterogeneity as a result of technical and biological variations. In this paper, a new method is presented for quantifying macromolecule bound to the basal-lateral region of the cell.

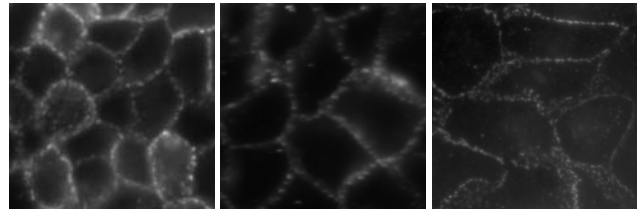


Figure 1. Membrane-bound macromolecules has a complex pattern of localization along the cell membrane. Notice that the signal has an additional punctate pattern on top of an existing diffused signal.

It is well known that symmetry, closure, and continuity are preattentive processes in the human vision system that can aid in object-level delineation and recognition [1]. In this paper, we propose a method that allows inference of curvilinear structures from noisy, often incomplete boundary information, such as those found in membrane-bound macromolecule signals. It involves perceptual grouping of pixels through voting. In the context of voting as a precursor for grouping and clustering, many methods have been developed. For example, Hough introduced the notion of parametric clustering in terms of well-defined geometry, which was later extended to the generalized Hough transform [4]. Guy and Medioni [5] proposed a general purpose tensor voting framework that uses deformable tensors to reveal perceptual structures. Loss *et al.* [6] extended this tensor voting

framework to accommodate iterative saliency thresholding and multi-scale analysis. Parvin *et al.* [9] developed an iterative voting system that employs tunable kernels to refine paths of low curvature in images. In general, voting operates on continuity and proximity, which can occur at multiple scales, e.g., points, lines, or lines of symmetry. One of the main advantages of voting frameworks is their reliance on relative simple models, which reduces the number of free parameters considerably. Our work builds upon the tensor voting and the iterative voting frameworks in order to produce an efficient method to detect and refine perceptually interesting linear structures in images. By coupling tensor and iterative voting fundamentals we leverage advantages of both methods to produce better results than those achieved by them individually. The novelty of our method lies on the extension of the tensor voting framework to precisely detect and refine linear structures at different scales, by iteratively tuning the tensor fields as pixel orientations are better defined.

Our method starts by encoding each pixel in an image as an unoriented tensor, whose size is proportional to pixel intensities. First, a tensor voting pass is executed using a ball field, i.e. votes are propagated radially, as no initial orientation is known. This allows tensors to start their characteristic structural deformation that consequently reveals, although still inaccurately, the presence or not of perceptual lines in the image. Although the classical tensor voting would stop at this stage, we proceed with consecutive tensor voting passes aiming at refining previous results. These iterations are performed with stick fields, i.e. votes are concentrated along the pixel’s tangent only, for this is the most natural continuation of a line passing through the pixel. The concentration of the votes through stick fields is possible because the first voting pass naturally produces an estimation of the orientation at each pixel. Note that, in contrast to [9], our method is independent of initial estimates of gradient or curvature. One interesting observation is that the stick fields are gradually tuned, i.e. the field aperture is reduced as the voting iterations proceed and the orientation estimations become more and more accurate. The method is applicable to detection of linear features, has excellent noise immunity, is tolerant to changes in target scale, and applicable to a large class of application domains.

In order to assess the improvements yielded by the iterative tuning of tensor voting, we perform experiments involving synthetic configurations and real microscopic images. Synthetic images are used to help us analyze and predict the method’s behavior, in particular, in the presence of different linear curvatures and junctions. Experiments on microscopic images of membrane-bound macromolecules aim at evaluating the method on real scenarios, providing an insight about its potential and effectiveness.

The remaining of this paper is organized as follows:

Section 2 describes the tensor voting and the iterative voting frameworks, along with their applications to perceptual grouping of linear structures. Section 3 introduces our method, extending the concepts of the tensor and the iterative voting frameworks. Experimental results are shown in Section 4, and conclusions are presented in Section 5.

2. Voting frameworks

2.1. The tensor voting framework

In the framework proposed by [5], perceptual grouping is achieved by vote casting between elements of an image. Such elements are represented as tensors, mathematical entities whose capability of encoding magnitude and orientation make tensor voting particularly efficient for detection of perceptually organized structures, such as edges, lines and regions. In 2D, tensors can be represented geometrically as ellipses or analytically as 2 by 2 matrices. Initialized with an arbitrary size and shape (given respectively by the eigenvalues, λ_1, λ_2 , and eigenvectors, e_1, e_2 , of its analytical representation), input tensors are gradually deformed due to the accumulation of votes cast by other neighboring tensors. Votes are also tensors composed of certain magnitude and orientation, which encode the Gestalt principles of proximity, smoothness and good continuation.

Depending on the nature of the input elements, a priori information about their orientation can be available or not. Therefore, tensor voting offers two possible vote casting configurations: one that concentrates the votes according to the input orientation (stick field - Fig. 2(a)) and another one that casts votes radially (ball field - Fig. 2(b)). The voting fields are the composition of all votes that can be cast from a tensor located in the center of the field to its neighboring tensors. Given two tensors positioned in the image, the angle θ , arc length s and curvature κ between them is used to produce the vote V from one another, as shown by Equation (1), where N is the vector normal to the smoothest path between the two tensors and is given by $[-\sin(\theta) \cos(\theta)]^T$. Note that the stick field exists only at $\theta \leq 45^\circ$.

$$V = e^{-\frac{s^2 + c\kappa^2}{\sigma^2}} NN^T \quad (1)$$

The tensor deformation imposed by accumulating the strength and orientation of the votes eventually reveals behavioral coherence among image elements. In other words, elements that lie on the same salient feature (e.g. a curve or a region) strongly support each other and deform the tensor at those sites according to the underlying structure orientation. Therefore, each kind of structure is expected to produce tensors of a particular shape: very elongated tensors (high $\lambda_1 - \lambda_2$) for lines, and more rounded ones (low $\lambda_1 - \lambda_2$) for regions. Fig. 3 exemplifies how a set of (a)

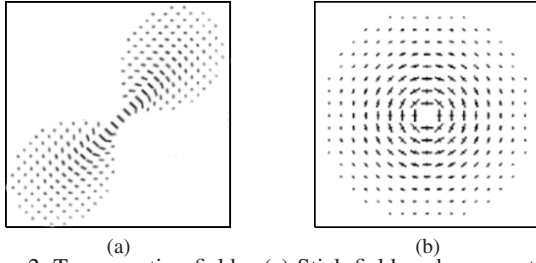


Figure 2. Tensor voting fields. (a) Stick field - when an estimate of the initial orientation is known, and (b) Ball field - when no orientation information is known.

input elements are (b) encoded as tensors, whose (c) deformations resulting from accumulated votes reveal an underlying salient linear structure. The voting process can also be either *sparse* or *dense*. Sparse voting restricts tensors to cast votes only on other encoded input tensors, while dense voting extrapolates the input configuration allowing tensors to cast votes everywhere within their neighborhood. Tensor voting has been shown to be robust to considerable amounts of noise and does not depend on critical thresholds. The only free parameter is the scale factor σ , which determines the range of the voting neighborhood. More detailed information can be found in [7].

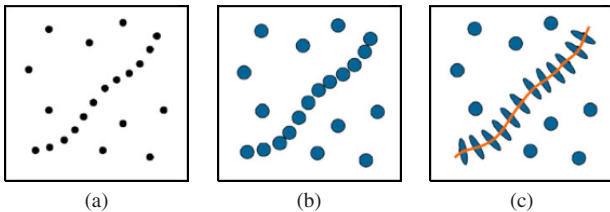


Figure 3. Example of perceptual grouping through tensor voting. A set of (a) input elements are (b) encoded as tensors, whose (c) resulting deformations reveal a curve.

2.2. Iterative voting

The framework proposed by Parvin *et al.* [9] also uses a voting approach to detect and group linear patterns. Similarly to the tensor voting framework, each pixel spreads its magnitude by casting votes to its neighbors. Here, however, pixels are initialized as negative curvature maxima and a set of voting kernels are utilized (Fig. 4). Each kernel represents the vote a pixel will cast to other pixels located in its vicinity. The voting kernels present a Gaussian-like energy profile that decays with the distance from its center. After voting, the orientation of each pixel is estimated from the direction of the maximum magnitude around this pixel. The estimated orientations are then used to realign the kernels in subsequent iterations. Since a better estimation of the pixel orientation is produced after each iteration, the kernels have their energy tuned from initially diffuse to eventually focused over linear structures emerging from the process.

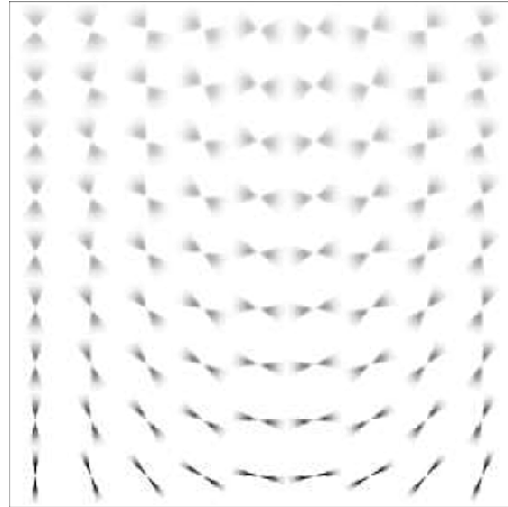


Figure 4. The voting kernels developed in [9] for the detection of linear patterns. The kernels' energy is (1) dissipated as a function of distance and (2) focused at consecutive iterations (top to bottom row).

A refined set of lines are shown to be produced from the iterative voting. The method is used to characterize membrane signals, and scaled to deal with 3D images. Detailed information can be found in [9].

3. Tunable tensor voting

We build upon the tensor voting [5] and iterative voting [9] frameworks in order to produce an efficient method to detect and refine perceptually interesting linear structures in images. The method is developed to tackle problems in which lines (1) may be punctate and noisy, (2) are highly surrounded by clutter, and (3) have nonuniform and diffuse intensities. It is based on iterative tuning of the tensor voting fields, eventually concentrating the votes only over real lines, producing better, enhanced results, when compared to the classical tensor voting (i.e. single run of ball voting).

Our method starts by encoding each pixel as an unoriented tensor, whose size is proportional to pixel intensities ($\lambda_1 = \lambda_2 = I_{ij}$). First, a tensor voting pass is executed using the ball field (Fig. 2(b)), as no initial pixel orientation is known. This allows tensors to start their characteristic structural deformation that consequently reveals, although still inaccurately, the presence or not of perceptual lines in the image. Although the classical tensor voting would stop at this stage, we proceed with consecutive dense tensor voting passes aiming at refining the previous results. These consequent iterations are performed with stick fields (Fig. 2(a)), as the first voting naturally produces an estimation of the orientation at each pixel. An interesting observation is that the stick fields are gradually tuned (i.e. the field aperture is reduced) as the voting iterations proceed and the orienta-

tion estimations become more and more accurate (Fig. 5). At each iteration, tensors that do not deform as lines (low $\lambda_1 - \lambda_2$) are eliminated so their influence is not accounted into the following iterations. Fig. 6 depicts the process.

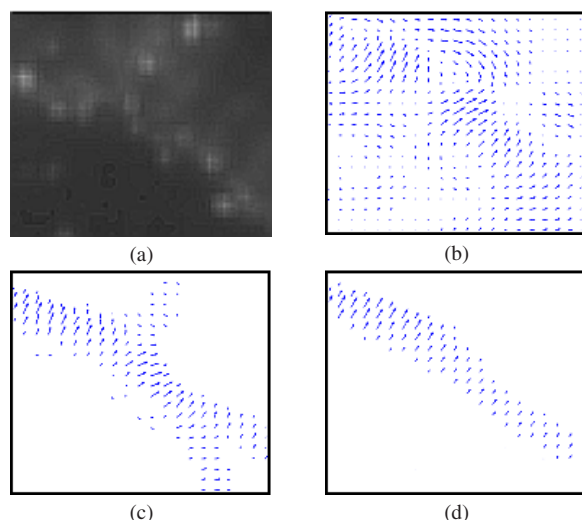


Figure 5. The signal along the curvilinear path is gradually refined by using tunable filters: (a) Original image. (b) After first iteration (ball voting). (c) After four iterations (at $\theta = 30^\circ$). (d) Final result (at $\theta = 5^\circ$).

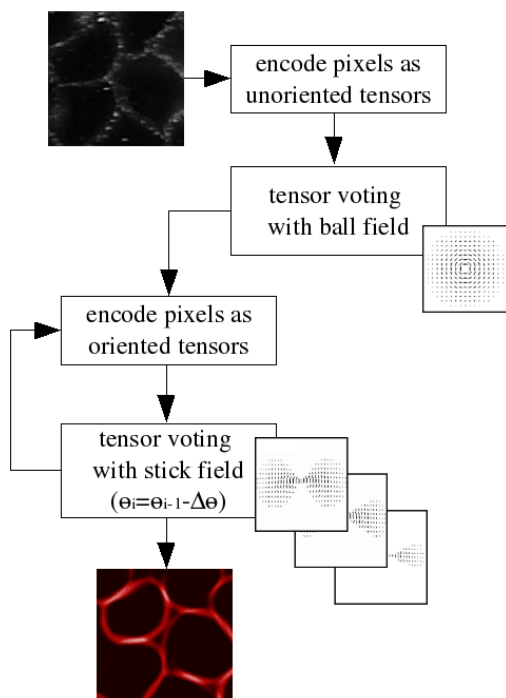


Figure 6. Detection of linear patterns by tunable tensor voting. The primary theme is the feedback loop for iterative change of the voting aperture for continuous refinements.

This tuning process eventually concentrates the votes

only over real lines, producing better, enhanced results. The iterations stop when the aperture of the voting field is small enough, producing the same field as in a previous iteration or at $\theta = 1^\circ$.

In the next section, we experimentally show that our tunable tensor voting produces refined, more accurate membrane lines than the previous methods.

4. Experimental results

We demonstrate here the benefits of the Tunable Tensor Voting (TTV) for detection and grouping of membrane-bound macromolecule signals (e.g. E-cadherin). We validate the utility of our system using synthetic and annotated real data. TTV's effectiveness is demonstrated on complex phenotypic signals that are representative of membrane-bound macromolecular structures.

4.1. Synthetic data

Fig. 7 shows how the TTV performs on a few synthetic configurations. The synthetic images (a) were generated in an attempt to simulate real punctate, diffuse, noisy signals found in real applications. The first row shows three signals of different curvature and strength, the second column has a punctate T-junction, and the third column has a punctate, diffuse X-junction contaminated with random noise. Intermediate results are shown on (b) ball voting (also classical tensor voting result) and (c) stick voting with $\theta = 45^\circ$ column. (d) shows the resulting signal obtained by TTV. The results combine pixels whose either stick or ball saliency is high. This is important to better preserve junctions [7]. One can notice how the tunable tensor voting is able to evolve from a punctate, diffuse signal to a more or less strong filament.

4.2. Real data

TTV was applied in quantifying E-cadherin that is bound to the basal-lateral region of the cell. E-cadherin is expressed along the cell boundary under normal mammary tissue homeostasis (Fig. 1). At optical resolution, the membrane signal is sometimes visualized as punctate and noisy, has nonuniform intensity, and is perceptual at some locations. Fig. 8 shows (a) localization of macromolecules at the membrane, (b) the result of first iteration (also from ball voting), (c) result of an intermediate tuning iteration, and (d) overlay of iterative voting on the original image. The enhancement achieved (Fig. 8(d)) is clear when compared with the original membrane signal (Fig. 8(a)). Note that by iterating over the result obtained first by a regular application of tensor voting, pixels belonging to the curvilinear structure are determined with better precision.

Fig. 9 shows examples of actual results produced by our method. Images on the top row are sub-images of the orig-

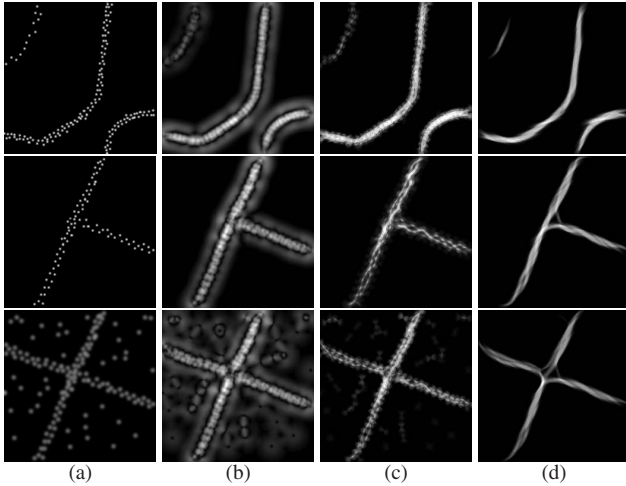


Figure 7. Synthetic signals processed by TTV. Examples of expected results produced by TTV on (first row) curves of different curvature, (second row) a T- junction, (third row) X- junction and noise. (a) original images; (b) ball voting results (also classical tensor voting result); (c) intermediate results with stick fields; (d) resulting signal.

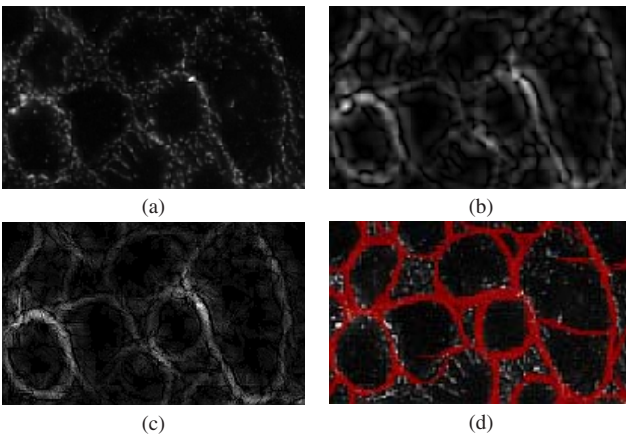


Figure 8. Membrane grouping through tunable tensor voting. (a) Original membrane signal. (b) Ball tensor voting result (Also classical tensor voting result). (c) Result of an intermediate tuning iteration. (d) Enhancement promoted by tunable tensor voting.

inal one, while those on the bottom are their respective results. In general, membrane signals are highly dispersed along the cell membrane. This is partially due to wide field microscopy, the influence of out-of-focus light, and/or sample preparation. Note that even in the presence of noisy and dim signals, our method was able to infer lines interpolating punctate and diffused patterns, filling up perceptual gaps.

4.3. Annotated data

Visual inspection clearly shows the benefits of detection and grouping by TTV. To quantify these benefits, we compared the results produced by TTV with manually marked results. Our data set consists of 274 1344x1024 images

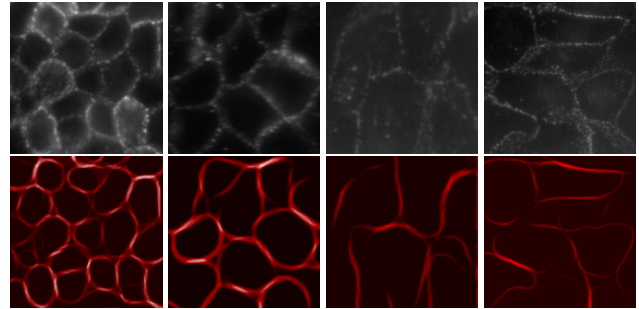


Figure 9. Results of tunable tensor voting on membrane grouping. First row - original membrane signal; second row - resulting membrane detected. Even in the presence of noisy and dim signals, TTV was able to infer lines interpolating punctate and diffused patterns, filling up perceptual gaps.

with membrane signals presenting similar characteristics. Ground truth was produced by two cell biologists, whose only instruction was to trace all cell membranes they could detect using a computer mouse. Annotation was performed to around 5% of the images. The discrepancy between the two sets of independently marked membranes was evaluated to determine human performance on the data set. Performance was reported by recall vs. precision plots. In the problem at hand, precision measures the probability of a marked line to be a true membrane. Recall is a measure of the probability of a membrane to be marked. In order to account for natural misalignment between results, evaluation was performed after dilating ground truth lines by a certain number of pixels. The amount of dilation was used as a parameter that varies from 1 to 20 pixels wide. Precision and recall rates were then computed after overlapping a set of marked lines with each dilated ground truth. The number of pixels in agreement between a set of marked lines and the dilated ground truth could then be quantified and compared to the ground truth, forming a curve in the recall vs. precision plot. The human performance is, therefore, the average curve produced by using one of the sets of manually marked lines as ground truth at a time, and the remaining one for evaluation. For instance, if the results produced by the biologists were identical, they would have achieved 100% precision and 100% recall rates. Since discrepancies did exist, human performance was used as quality assessment to computer algorithms, whose performance could be evaluated in the same fashion described (i.e. average curve resulting from each human set of lines as ground truth at a time).

Fig. 11 shows four of the most challenging images and the computed result by TTV. The top row shows the original (enhanced for a better display) images. The bottom row shows the lines extracted by TTV and the specialists. Red and green lines were marked by each specialist, while the white one shows the thinned line resulting from TTV. One

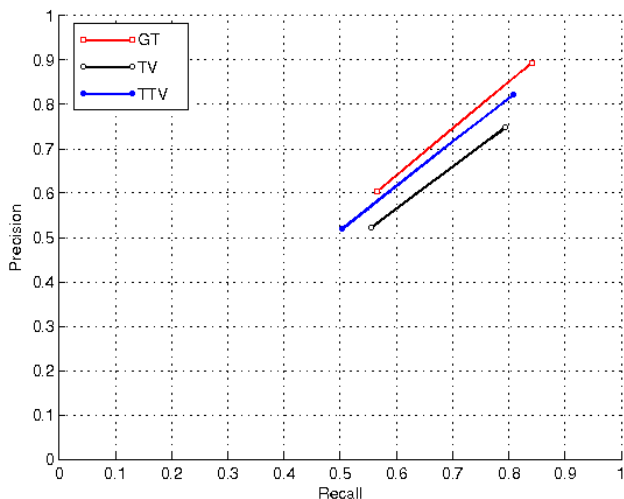


Figure 10. Recall vs. precision plot of the overall performance. GT stands for Ground Truth, i.e. the specialists’ performance, TV for tensor voting (ball voting) alone, and TTV for the Tunable Tensor Voting. TTV not only performs better than TV but also produces as high as 91% rate of agreement with GT in precision and 94% in recall.

Table 1. Rate of agreement between the methods and GT.

	Precision	Recall
TV	86%	95%
TTV	91%	94%

can notice that TTV produced results in fairly high agreement with the specialists in most of the cases. A larger disagreement happened at junctions, as predicted by the synthetic results from Fig. 7. Results were quantitatively measured and the overall average is shown in Fig. 10. GT stands for Ground Truth, i.e. the specialists’ performance, TV for tensor voting (ball voting) alone, and TTV for the Tunable Tensor Voting. Analysis of the behavior of TTV reveals that it was the closest curve to GT and had similar improvement rates as produced by the specialists, resulting in similar decay as we reduced the amount of GT dilation. Furthermore, comparing the rate of agreement between the methods and GT (Table 1), we can conclude that TTV had a higher agreement with the specialists, showing 91% precision and 94% recall rates of agreement. The rate of agreement was computed by averaging the differences of results from each dilated version of the ground truth and the method being evaluated. In other words, a method might have had achieved 65% precision, but if the human performance was only 75%, the rate of agreement between the two was 90%.

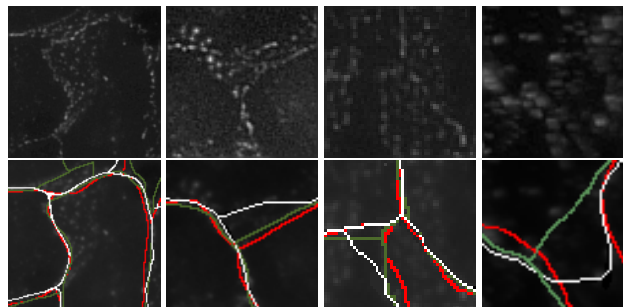


Figure 11. Results of the detection of punctate, diffuse membrane signals. The top row shows original (enhanced for a better display) images. The bottom row shows the lines extracted by TTV and the specialists. Red and green lines were marked by each specialist, while the white one shows the thinned line resulting from TTV.

5. Conclusions and future work

In this paper, we introduced a tunable tensor voting method, able to detect and refine punctate, diffused linear signals, including membrane-bound macromolecules. The method couples tensor and iterative voting frameworks to leverage advantages of both methods. As a result, complex patterns along a curvilinear path could be perceptually grouped and regularized. Such regularization can serve as pre-processing step, making possible to membrane-bound macromolecules to be quantified on a cell-by-cell basis, for example.

The method was shown to produce fairly high quality detection of membrane-bound macromolecule signals when compared to its previous method and the performance achieved by the specialists. In particular, our method achieved as high as 91% precision and 94% recall rates of agreement with the latter. One observation is that junction localization is particularly affected by our method, lowering its overall performance. This issue could, however, be overcome by including an *ad hoc* junction detector in the framework. For future work, we plan to extend our method to detect membrane signals in 3D cell cultures. One important remark is that membrane lines in 2D become surfaces in 3D. Similarly to the tensor voting framework, our method can be naturally scaled to deal with 3D signals.

References

- [1] I. Biederman. Human image understanding: recent research and theory. Academic Press, Rosenfeld, 1985.
- [2] U. Cavallaro and G. Christofori. Cell adhesion and signaling: implications for tumor progression. volume 11, pages 118–132. Nature Rev. Cancer, 2004.
- [3] H. Chang and B. Parvin. Segmentation of three dimensional cell culture models from a single focal plane. volume 1. Proc. of Int. Symp. on Visual Computing, 2006.

- [4] R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in pictures. volume 15, page 1115. *Comm. ACM*, 1972.
- [5] G. Guy and G. Medioni. Inferring global perceptual contours from local features. volume 20, pages 113–133. *Int. Journal of Computer Vision*, 1996.
- [6] L. Loss, G. Bebis, M. Nicolescu, and A. Skurikhin. An iterative multi-scale tensor voting scheme for perceptual grouping of natural shapes in cluttered backgrounds. volume 113, pages 126–149. *Journal of Computer Vision and Image Understanding (CVIU)*, 2009.
- [7] G. Medioni and S. B. Kang. *Emerging topics in computer vision*. Prentice Hall, 2004.
- [8] R. Murphy. Automated interpretation of subcellular location patterns. volume 1, pages 53–56. *IEEE Int. Symp. on Biomedical Imaging*, 2004.
- [9] B. Parvin, Q. Yang, J. Han, H. Chang, B. Rydberg, and M. H. Barcellos-Hoff. Iterative voting for inference of structural saliency and characterization of subcellular events. *IEEE Trans. on Image Processing*, 2007.
- [10] N. Prigozhina, L. Zhong, E. Hunter, I. Mikic, S. C. D. Roop, M. Mancini, D. Zachariasa, J. Price, and P. McDonough. Plasma membrane assays and three-compartment image cytometry for high content screening. volume 5, pages 29–48. *Assay Drug Dev. Technologies*, 2007.
- [11] Q. Yang and B. Parvin. Harmonic cut and regularized centroid transform for localization of subcellular structures. volume 50, page 469476. *IEEE Trans. on Biomedical Engineering*, 2003.