Probabilistic segmentation of horizontal cells

Vebjorn Ljosa and Ambuj K. Singh
University of California, Santa Barbara

Abstract

We propose a novel segmentation algorithm that can segment confocal micrographs of horizontal cells. A unique aspect of the algorithm is that it produces a probabilistic mask, *or* pmask, *which gives for each pixel the probability that the pixel belongs to the cell. The pmask enables analysis modules to compute confidence values for their results.*

1 Introduction

The role of microscopy, a cornerstone of many fields of biology, is changing. Rather than a likeness subject to visual inspection, the micrograph is becoming a quantitative measurement subject to formal analysis. Together with high-throughput acquisition techniques, this change in mind set may bring data-driven research, and the spectacular success it has had in genetics, to other fields. Morphological neuroscience is a prime candidate for this transformation.

One approach to understanding the vast complexity of the brain is to study the retina, which is part of the central nervous system, but is accessible and contains relatively few classes of cells organized in well-defined layers. The electrical signals generated by the photoreceptors in response to light go through the bipolar cells and ganglion cells before they are finally passed through the optic nerve into the rest of the brain. There are other types of neurons in the retina in addition to these three, however. In particular, *horizontal cells* are fairly flat neurons located in the outer plexiform layer, where the photoreceptors and bipolar cells meet [2]. As their name indicates, they provide connections that are horizontal, i.e., perpendicular to the main direction that signals are conducted. Figure 1 shows two horizontal cells.[1]

Horizontal cells have many long, thin neurites that branch several times. Each image shows several cells, their neurites intertwined, and also neurites from numerous cells the bodies of which are outside the field of view. Measurement and analysis is difficult because it is hard to determine which parts of the image belong to a certain cell.

[1] All images have had their dynamic range adjusted so that the morphology of the cells would be visible in print.

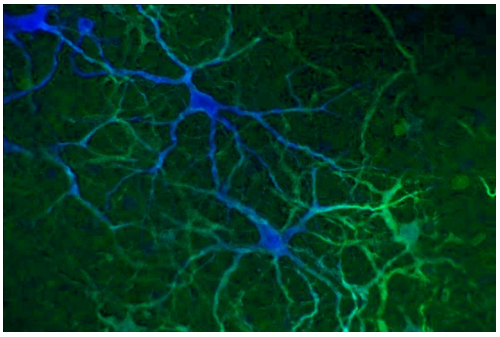

Figure 1. Confocal micrograph of two horizontal cells in a detached cat retina, labeled by neurofilament (green) and calbindin (blue).

In this paper, we propose a novel segmentation algorithm that can successfully segment confocal micrographs of horizontal cells. A unique aspect of the algorithm is that it does not output a binary mask, i.e., a set of pixels that definitely belong to a particular cell. Instead, the algorithm produces a *probabilistic mask*, or *pmask*, which gives for each pixel the probability that the pixel belongs to the cell. The *pmask* is important because it enables an analysis to compute a confidence value for its result.

2 Probabilistic segmentation

The proposed segmentation algorithm is based on repeated random walks. Starting in the center of the cell, the algorithm takes a sequence of steps. In each step, it randomly moves to one of the eight neighboring pixels, but the decision is biased by the intensities of the neighbors so that the next step is more likely to visit a bright neighbor pixel. After each step, the intensity of the pixel in the pmask that corresponds to the current location is incremented. Then, with a small probability c, the walk "restarts," i.e., returns to the center of the cell. The algorithm stops when the normalized pmask converges.

The restarts mitigate the effect of crossing into another cell—an event that is bound to happen eventually when the neurites of two cells intersect. Restarts solve the problem because a walk that strays into the wrong cell will soon return to the original starting point. Because most cells will

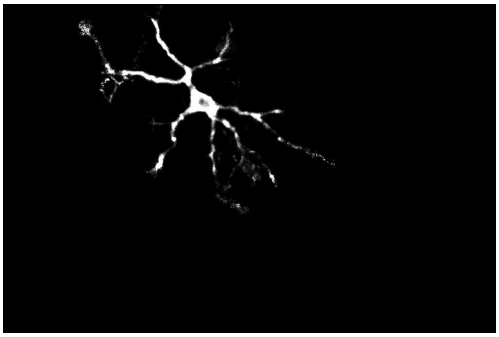

Figure 2. Probabilistic mask for a horizontal cell.

never cross the bridge, the original cell will be visited much more than the other.

It may appear that crossing into and walking around in the wrong cell for even a little while is a misfortune. However, this behavior is actually a blessing in disguise. The reason is that a connection between neurites of different cells can look quite similar to a branching point in a neurite. If the algorithm had to make a definite choice between staying in the original cell or crossing into the other, a wrong choice would have huge impact on the segmentation result and consequently on the analysis results. The randomized solution avoids this by visiting each pixel proportionally to how well it is connected to the original cell. An analysis algorithm can recognize that there is doubt about the extent of the cell and conclude, for instance, that there is a 0.8 and 0.2 probability that the neurite is 30 μm and 40 μm long, respectively.

Figure 2 shows the pmask for one of the cells in Figure 1. Pixels far from the center are generally visited less. This makes sense from a probabilistic point of view, for a path from the center to a pixel far away has more opportunities for making mistakes and crossing into other cells, so we can be less certain that the pixel actually belongs to the original cell. Notice also that wide neurites are followed more often—again consistent with what we expect.

Although a simulation-based implementation of the algorithm, as described above, is sufficiently efficient (5 s per cell for a 768-by-512-pixel image), it is interesting to note that the pmask can also be computed by solving an eigenvector problem. Each iteration of the walk can be written

$$\boldsymbol{x} := (1-c)\boldsymbol{P}\boldsymbol{x} + c\boldsymbol{s}. \tag{1}$$

Here, $\boldsymbol{x}$ is the pmask, P the transition matrix, c the restart probability, and s a vector indicating the center of the cell.

Çamoğlu et al. [1] show that Eq. (1) converges to the stationary probability distribution of the Markov chain with transition matrix $\boldsymbol{Q} = \{Q_{ij}\}$, defined by

$$Q_{ij} = \begin{cases} (1-c)P_{ij} & \text{if } s_i \neq 1 \\ (1-c)P_{ij} + c & \text{if } s_i = 1. \end{cases} \tag{2}$$

At convergence, $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{x}$. Because Q is column-normalized, its largest eigenvalue is 1, so $\boldsymbol{x}$ is the corresponding eigenvector. The eigenvector problem can be computed quickly because $\boldsymbol{P}$ and $\boldsymbol{Q}$ are very sparse: each row has only eight non-zero elements, one for each possible next step.

3 Analyzing probabilistic masks

The pmask facilitates quantitative analysis of the cell's morphology. For lack of space, we only discuss a simple analysis: how to measure the thickness of a neurite at a certain point q. Let $\boldsymbol{m}$ be a sequence of values read from the pmask along a line perpendicular to the neurite at q, and let m_q be the element of $\boldsymbol{m}$ read from the pmask at q. Assume that the value of an element m_i in $\boldsymbol{m}$ is proportional to the probability p_i that the pixel belongs to the cell, i.e., $p_i = \alpha m_i$.

One way to measure thickness would be to count the number of consecutive elements in $\boldsymbol{m}$ that exceed a threshold τ and include m_q. This will compute a thickness, but as a single number, with no information about how much confidence we can place it it or how sensitive it is to τ.

Instead we propose to calculate the probability that the neurite is w pixels thick, for each possible w. The probability that the neurite is w pixels thick, starting with element s in $\boldsymbol{m}$, is

$$P(w,s) = (1-\alpha m_{s-1}) \prod_{i=s}^{s+w-1} \alpha m_i (1-\alpha m_{s+w}). \tag{3}$$

We can then compute the probability $P(w)$ that the neurite is w pixels thick by trying all relevant start positions:

$$P(w) = \max_{s=q-w+1}^{q} P(w,s) \tag{4}$$

4 Conclusion

The proposed algorithm, based on repeated intensity-biased random walks, shows promise for segmenting horizontal cells and similar structures, and the resulting probabilistic masks enable analyses that yield probabilistic values and to indicate the confidence that can be placed in their results. Future work will compare to other seeded segmentation algorithms and develop analysis and mining techniques for probabilistic masks.

Acknowledgements. The authors thank Geoffrey P. Lewis from the laboratory of Steven K. Fisher at UCSB for providing micrographs. This work was supported in part by grant no. ITR-0331697 from the NSF.

References

[1] O. Çamoğlu, T. Can, and A. K. Singh. Integrating multi-attribute similarity networks for robust rrepresentation of the protein space. *Bioinformatics*, 22(13):1585–1592, 2006. The proof is in the supplement.

[2] J. E. Dowling. *The Retina: An Approachable Part of the Brain*. Belknap, 1987.