2020

Examining Artifacts of the Watershed Segmentation

Emily Jo Armitage

Examining Artifacts of the Watershed Segmentation

A Thesis

Presented to

the Faculty of the College of Natural Sciences and Mathematics

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

Emily Jo Armitage

June 2020

Advisor: Dr. Dinah Loerke

Author: Emily Jo Armitage
Title: Examining Artifacts of the Watershed Segmentation
Advisor: Dr. Dinah Loerke
Degree Date: June 2020

Abstract

The watershed segmentation is an algorithm used to systematically track cell intercalary behaviors during germ band extension of the *Drosophila* embryo. Neighboring cells share a contracting vertical interface, called a T1, which continues contracting to a single point, a T2, and extending in the horizontal direction to create what is called a T3 interface (Fig. 1). Additionally, higher order vertices called rosettes occur when five or more cells meet at a common vertex. Simulated T2 events demonstrate that cell angle and not noise level in the image contributes to the incorrect detection of artifactual T1s in more acute angled cells and T3s for obtuse angled cells. Short T1 simulations show a systematic overestimation of T1 lengths detected by the watershed segmentation. Order three vertex simulations show central vertex displacement is biased toward the smallest angled cells. Rosette simulations of order 5 to 11 provide a working definition of rosettes in the context of the watershed segmentation in terms of short interface frequency, length, and radius of artifactual vertices.

Acknowledgements

I wish to show my deepest gratitude to my advisor, Dr. Dinah Loerke, for welcoming me into her lab as well as for her support and graciousness as I learned to think quantitatively and code in MATLAB.

I am forever indebted to my lab mates Tim Vanderleest, Roopa Madhu and Noah De Leeuw for all of the time they spent patiently explaining concepts to me, sending me code, and working through bugs with me.

I want to thank my committee: Dinah Loerke, Todd Blankenship, Michelle Knowles, and my committee chair Natasha Dobrinen.

I would finally like to thank my friends and family for their unending support.

Table Of Contents

List of Figures

Introduction

In the early *Drosophila* embryo, germ band extension (GBE) occurs to elongate the body axis during gastrulation (Fig. 2A) (Irvine & Wieschaus, 1994). This results from a remodeling of cell topologies driven by cell intercalation, a phenomenon where cells insert themselves between previously neighboring cells (Fig. 2B) (Butler, et al., 2009; Irvine & Wieschaus, 1994; da Silva & Vincent, 2007; Wang, et al., 2017). Neighboring cells share a contracting vertical interface that lies along the anterior-posterior (AP) axis, referred to in developmental biology as a T1 event, which continues contracting to a single point, until four cells share a common vertex, a T2 event. These previously separated cells create a horizontal interface between them along the dorsal-ventral (DV) axis, and this is called a T3 event (Fig. 1, 2C) (Bertet, et al., 2004; Blankenship, et al., 2006; Collinet, et al., 2015; Irvine & Wieschaus, 1994; Jessica & Fernandez-Gonzalez, 2016). In polymer physics, this is collectively referred to as a topological T1 process and results in a systematic contraction of the epithelium along the DV axis and a perpendicular elongation along the AP axis of the embryo (Weaire & Hutzler, 2000). Higher order intercalation

events deemed rosettes also occur, and form when successive vertical interfaces contract to bring five or more cells together to a single vertex, which finally resolve into many elongating horizontal interfaces (Fig. 2C) (Blankenship, et al., 2006; Irvine & Wieschaus, 1994).

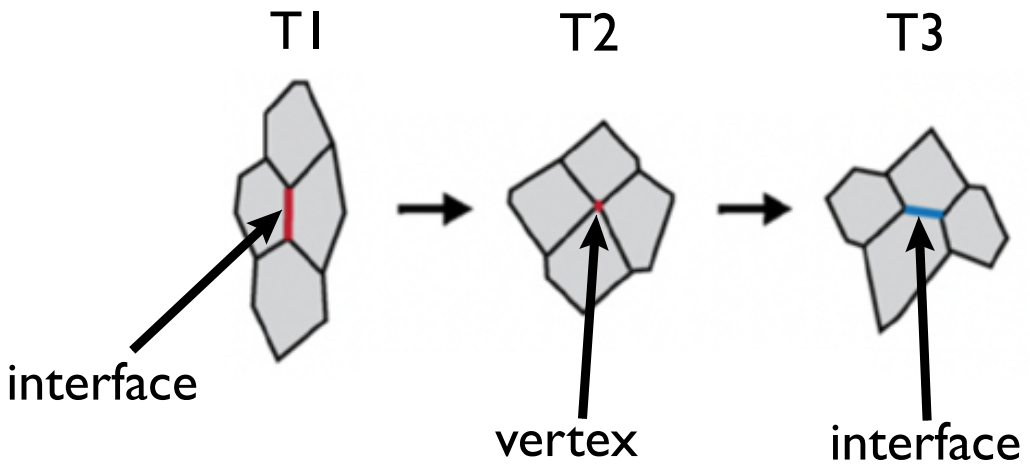

Figure 1: Illustration of topological T1 process. During cell intercalation, neighboring cells share a vertical T1 interface, which contracts to a single point, and four cells meet at a single vertex, a T2. This then resolves into a horizontal T3 interface. Illustration modified from (Vichas & Zallen, 2011).

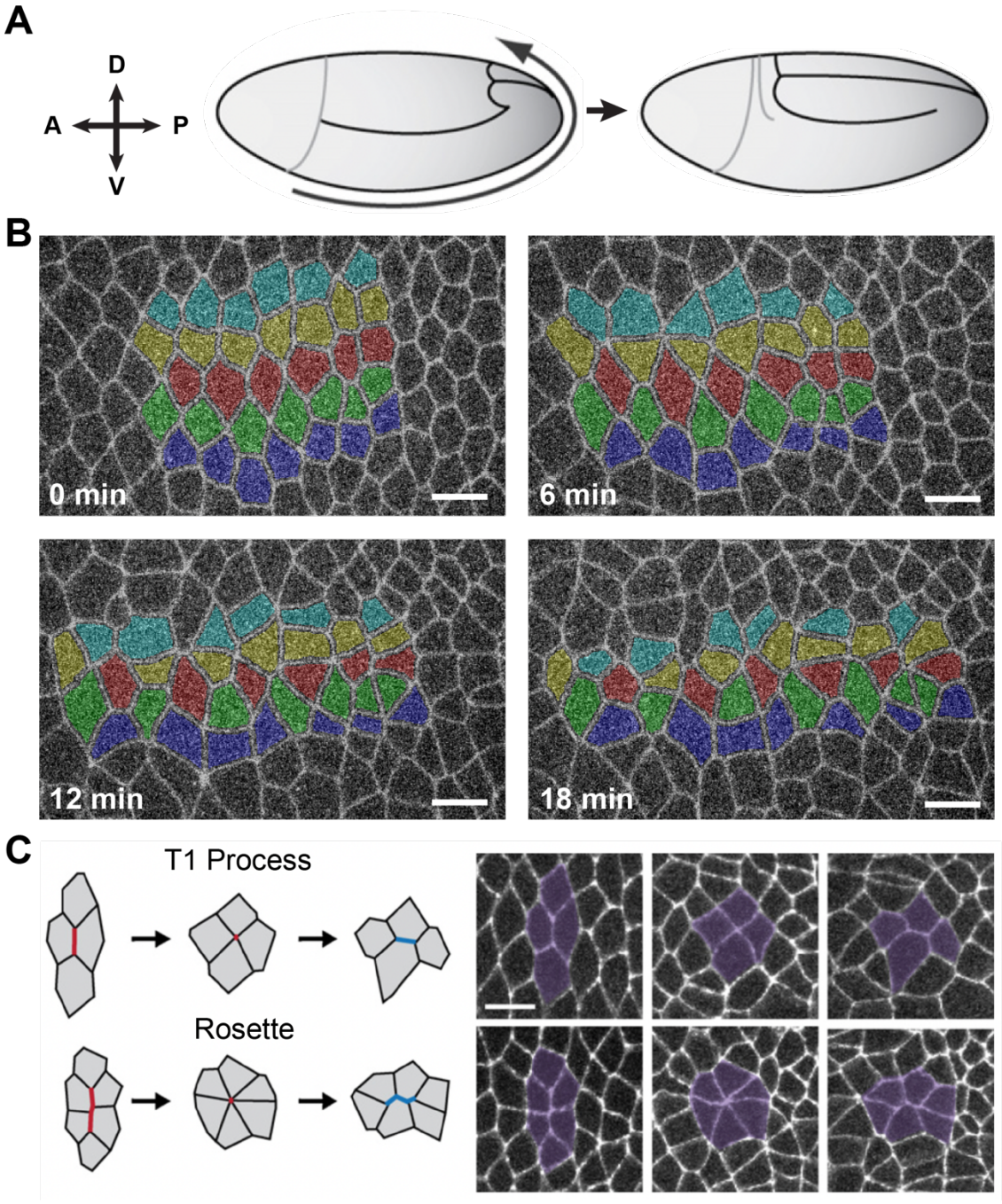

Figure 2: GBE occurs via cell intercalation. (A) Graphic of the germband elongating towards the posterior end and folding upward and around to the dorsal side of the embryo. Anterior, posterior, dorsal, and ventral are labeled A, P, D, and V respectively. Illustration is modified from (Vichas & Zallen,

2011).(B) Intercalation shown through time lapse images of the germband with rows of cells tracked and colored, by Dr. Timothy Vanderleest. (C) Illustration of a T1 process (top) and a rosette (bottom). Illustration courtesy Dr. Timothy Vanderleest and is modified from (Vichas & Zallen, 2011). Scale bars are 10 μm.

In order to study the molecular mechanisms that are directing these topological transitions, we rely on image analysis tools. These tools allow us to analyze large numbers of cells in a reproduceable fashion rather than relying on the biased segmentation of one individual. One fundamental image analysis tool is the watershed segmentation, which does the first step of cell fate analysis by identifying cell interfaces (Fernandez-Gonzalez & Zallen, 2011; Jewett, et al., 2017; Leung & Fernandez-Gonzalez, 2015).

The watershed segmentation, first introduced in the late 1970s, is a powerful tool for segmenting grayscale images (Beucher & Lantuéjoul, 1979). It has origins in mathematical morphology (Serra, 1982), and has been developed in terms of accuracy and efficiency since it emerged (Beucher, 1992; De Smet & Pires, 2000; Meijster & Roerdink, 1995; Meyer, 1994; Vincent & Soille, 1991).

This algorithm is used to systematically interrogate interface dynamics and intercalary behaviors in microscopy images by systematically differentiating between objects of interest and the background. It treats a grayscale image as a topographic map where pixel brightness represents elevation, and each regional

intensity minimum on the image is considered a "catchment basin" holding water. If the water levels were to rise, the points at which the water from neighboring catchment basins meet become the watershed segmentation lines (Fig. 3). Upon finding the areas where these theoretical water sources would meet, this algorithm segments the image according to each regional minimum.

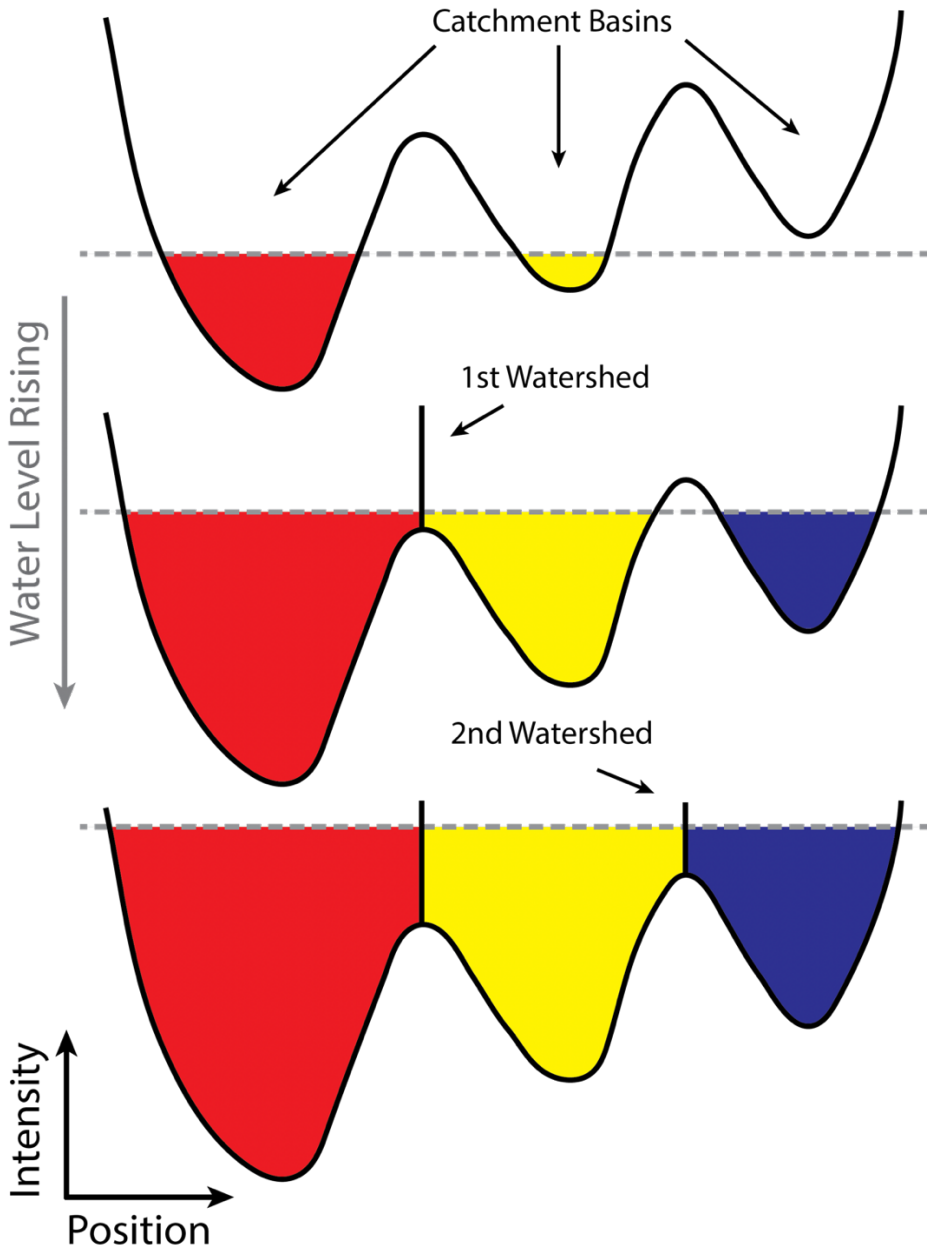

Figure 3: One-dimensional illustration of the Watershed Transform. The black curve represents the image intensity, where the minima represent "catchment basins". When the gray dashed line representing water rises, these points where water from adjacent catchment basins meet becomes the watershed segmentation lines. In this example the curve is segmented into the three regions indicated by the colors red, yellow, and blue. Illustration courtesy of Dr. Timothy Vanderleest.

Watershed segmentation is used to differentiate cell interfaces from the background in fluorescence microscopy images and movies. As proteins in the cell membranes of embryos are tagged with fluorescent probes, the resulting microscopy images are brightest at the cell interfaces (Fig 4A). Because those pixels in the image along the cell interfaces have the highest intensity, the watershed segmentation applies segmentation lines along those pixels, differentiating them from the local intensity minima inside the cells (Fig 4B). The result is an image with lines one pixel in width representing all locations in the image where two cells are touching one another, a cell interface, or multiple cells are touching one another, a vertex (Fig. 4B, subset). The relevant information from the segmented images such as vertex position and interface length can then be extracted and analyzed.

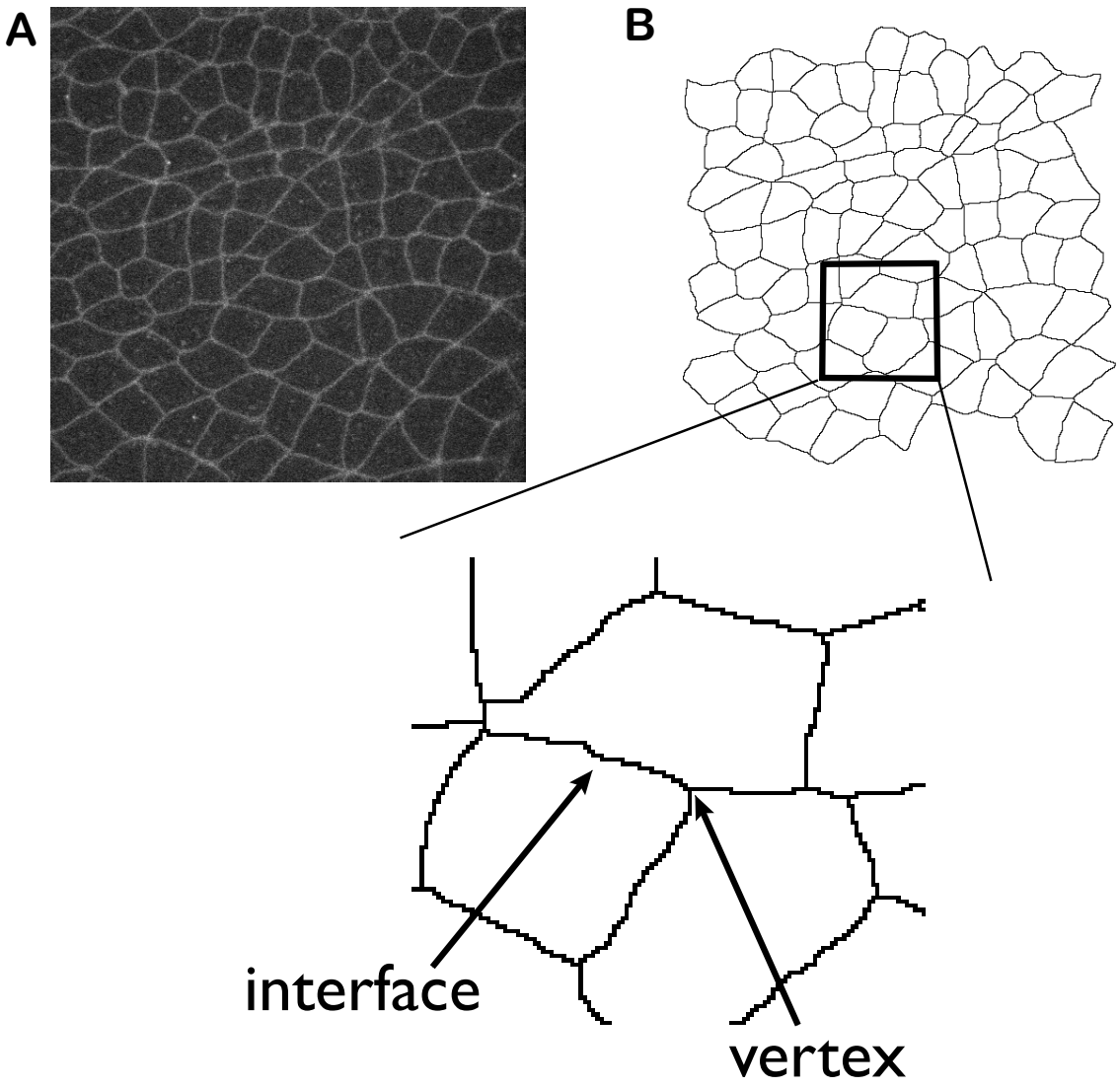

Figure 4: Raw data and the result of the watershed transform on that image. (A) Spinning disc confocal fluorescence microscopy image of mCherry tagged Gap43 in the *Drosophila* embryo. (B) Segmented image yielded from watershed segmentation. A and B courtesy of Dr. Timothy Vanderleest. Zoomed in portion of panel B illustrates examples of an interface and a vertex.

Now that cell interfaces can be segmented from cell areas, we can apply this analysis to long time resolved movies in order to track interfaces over time.

To examine the dynamics of interface transitions, we look at the length of these interfaces, which is defined as the distance between vertices.

Upon tracking interface lengths in movies demonstrating T1 to T3 processes in GBE of the *Drosophila* embryo, the length traces show steadily contracting T1 interfaces. As the interface length approaches zero, a T2 event, the length traces near this supposed T2 time point are unstable, as the watershed segmentation measures alternating positive and negative values as if the interface was switching back and forth between a vertical T1 to a horizontal T3 interface before steadily expanding in length in the horizontal direction into a T3 event (Fig. 5). This observation led to the question: is this phenomenon an instability of the vertex itself or the watershed segmentation algorithm? In order words, is this an inherent biological process, or is this an error with the way the vertex is being defined by the watershed segmentation algorithm?

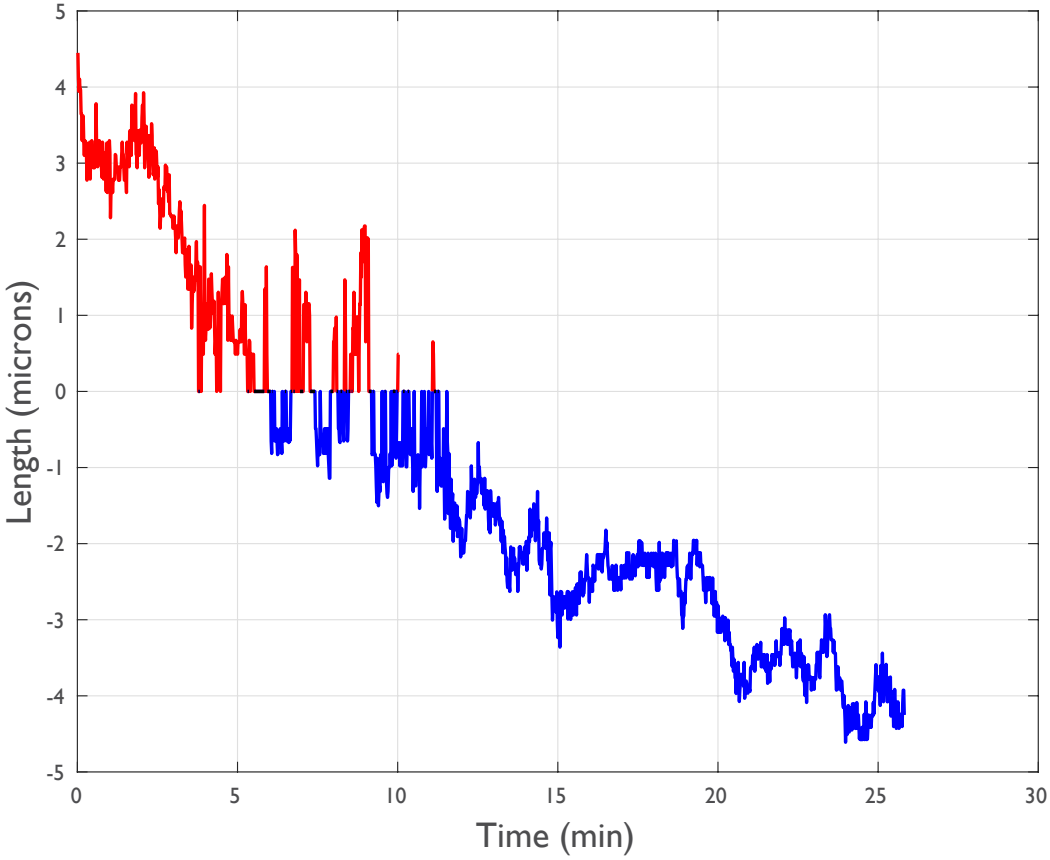

Figure 5: Interface length trace of *Drosophila* cells during a T1 to T3 transition of GBE. Red (positive) length values represent a vertical T1 interface and blue (negative) length values represent a horizontal T3 interface. Surrounding length zero, a supposed T2 event, the length trace is highly unstable, moving from relatively high positive to relatively low negative values before decreasing uniformly as time passes and the interface extends in the horizontal direction. Courtesy of Dr. Timothy Vanderleest.

In order to interrogate the source of this error, I created simulated images of T2 and T1 events with varying levels of noise and internal angles of the cells as they meet at the common vertex. I then subjected them to the watershed segmentation and analyzed the manner in which the algorithm

segmented the cells in order to identify any possible error in interface identification and measurement.

This experimentation led to the next question, which was: What is the precision and accuracy of watershed-based vertex measurements under this experimental framework? To explore this question, I simulated three-cell junctions and measured watershed identified vertex position displacement from true vertex position as a function of internal cell angle.

Furthermore, higher-order rosette structures composed of five or more cells meeting at a common vertex have been implicated in linking local cell interactions to global tissue reorganization during morphogenesis (Blankenship, et al., 2006). Although rosettes are easily detectable by eye, the watershed segmentation algorithm only allows for an accurate visualization of third and fourth order vertices, or cell junctions in which only three or four cells share a common vertex. Due to the square nature of pixels, vertices with an order of 5 or higher are broken down into a series of short interfaces. This precludes an accurate systematic identification of rosette events, requiring all rosette identifications to be manually done by eye, introducing the possibility of bias and error. For this reason, I created rosette simulations where at least five cells shared a common vertex and analyzed the manner in which the watershed segmentation broke down the central vertex into a series of short interfaces.

This was done in order to propose a working definition of a rosette in the context of the watershed segmentation's output to aid in systematic detection.

Methods

T2 simulations

In order to accurately emulate experimental data, I analyzed the intensities of cell interfaces and background found in real data. This was done by analyzing an image of an mCherry tagged Gap43 *Drosophila* embryo and its accompanying watershed segmented image (Fig. 6A, 6B). I first determined the width in pixels of the cell interfaces by performing a Euclidean distance transform on the segmented image (Fig. 6C). For each pixel in the image, this process assigns a number that is the distance between that pixel and the nearest nonzero pixel. This yields a matrix with cell interface lines represented by zeros, each pixel surrounding those zeros are ones, and each around those ones are twos, etc. I was able to use that information to index pixel positions on the experimental image in descending order from the center of the interfaces.

I next interpolated the distance transform values against the intensity values of the image using a gaussian weighting of distances, due to the diffraction limit of light emitted by a fluorophore as observed by a lens with a circular aperture (Fig 6D). The profile of intensities emitted by each

fluorophore on the plasma membrane of these cells is characterized by the Airy Disc, with a bright central region that falls to zero with increasing distance from the center and a series of concentric rings that surround it. An alternative measure of this is to ignore the small outer rings, and to approximate the central lobe to a gaussian profile. I used a bin size of 0.5793 pixels, consistent with the diffraction limit of light of the spinning disc confocal microscope used to capture the image that has a numerical aperture of 1.4 and the emitted wavelength of the mCherry fluorophore, 650 nm. I calculated the standard deviation of this distribution, and when multiplied by three to encompass three standard deviations from the mean (or 99.7% of the data), was calculated to be 6.9607 pixels. This represented the distance from the center of the interface to the edge of the interface, meaning that the full width of the interface was approximately 14 pixels. With that information I was able to measure the average intensity within that 14-pixel range, which was 3279.5 arbitrary fluorescence units. Excluding the range of values that lied within the interfaces, I calculated the average background fluorescence to be 2,765.2 units.

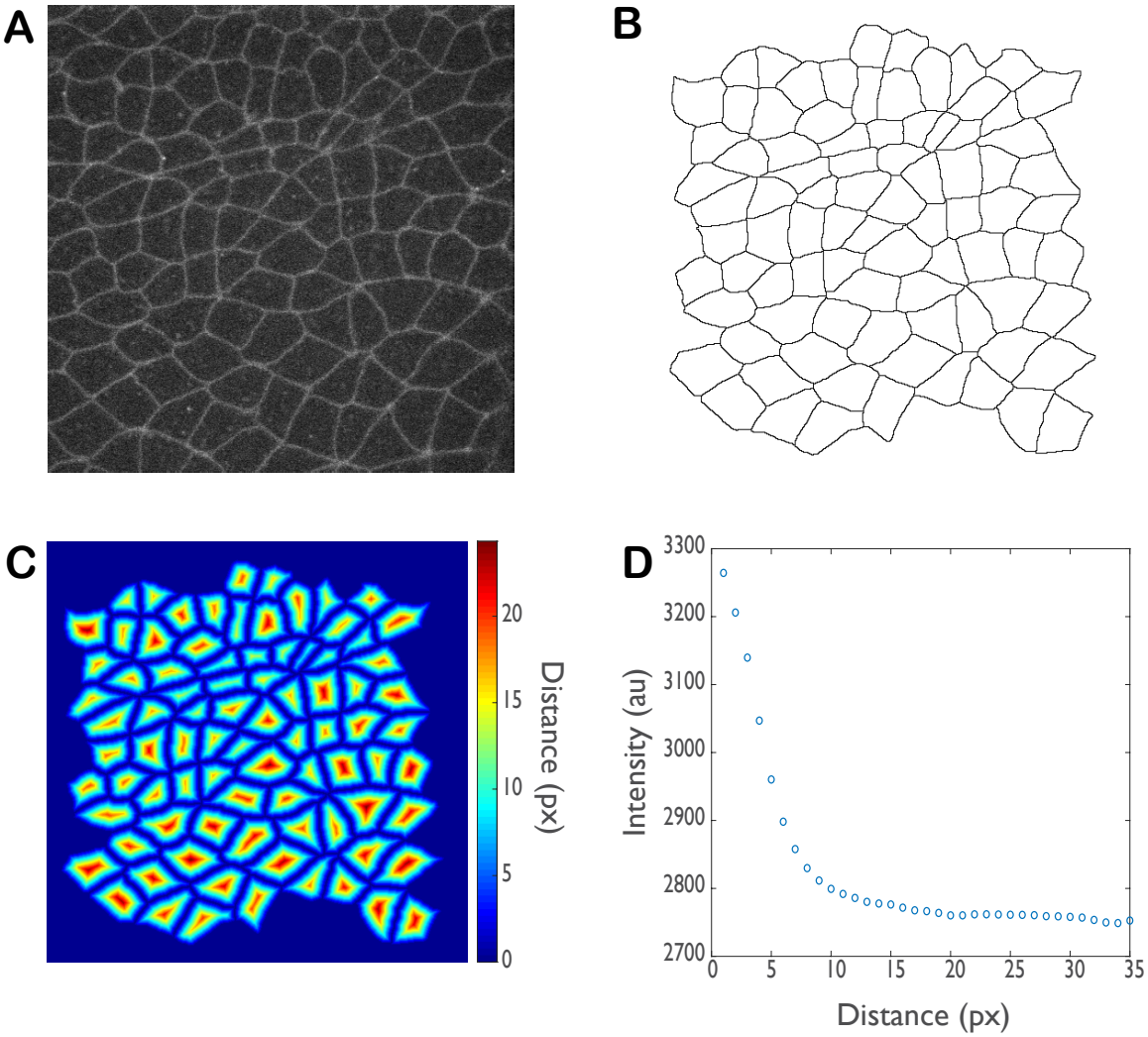

Figure 6: Scaling simulations to experimentally relevant values. (A) Spinning disc confocal fluorescence microscopy image of mCherry tagged Gap43 in the *Drosophila* embryo (B) Segmented image yielded from watershed segmentation. A and B courtesy of Dr. Timothy Vanderleest. (C) Distance transform of B. (D) Plot of interpolated distance transform values against the intensity values of the image using a gaussian weighting of distances, bins of 0.5793.

Once I had these experimentally relevant values, I created binary images of polygons that mimicked four cells meeting at a common vertex (Fig. 8A). The lines in the image represented cell interfaces; each pixel composing these

lines had an intensity value of one, and all background pixels had an intensity of zero. I then scaled the simulated interface and background to the experimentally relevant intensity values calculated above.

After creating this polygonal representation of a T2 event, I tilted the images ten, twenty, thirty and forty degrees to avoid any privileged situation for the watershed segmentation (Fig. 7). If an interface line perfectly lined up with one column or row of pixels, this would be a particularly stable configuration for the watershed segmentation to draw from when segmenting the image and is unrealistic in terms of configuration of interfaces in actual data, which are distributed in an isotropic manner across both stable and unstable configurations.

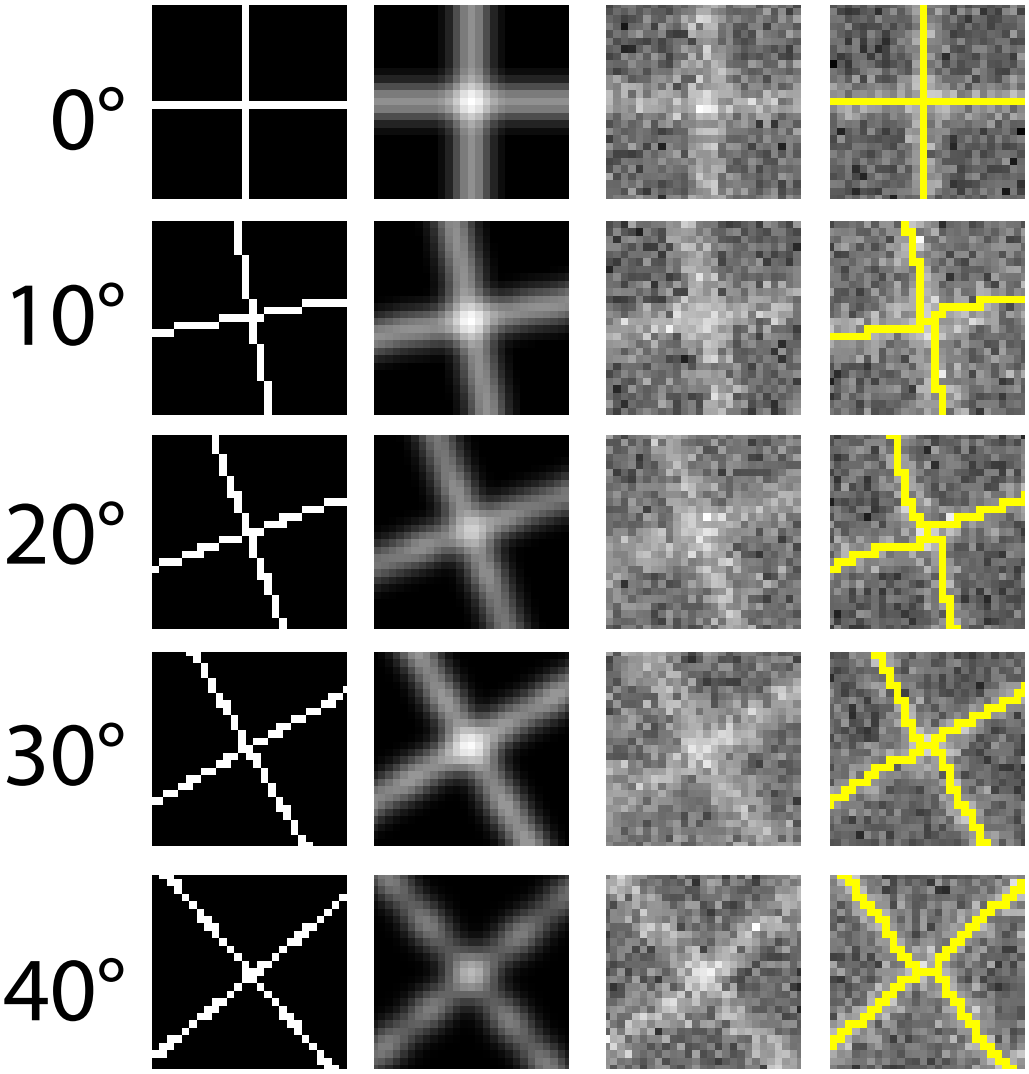

Figure 7: Tilting simulated images to avoid privileged situations for watershed segmentation. First column shows polygonal image simulating a central vertex of a T2 event with internal angles of 90 degrees, at various degrees of tilt. Second column shows the implementation of a gaussian filter, third column shows addition of random noise, and rightmost column shows the result of the watershed segmentation overlaid on top of simulated T2 event.

Next, I added a gaussian filter to my binary image that mimicked the blurring effect due to the diffraction limit of light as described above (Fig. 8B). The standard deviation of interface fluorescence obtained from the

interpolation procedure above gave me an experimentally relevant sigma value to input in the imgaussfilt command in MATLAB. This filter expanded the interface to a similar width of the interfaces in the real data.

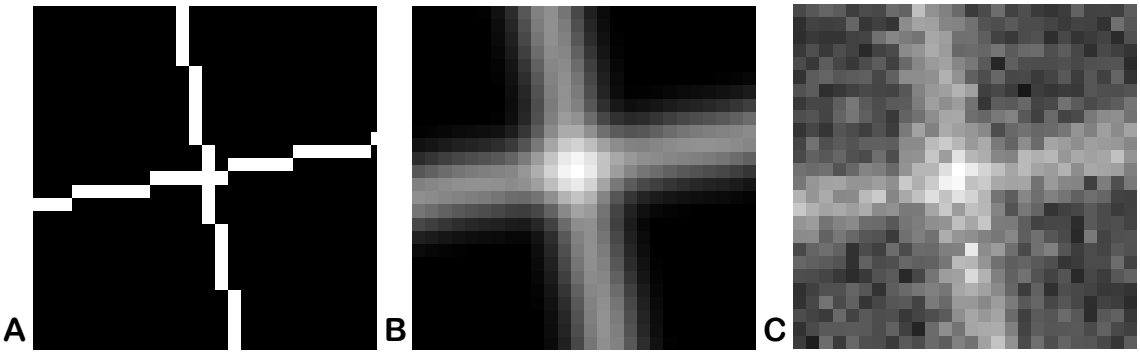

Figure 8: Addition of gaussian filter and noise to emulate real data in T2 simulations. (A) Polygonal simulated T2 event tilted 10 degrees. (B) Addition of gaussian filter to image. (C) Addition of gaussian random noise to image.

In order to incorporate the background noise in the measured fluorescence signal into the simulated image, I added random gaussian white noise with the standard deviation of background intensity in the Gap43 image (Fig 8C). This image now accurately simulated experimental data (Table 1).

	Gap43 image	Simulation
Mean interface intensity	3.2795e+03	3.2789e+03
Mean background intensity	2.7652e+03	2.7637e+03
STD of background intensities	164.9333	164.0479
STD of interfaces	262.2585	253.4403
Signal to noise ratio	1.1860	1.1864

Table 1: Comparison of experimental data to simulation. Intensity units are arbitrary fluorescence units. Signal to noise ratio was calculated as the ratio between mean interface intensity and mean background intensity. STD: standard deviation.

Next, I created this image 100 times using 100 different iterations of random noise, then subjected each simulation to segmentation.

The first step in the watershed segmentation is to select the seeds for segmentation that designate local minima. The seeds were selected using the point seeding method, where small circular spots several pixels in diameter were manually placed within the cells, and polygon mask seeding to exclude the background area surrounding the four cells of interest (Fig. 9A). This was manually done for the first frame of the 100-frame collection, and these seeds were propagated to the other 99 frames that differed only in the new iteration of random noise, and segmentation commenced (Fig 9B).

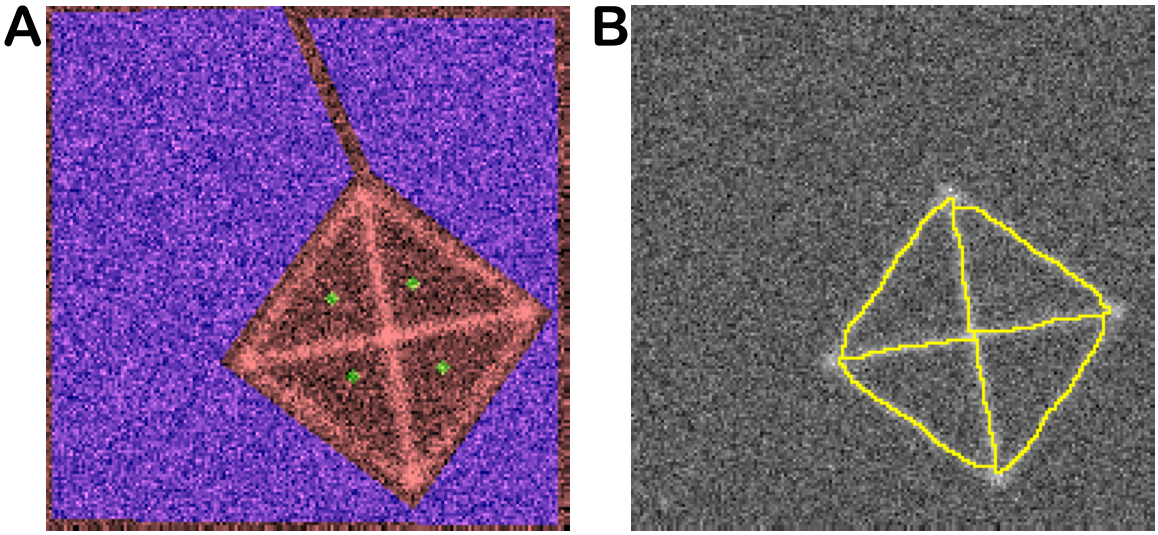

Figure 9: Initializing seeds and mask and segmentation output. (A) The green points are the locations where local minima were designated for the watershed segmentation algorithm to recognize. The blue area is the mask applied to exclude that area from segmentation. (B) Output of segmentation based on designated seeds and mask.

The segmented representation of the simulated T2 event yielded vertex position and interface lengths in units of pixels, which were stored together with cell-cell and vertex-vertex connectivity matrices. Vertical interfaces were recorded as positive values, and horizontal interfaces were recorded as negative values to distinguish the two classes of artifacts. This data simulation process was repeated to emulate data with different levels of background noise: signal to noise ratios of 1.0729 and 1.3831, as well as different internal angle measurements of cells in combinations ranging from 30 degrees to 140 degrees by intervals of 10 degrees (Fig. 10).

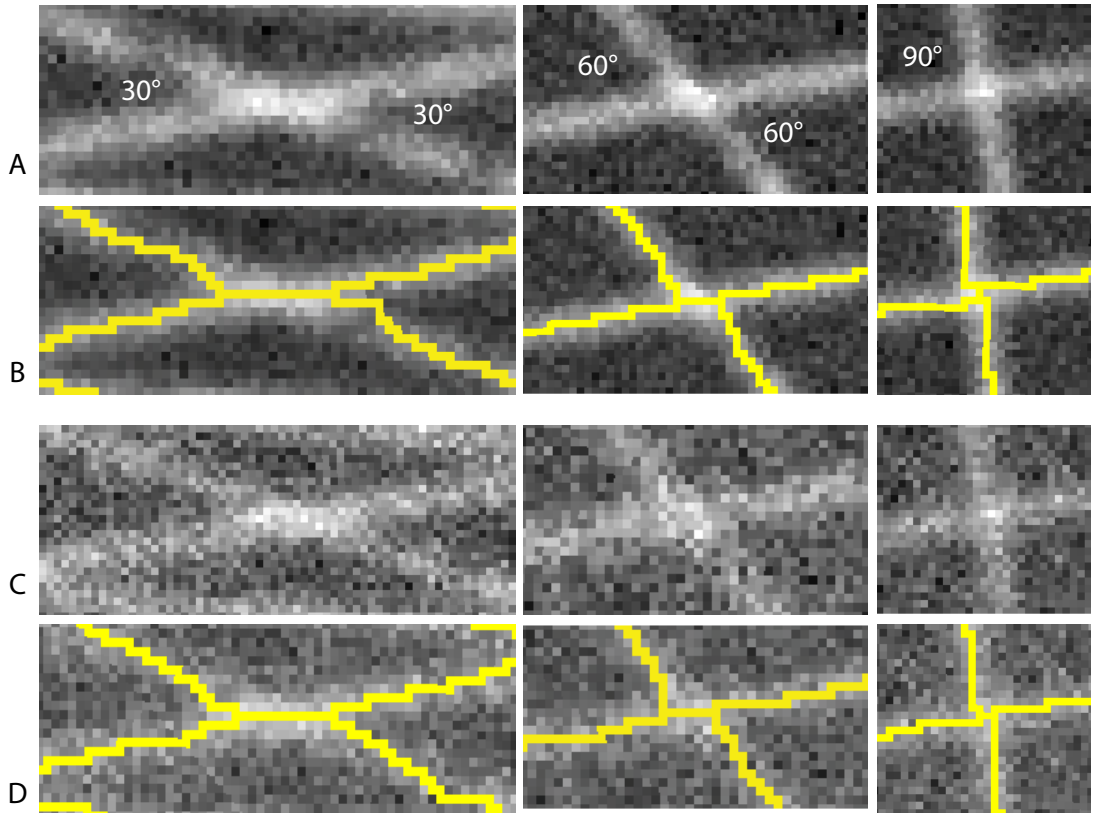

Figure 10: Examples of T2 simulations with varying levels of noise and internal angle measurements. Rows A and B were simulations with a SNR of 1.864, consistent with real data, and C and D had a lower SNR of 1.0729. Watershed segmentation lines overlaid in yellow.

Short T1 simulations

I created short T1 simulations by adding a three-pixel interface between one pair of oppositely juxtaposed cells (Fig. 11). Simulations were tilted 10 degrees. All methods used to scale simulation to experimentally relevant

conditions (with a signal to noise ratio of 1.864), segment and extract data as described in the T2 simulations were used.

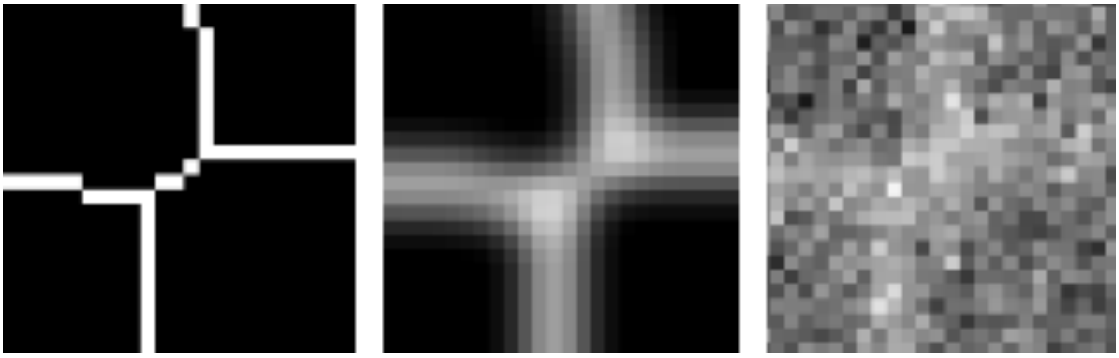

Figure 11: Simulated image of short T1 interface 3 pixels in length. Left represents the skeletonized image, middle shows gaussian filtering and right shows addition of random noise.

Three cell vertex simulations

I created order-three vertex simulations (Fig. 12) with internal angles in intervals of 20 degrees in the range 80 to 160 degrees. Simulations were tilted 0, 10, 20 30 and 40 degrees. All methods used to scale simulation to experimentally relevant conditions (with a signal to noise ratio of 1.864), segment and extract data as described in the T2 simulations were used.

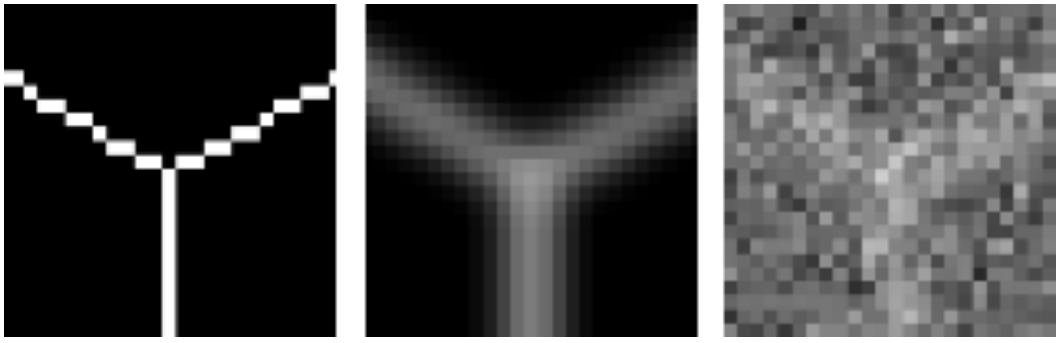

Figure 12: Simulated three-cell junction. Left represents the polygonal image, middle shows gaussian filtering and right shows addition of random noise.

Rosette Simulations

I simulated rosettes composed of 5 to 11 cells with equal internal cell angles (Fig. 13). All methods used to scale simulation to experimentally relevant conditions (with a signal to noise ratio of 1.864), segment and extract data as described in the T2 simulations were used.

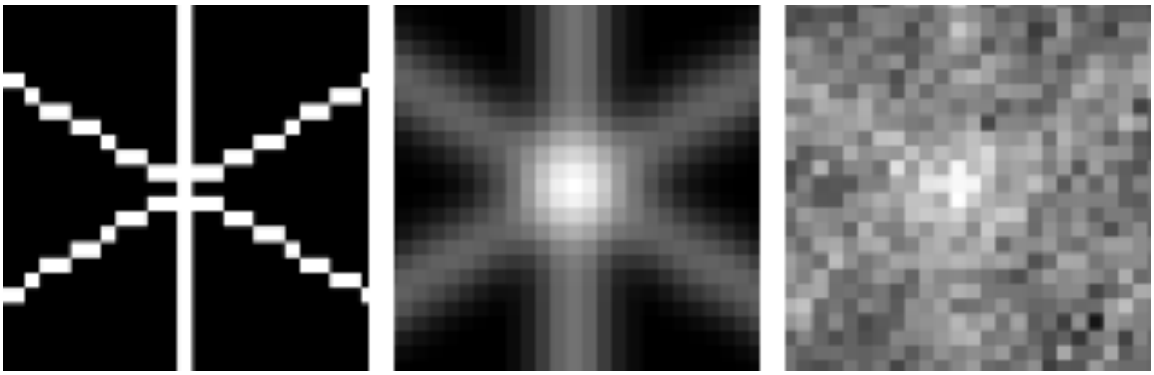

Figure 13: Simulated 6-cell rosette with equal internal angle measurements. Left represents the polygonal image, middle shows gaussian filtering and right shows addition of random noise.

Results

Internal cell angle has significant influence on artifactual interface length in T2 simulations, not level of background noise

Although all simulations were done with all four cells meeting at a common vertex, the watershed segmentation often detected a T1 interface between either cells C and D or a T3 interface between cells A and B (Fig. 14). The length of that artifactual interface was averaged over each 100-frame collection, assigned a positive value for a vertical T1 interface between cells C and D and a negative value for horizontal T3 interfaces between cells A and B. The parameters tested, level of noise and cell angle, each influenced the result of the watershed segmentation. Different iterations of noise can result in differently oriented artifactual interfaces, with one realization of noise resulting in an interface separating cells A and B, and another realization of random noise resulting in the separation of cells C and D (Fig. 14). Additionally, the cell angle influences both the length and direction of the artifactual interface. The more acute oppositely juxtaposed cells are, there is a higher likelihood of a long artifactual interface that separates the neighboring cells (Fig. 15).

There was a much stronger relationship between artifactual interface length and internal cell angle than there was with interface length and level of noise in the simulated image (Fig. 16). In these simulations with equivalent opposite cell angles, regardless of the level of noise added to the simulations, the highest artifactual interface length is a result of more acute internal angles in oppositely juxtaposed cells. The error in interface measurement can be quite large, as seen in the artifactual interface length measurement of approximately 16 pixels in the simulation with opposite angles of 30 degrees although the true length of the interface was zero. However, as the simulations approach a more isotropic balance of cell angle, when all cells are the same measure of 90 degrees, the average artifactual interface length was very near the true length of zero.

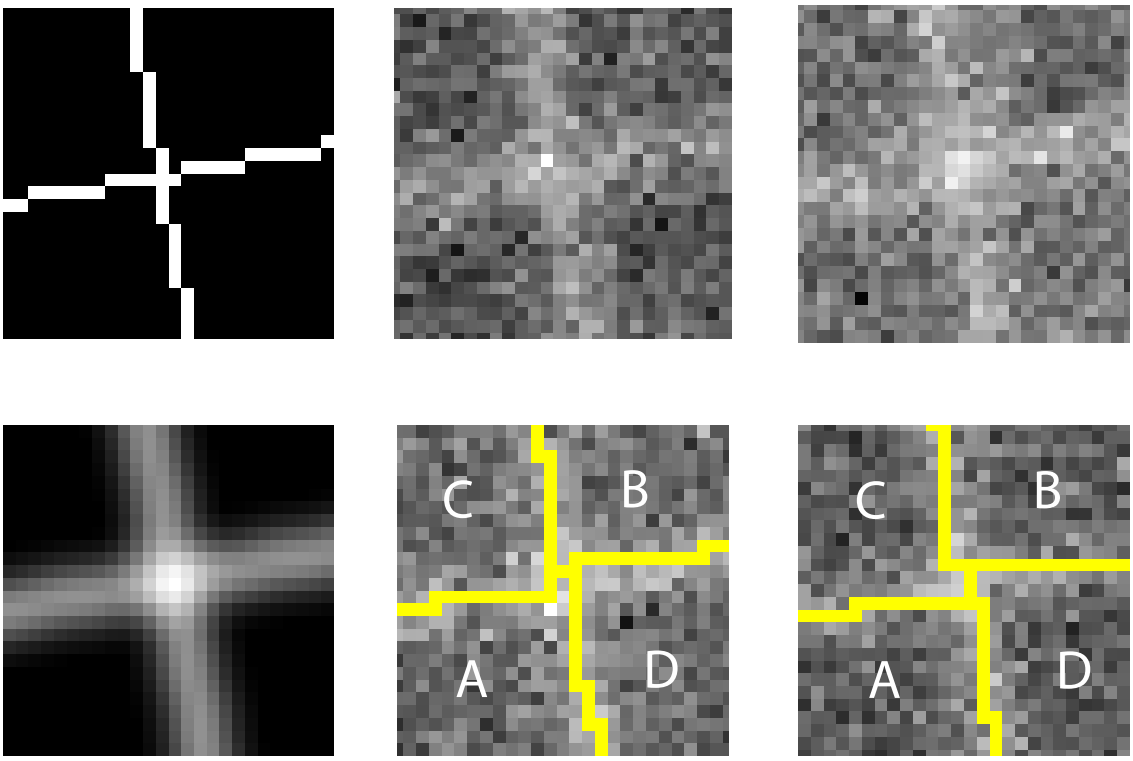

Figure 14: T2 simulations with cell ID letters and watershed segmentation lines overlaid. The left column shows the skeleton image, top, and the resulting blurred appearance after addition of the gaussian filter, bottom. The middle column shows the implementation of random noise in one frame, top, and the resulting segmentation lines overlaid in yellow with cell ID letters, bottom. Note that the artifactual interface is a horizontal T3 in this example, separating cells A and B. The rightmost column shows another iteration of random noise, that results in segmentation that separates cells C and D, a vertical T1 interface.

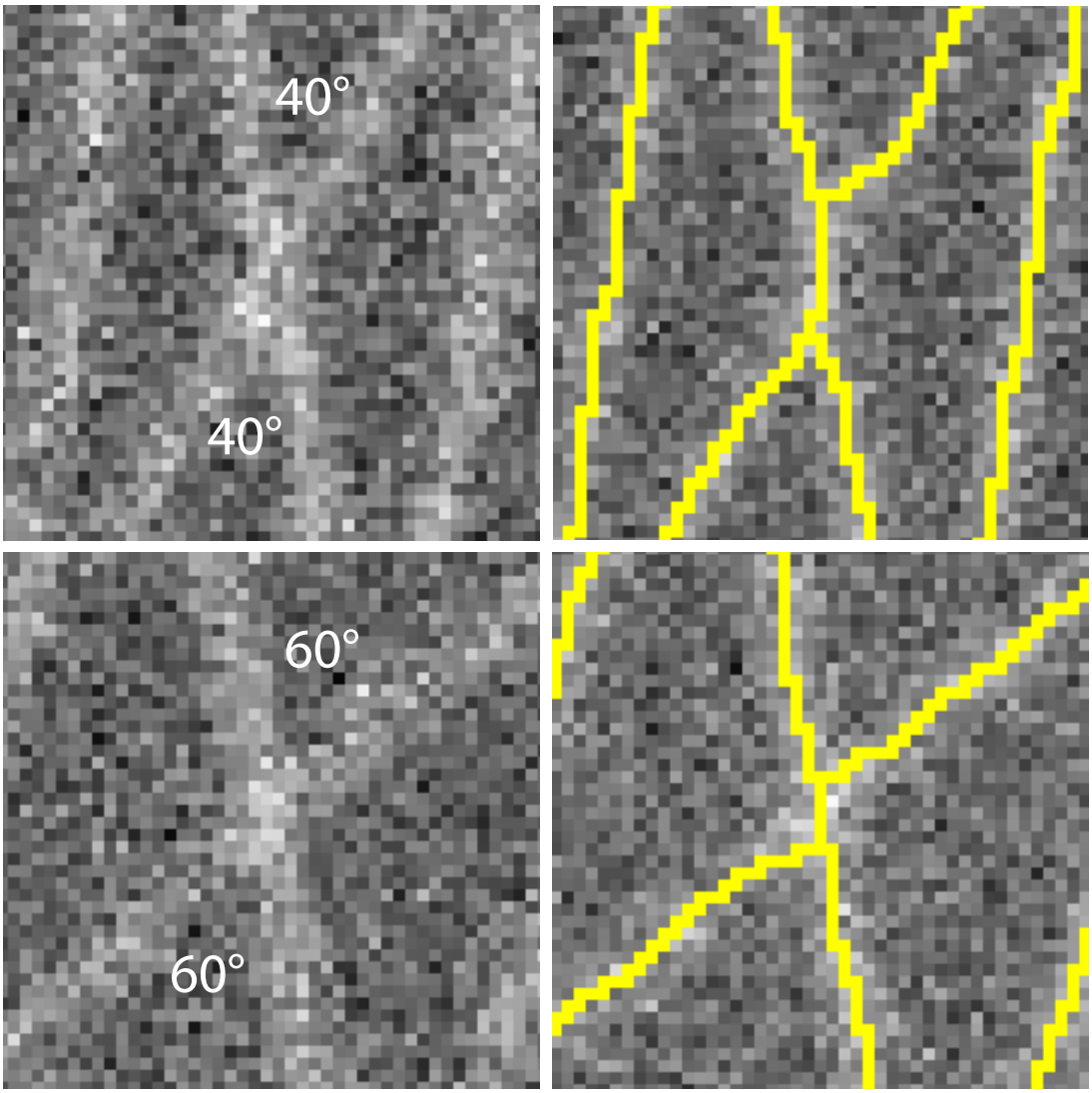

Figure 15: Internal angle measurement and its effect on artifactual interface length. The top row of images shows a T2 simulation with opposite angle measurements of 40 degrees, and its subsequent watershed segmentation results with a several pixel long artifactual interface separating two cells. The bottom row of images shows a similar T2 simulation, but with cells measuring 60 degrees. The subsequent watershed segmentation still shows an artifactual interface separating two cells, however, it is shorter in length than it was with more acute angle measurements.

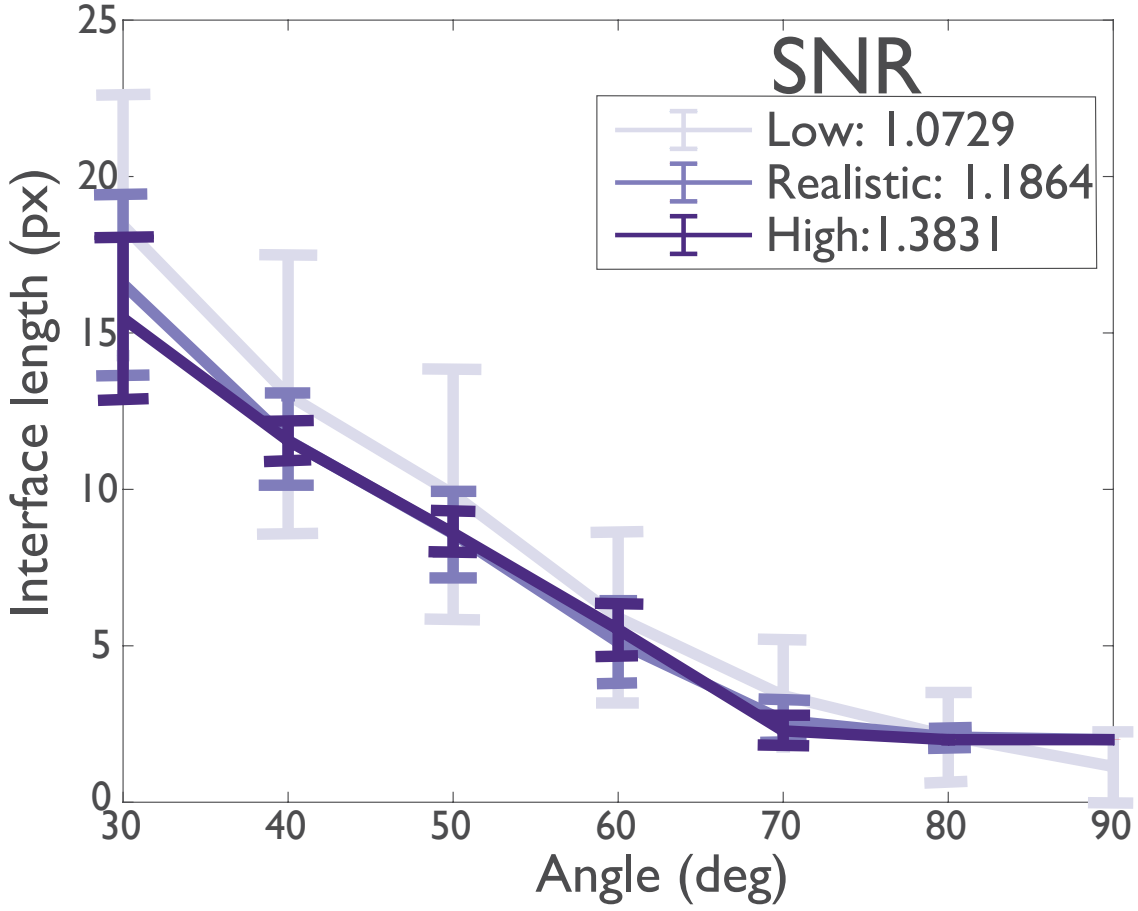

Figure 16: Mean artifactual interface length detected by the watershed segmentation of 100 simulated T2 images as a function of internal cell angle, for low, realistic and high signal to noise ratios. The value on the x axis corresponds to the angle measurements of cells C and D (Fig. 13) of the T2 simulation. Error bars: standard deviation. SNR: signal to noise ratio.

Artifactual interface lengths and direction depend on internal angle measurements

Successive T2 simulations were done at the SNR of 1.1864, comparable to real data, and with different C and D angle measurements combinations, again in ten-degree intervals. The artifactual interface lengths that occurred between angles C and D (vertical T1s) were again measured as positive values

and interface lengths between angles A and B (horizontal T3s) were measured as negative values. The heatmap in Figure 17 shows the artifactual interface length and direction for each combination of angles between 40 and 140. When angle C and D are more acute, the artifact tends to be longer, up to approximately 12 pixels, and is in the vertical T1 direction. When angles C and D are more obtuse, the artifact tends to occur in the horizontal T3 direction and can be about 12 pixels in length as well. When all four angles are 90 degrees, the error is quite small, with interfaces being on average less than a pixel long and very close to the true length of zero.

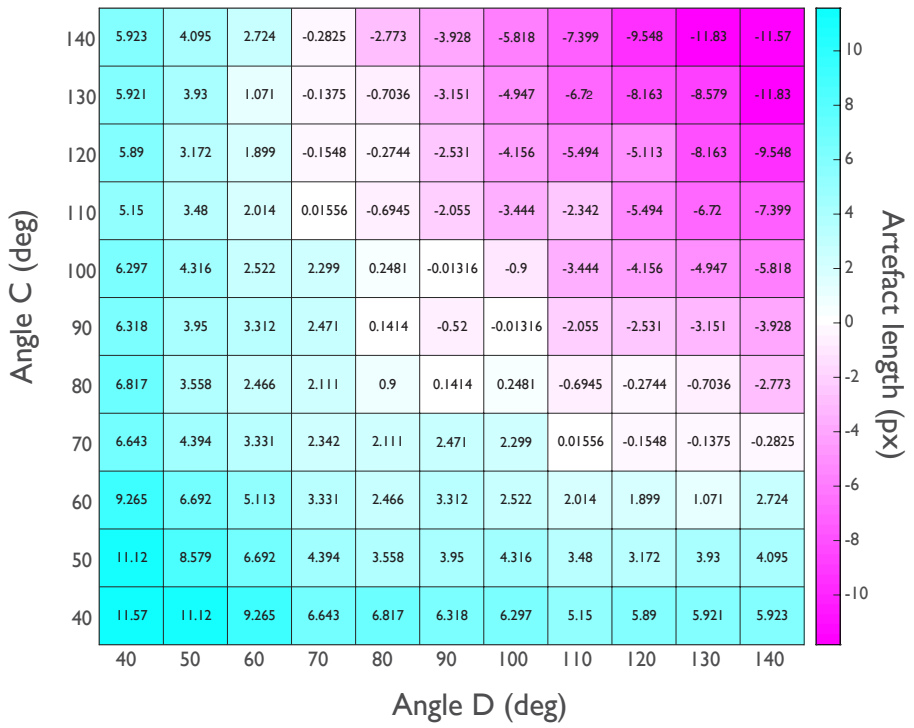

Figure 17: Heatmap of artifactual interface lengths of T2 simulations with varying combinations of cell C and D angle measurements. All simulations

were tilted 10 degrees. Each data point is an average of 100 simulations, each with a unique realization of noise. Positive values denote a vertical T1 interface between cells C and D, and negative values denote a horizontal interface between cells A and B.

The watershed segmentation systematically overestimates T1 interface lengths

When examining real angle measurements of cells at the T2 timepoint in actual data, we can analyze what angle combinations are occurring most often. In figure 18, the angle measurements of cells C and D (situated opposite each other) during the T2 timepoint were plotted. The data is concentrated around the cell pairs that both measure 80 degrees. Far fewer cell combinations are composed of one acute and one obtuse angle, and few have both highly acute or obtuse angles.

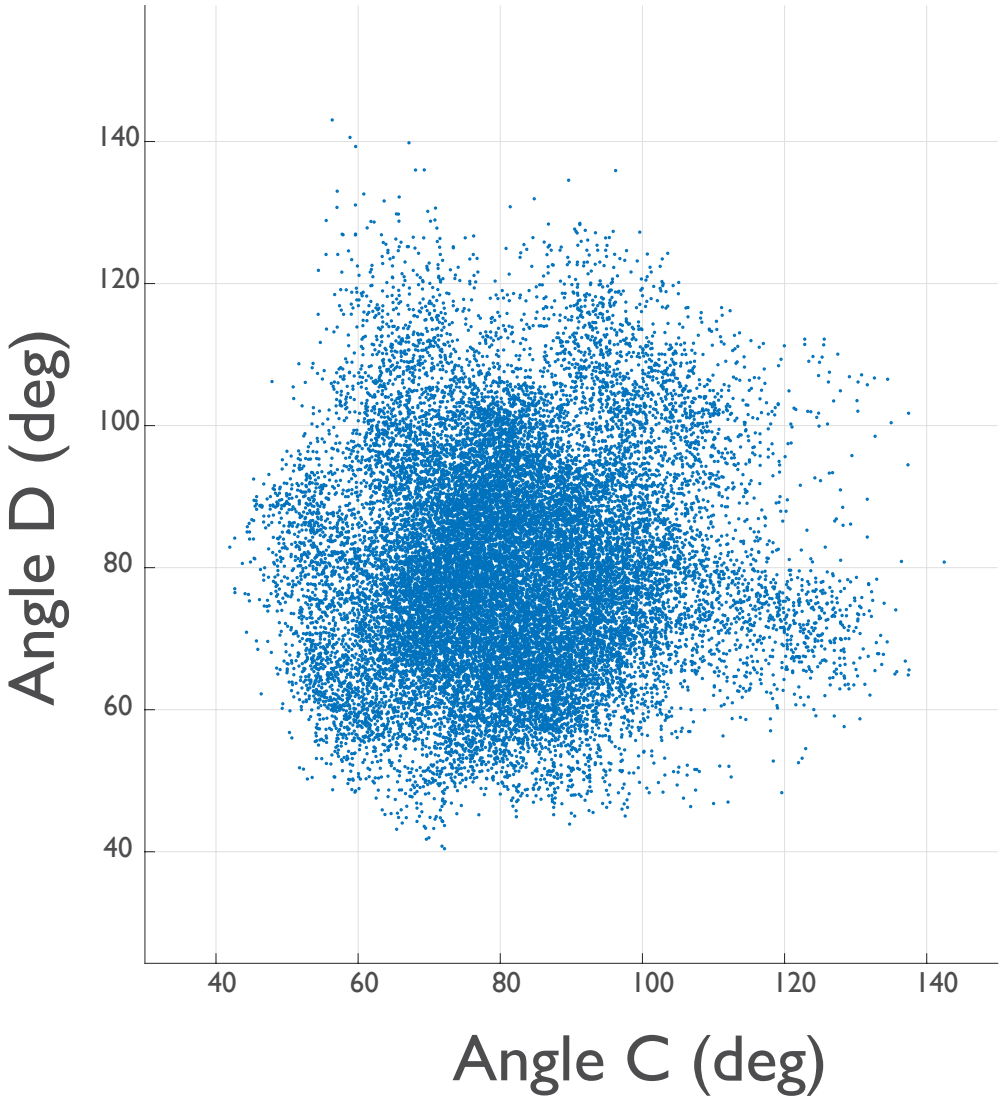

Figure 18: Scatter plot of internal angle combinations of cells C and D (See Fig. 14) during a T2 timepoint in real data. N = 26067 cells. Data courtesy of Dr. Tim Vanderleest.

I next calculated the probability of each angle combination occurring in real data based on the distribution of angle pairs at the T2 timepoint. The most likely angle combinations at the T2 timepoint are concentrated in the 70 to 90

degree range for both cells (Fig. 19). These angle combinations have corresponding artifact lengths ranging from -0.5 to 2.3 pixels. However, the most common angle combination is not the situation of complete isotropy, where all four cells meet at 90 degree angles. The distribution is biased toward angle C and D combinations that are more acute, with the most common angle combinations being 80 and 70 degrees.

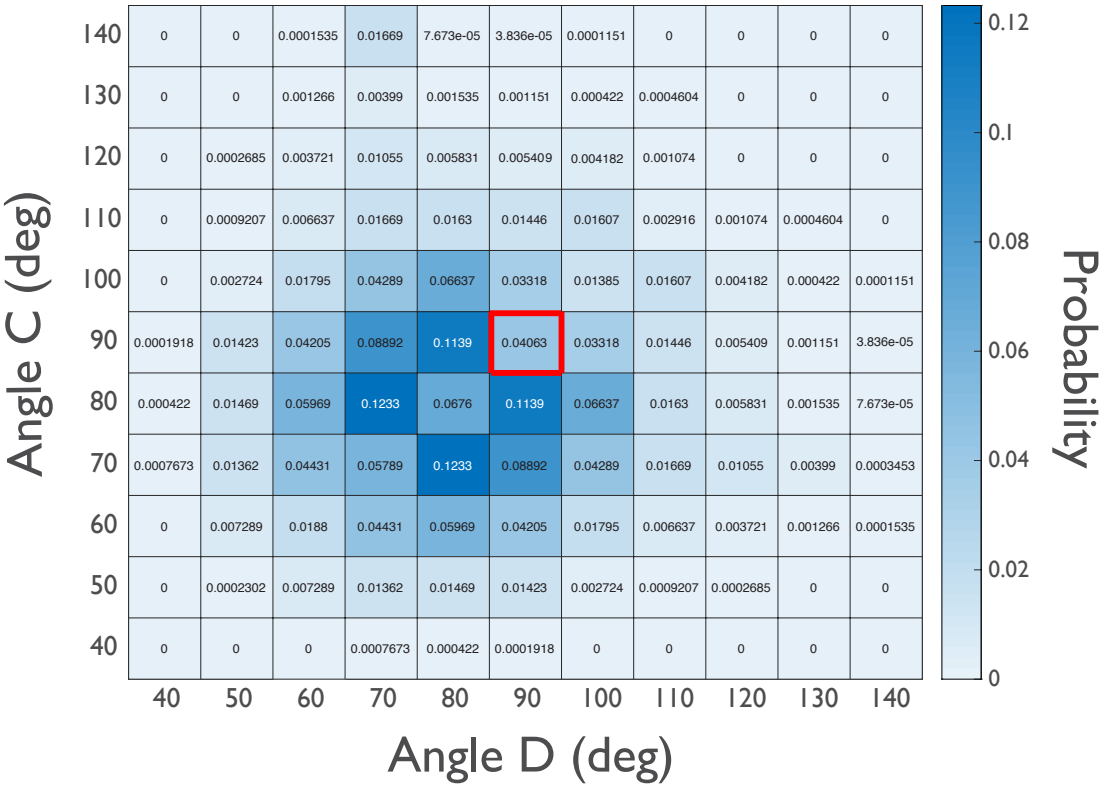

Figure 19: Probability of angle combinations at the estimated T2 timepoint in real data of 26067 cells. Red box around the 90 degree – 90 degree case shows the situation in which the T2 event has complete isotropy, where all four cells have equal angle measure.

Measurements of actual cell angles C and D at the T2 time point in real data were then convolved with their corresponding calculated artifact length. From this I was able to obtain a distribution of possible artifact lengths at the estimated T2 time point (Fig. 20). This distribution has a mean of 0.4195 and a standard deviation of 1.9490 pixels.

The areas with no data or "holes" in the distribution are present because the watershed segmentation does not measure an interface of length 1 pixel – it either measures a length of 0 or 2 pixels for an interface. This distribution is shifted toward the right side due to the bias toward acute angles at the T2 timepoint, and therefore longer positive T1 length measurements.

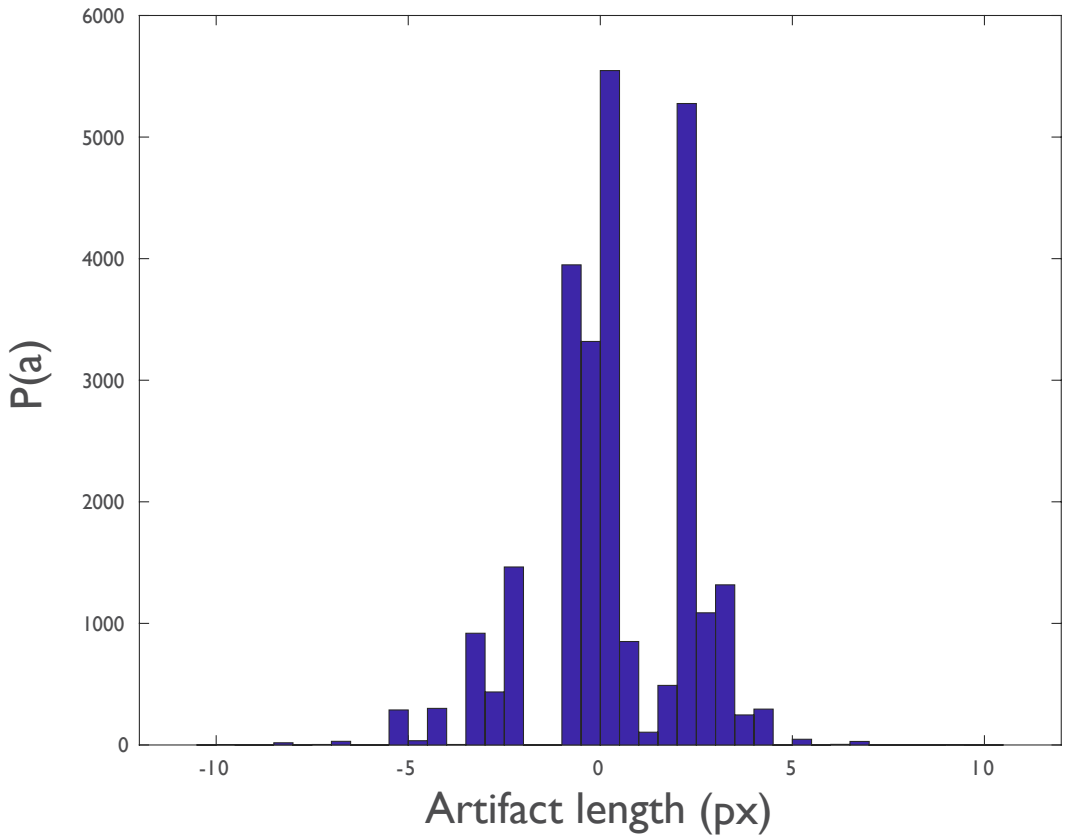

Figure 20: Convolved angle probabilities at estimated T2 time with corresponding artifact lengths of 26067 cells in real data. Half-pixel bins. Mean: 0.4195 pixels and standard deviation: 1.9490 pixels.

There are three areas of note in this distribution. One is the tail on the left side of the distribution. This 13.4% of data shows the probability that the watershed segmentation will underestimate the length of the interface, as it calculates a value less than -1.5 for the interface, meaning that it is incorrectly detecting a horizontal T3 interface. This region has a mean of -3.05 and a standard deviation of 1.16 pixels. The central lobe that holds 52.4% of the data shows the probability that the watershed segmentation essentially measures a

T2 interface, measuring between -1.5 and 1.5 pixels, with a mean of -0.11 pixels and standard deviation of 0.45 pixels. The right lobe that holds the remaining 34.2% of the data shows the probability that the algorithm is overestimating the measurement as a T1 1.5 pixels or longer, with an average overestimation of 2.6 pixels and a standard deviation of 0.67 pixels.

The mean of the entire distribution is positive, demonstrating that in real data, there is a systematic overestimation of interface lengths in the positive T1 direction. This is corroborated by the cumulative probability distribution of these artifact lengths (Fig. 21) which shows that approximately 30% of the data are above 1.5 pixels, shown in the rightmost blue line, while only 13.4% are lower than -1.5 pixels. This demonstrates a systematic bias toward measuring a T1 interface even though the interface has a true length of zero.

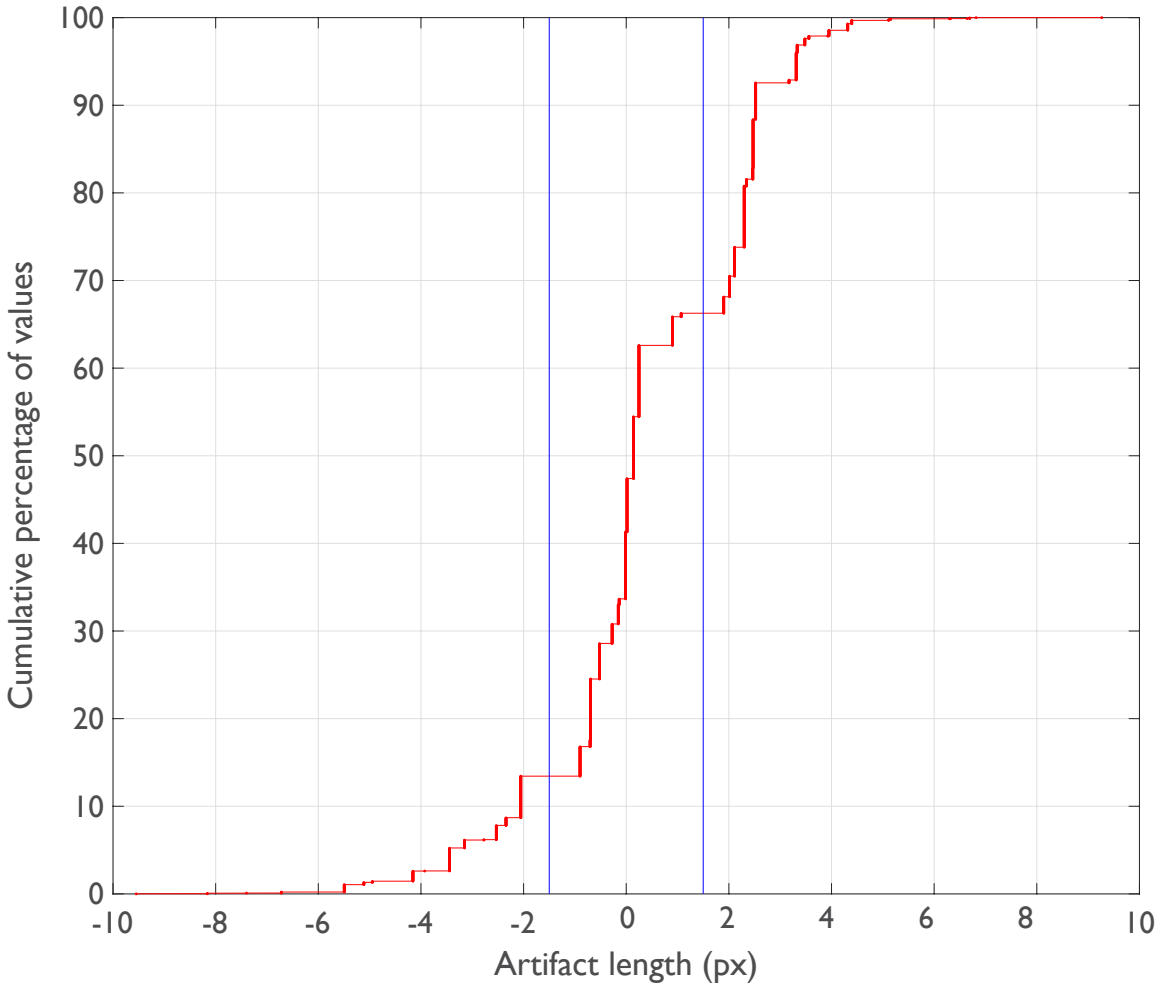

Figure 21: Cumulative probability distribution of artifact length. Blue lines represent values -1.5 pixels and 1.5 pixels.

Short T1 Simulations: a substantial overestimation of T1 lengths, and each vertex is segmented independently

The short T1 simulations and watershed segmentation yielded measurements of the T1 interface that were often different from the true length of 3 pixels, dependent on cell angle. The same distribution of angle combinations was used for these simulations, and the change in measurement

of the T1 interface from the true value of 3 pixels was calculated (Fig. 22). There is a substantial overestimation of T1 lengths by the watershed segmentation in these simulations, especially with more acute angle pairs. Additionally, it is evident that each of the two vertices composing the T1 interface are segmented independently of one another, due to the error in interface measurement at the 90-90 degree combination compared to the T2 simulation with the same angle measurement (Fig. 17). In the latter case, the isotropy of all four cells measuring 90 degrees resulted in the most accurate measurement of the interface as it was quite close to zero. In the T1 simulations, the measurement is not as accurate in the 90-90 degree case – the length is overestimated by approximately 3 pixels. The measurement is the most accurate when the angle pairs are 120 and 120 degrees, because every cell surrounding both vertices composing the short T1 interface measures 120 degrees. This isotropy of angle measurements around each vertex serves to reduce the bias of the watershed segmentation.

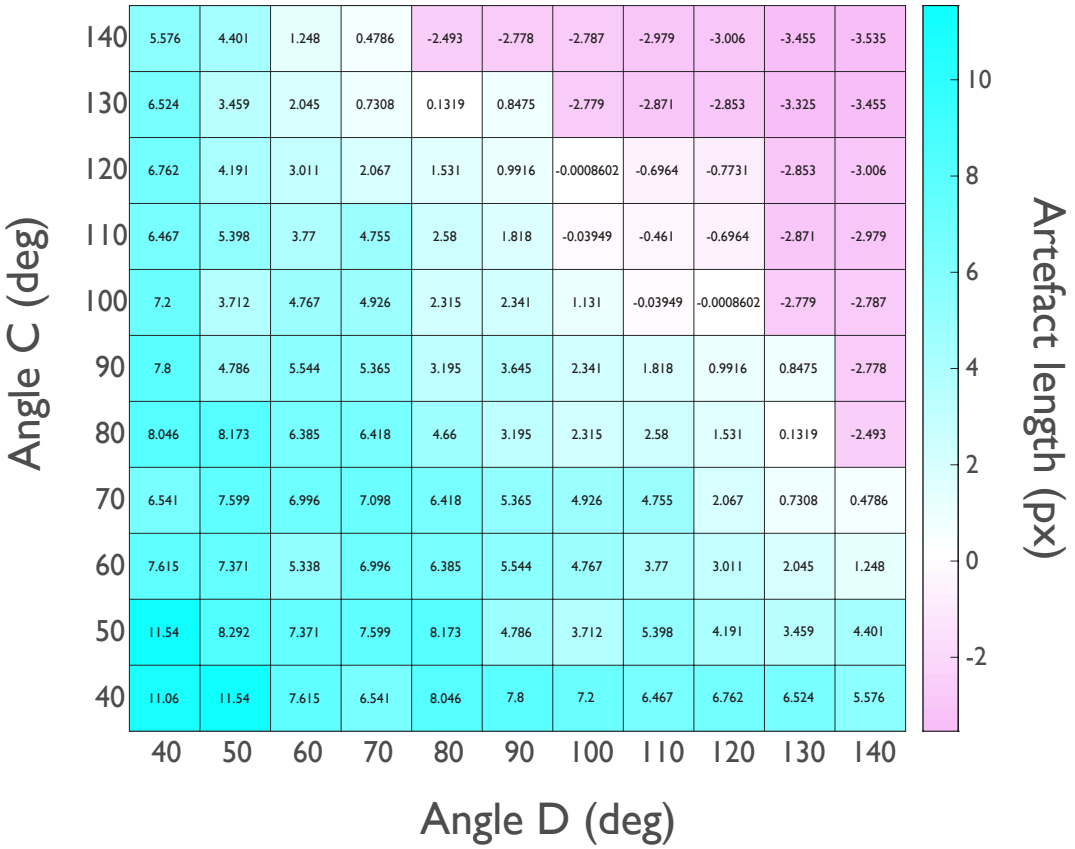

Figure 22: Heatmap showing the error in interface length measurement after watershed segmentation of three-pixel T1 simulations. Each data point is the average of 100 simulations, each with a different realization of noise. Positive values represent increase in T1 length from true length and negative values represent a decrease in measured T1 length from true length.

These data show a systematic overestimation of T1 interface length measurements. The distribution of artifact length for both the T2 and T1 simulations is biased toward the positive direction (Fig. 23) for both true lengths of 0 and 3 for interface length between cells. The possibility of incorrectly measuring a T3 interface is reduced drastically with the addition of a short T1 interface between cells C and D, however, the T1 length is still greatly

overestimated in some cases. The distribution of T2 error in length is shifted down slightly when incorporating data from other orientations of the simulated image by tilting the simulation 0, 10, 20, 30 and 40 degrees when compared to the distribution of error in measurements for the simulations only tilted 10 degrees. This is due to the slight bias that the watershed segmentation has in segmentation due to the discrete nature of the algorithm when choosing pixels to serve as segmentation lines (Fig. 7). However, both distributions show a bias toward over estimating the length of the interface by measuring it as a positive T1 interface.

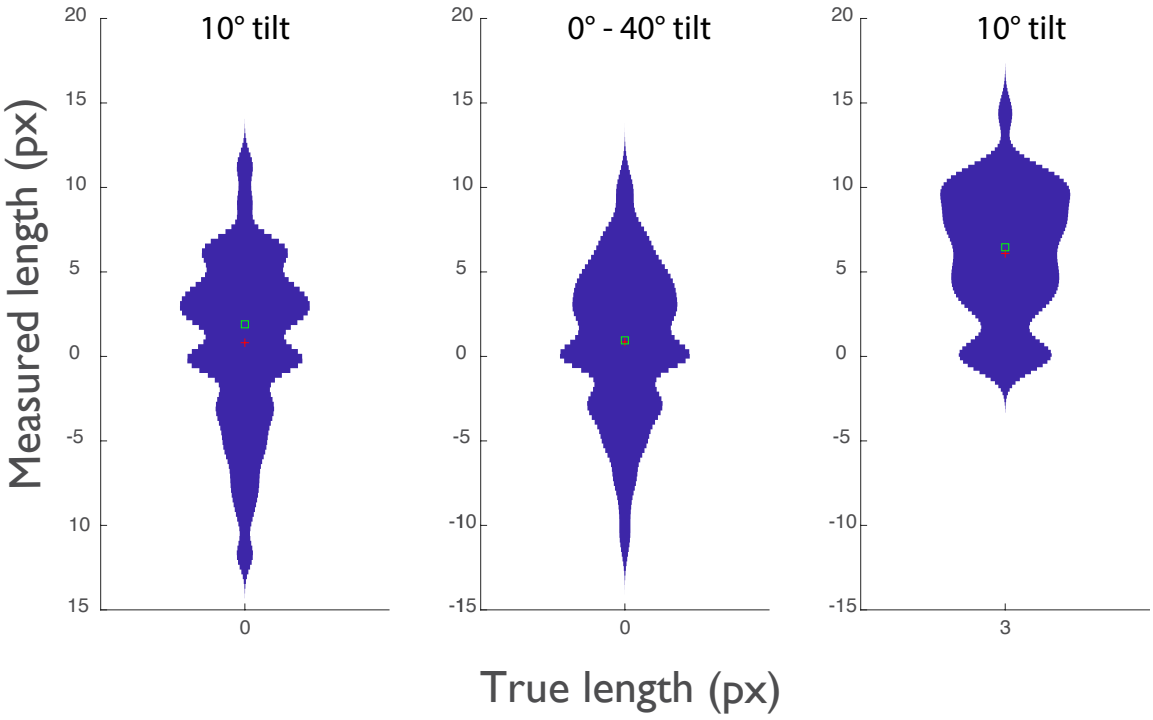

Figure 23: Violin plots showing distribution of artifact lengths. T2 simulations tilted 10 degrees, left, (n=121) tilted 0, 10, 20, 30 and 40 degrees, middle (n=605) , and short T1 simulations tilted 10 degrees, right (n=121). Each data

point of different cell angle combinations is the average of 100 simulations, each with a different realization of noise. Positive values represent vertical T1 interfaces and negative values represent horizontal T3 interfaces. Green box: mean, and red cross: median.

Order three vertex simulations: measured centroids are biased toward most acute angle

In order to understand why the watershed segmentation is biased toward overestimating T1 lengths, three-cell junctions were simulated, and the error in measurement of the true vertex position was measured and plotted relative to true vertex and interface positions (Fig. 24).

In the isotropic situation with all three cells having equal measure, when the simulation was tilted 0-40 degrees, the centroids were recognized by the watershed segmentation at the correct position 12.2% of the time. The points were plotted in the bottom cell 33.4% of the time, the left cell 1.6% of the time and the right cell 52.8% of the time (Fig 25).

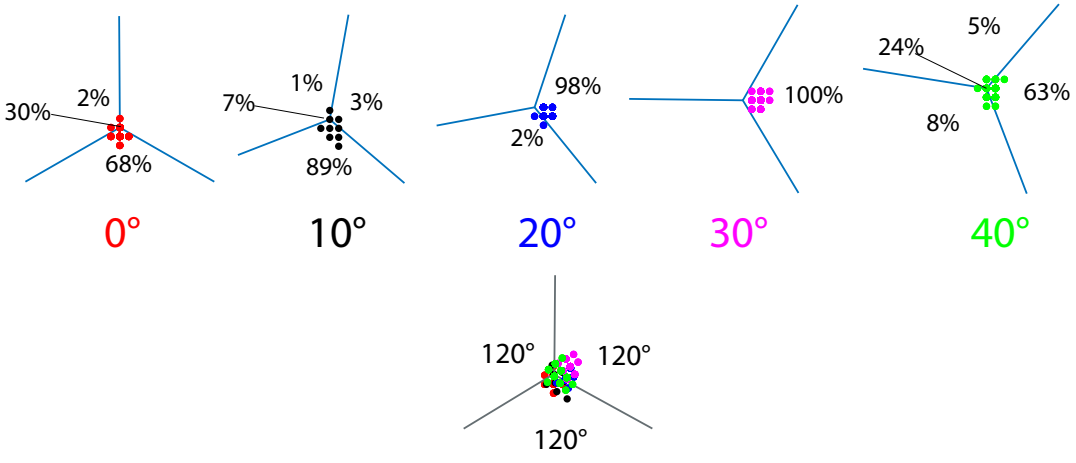

Figure 24: Three cell junction simulation and central vertex position measurements with varying image tilt. Internal cell angle measurements were 120 degrees for each cell. Percentages represent the fraction of points that were measured inside each cell or were measured correctly at the true centroid for each simulation orientation, averaged over 100 iterations for each orientation.

In the simulations with one two obtuse angles, the measured vertex positions were again plotted on top of true interface and vertex positions (Fig. 25). In this case, a vast majority of points were plotted inside the most acute cell. In the simulation that wasn't tilted, 63% of the centroids were detected along the one of the interfaces, and 1% of centroids were detected in the correct position. In the other simulations tilted 10-40 degrees, the vast majority of centroids were detected inside the bounds of the most acute cell that measured 80 degrees. Over all, in the entire simulation pool incorporating tilts 0-40 degrees, 90.5 % of centroids were detected incorrectly within the bounds of the 80 degree cell by the watershed segmentation (Fig. 26). This pattern is upheld with other angle combinations, and the majority of centroids are incorrectly detected inside the cell with the most acute angle measurement (Fig. 26).

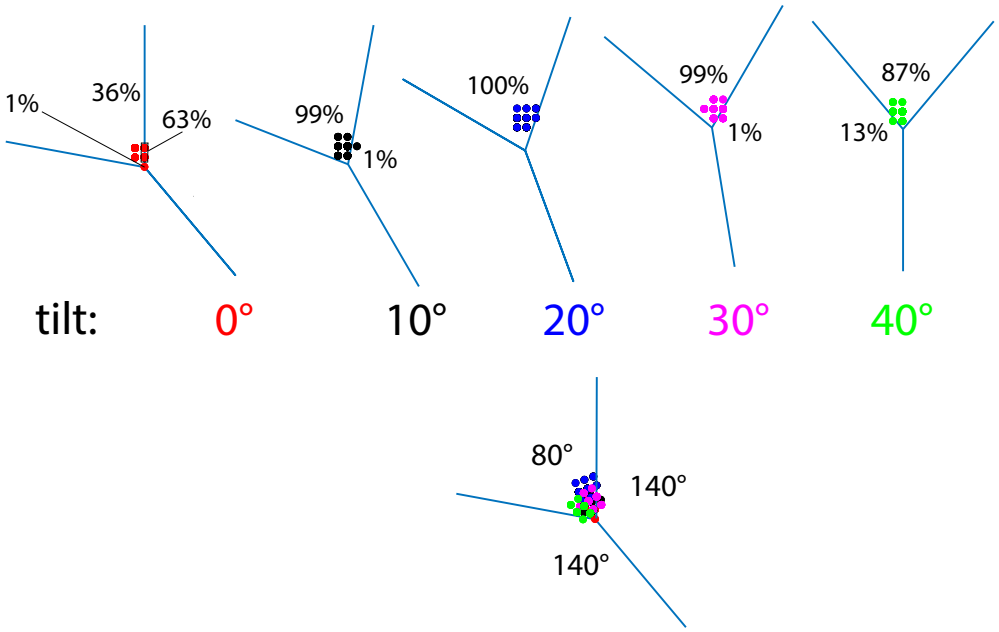

Figure 25: Three cell junction simulation and central vertex position measurements with varying image tilt. Internal cell angle measurements: 80 degrees, 140 degrees and 140 degrees. Percentages represent the fraction of points that were measured inside each cell or were measured correctly at the true centroid.

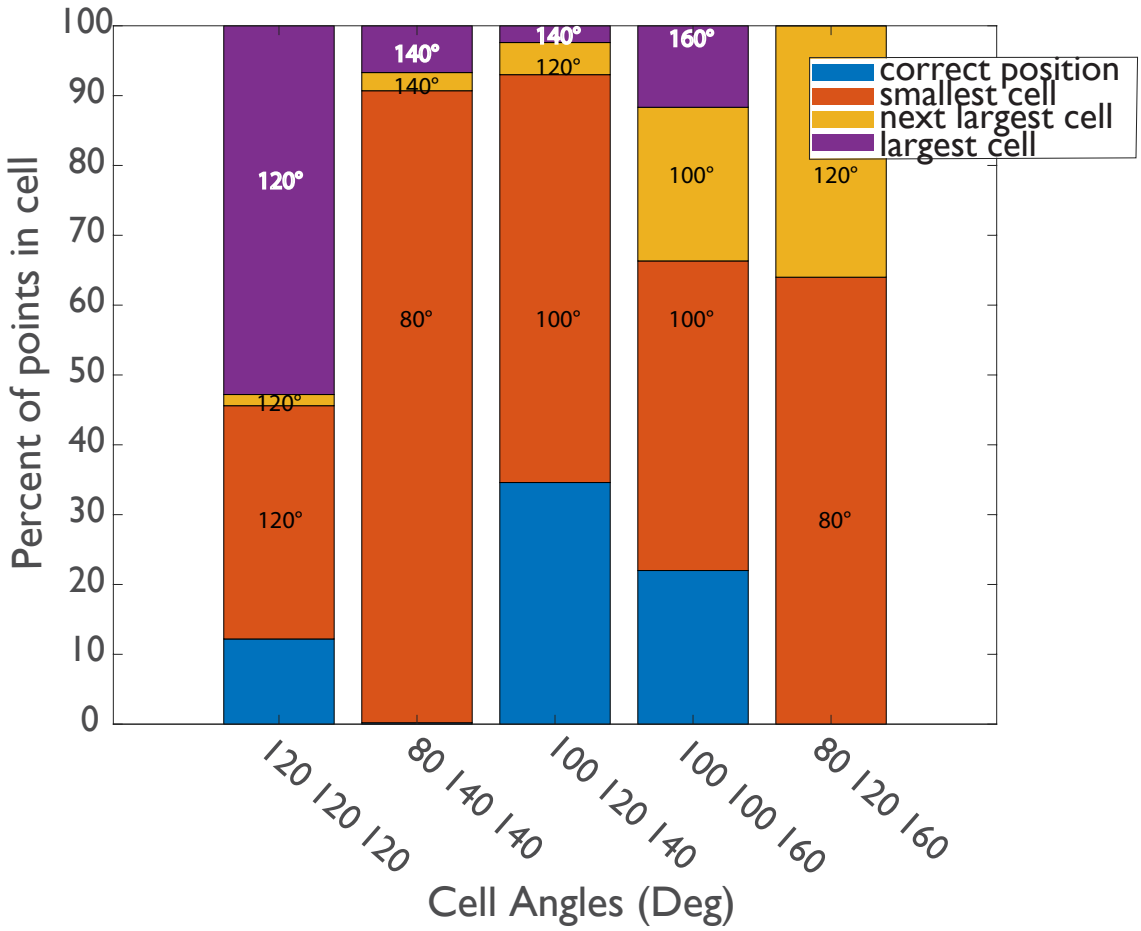

Figure 26: Bar graph of watershed segmentation detected centroids displaced into each cell's area defined by their angle measure. 500 simulations for each set of cell angle combinations, 100 for each degree of tilt: 0, 10, 20, 30 and 40 degrees (See fig. 7).

Rosette simulations: threshold of vertices and interfaces as definition of a rosette

In order to propose a working definition of a rosette structure in the context of the watershed segmentation, I created rosette simulations composed of 5 to 11 cells, as this is a possible range of rosette orders identified during

morphogenesis (Blankenship, et al., 2006). After subjecting them to the watershed segmentation, I analyzed the manner in which the true centroid broke down into a series of short interfaces.

In a five-cell rosette, 6 or 7 interfaces were present near the true centroid including the interfaces separating each of the five cells. This means up to two artifactual interfaces are present in a true five cell rosette 1 to 5 pixels in length (Fig. 27). The distance of the vertices on these interfaces from the true centroid of the rosette was also calculated and can occur up to 8 pixels away from the true vertex. Therefore, the identification of 7 interfaces up to length 5 pixels in a radius of 16 pixels can be considered indistinguishable from a five-cell rosette.

For the six-cell rosette simulations, up to 9 interfaces including the interfaces separating each of the six cells were detected. The short interfaces reached five pixels in length, and the maximum distance of the artifactual vertices were displaced by over 8 pixels. The distribution of these data was shifted right compared to the distribution for the five-cell rosette, as the average number and length as well as vertex displacement increases when the number of cells in the rosette increases. This pattern was consistent throughout the simulations of up to 11 cells (Appendix).

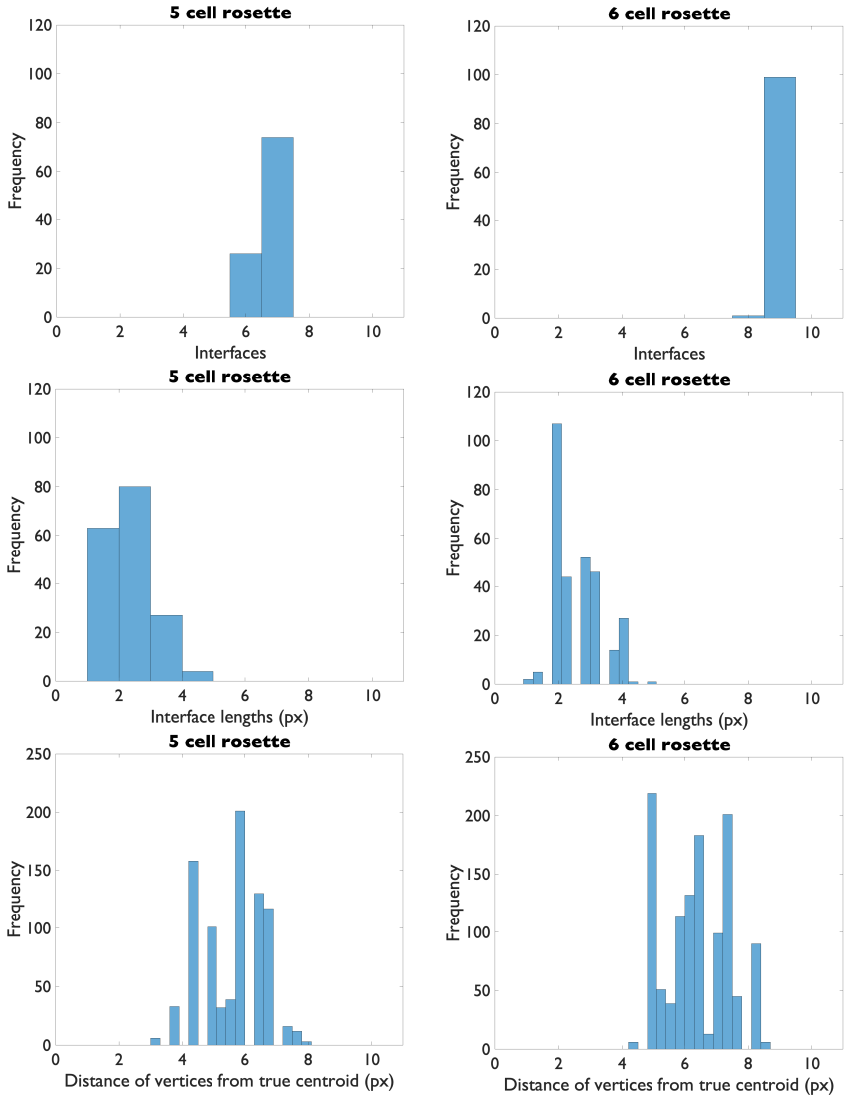

Figure 27: Distribution of detected number of interfaces, length of interfaces in pixels, and distance of detected vertices from true centroid in 5 and 6 cell rosette simulations. 100 simulations each.

Discussion

Artifactual interface length and orientation depends on internal angle of cell, not noise

In simulated T2 events, where all four cells meet at a common vertex, the watershed segmentation commonly incorrectly identifies an interface between two cells that does not actually exist (Fig. 28). The length of this artifactual interface was not influenced by the level of noise added to the simulated images and remained quite similar across varying signal to noise ratios. Strikingly, however, there was a direct relationship between the acuteness of opposite cells in the T2 event and length of the artifactual interface. As the internal angle measurements decreased in oppositely juxtaposed cells, the longer the artifactual interface separating their neighboring cells became.

Although the ground truth is that all cells are meeting at a common vertex, the angles of the cells as they meet have a strong influence on the direction and length of the artifactual interface. If opposing cells in a T2 structure are both obtuse, a horizontal artifactual interface or T3 event is detected by the watershed segmentation and gets longer as the sum of the two

obtuse angles become larger. If opposite cells in a T2 structure are both acute, a vertical T1 interface is incorrectly detected by the watershed segmentation, separating cells with an interface that in reality share a common vertex. This artifactual T1 becomes longer as the two opposite cells become more acute, or the sum of their angle measures become smaller. The most accurate detection of interface length, or when the algorithm was closest to detecting a true length of zero, was in the isotropic situation of all four cells measuring 90 degrees.

Real data of cell angles at the T2 timepoint show a bias toward slightly acute angles in cells C and D (Fig. 14) that are opposite each other. When these angle combinations were convolved with their corresponding artifact length discovered by the simulations, it became clear that the watershed segmentation systematically detects T1 interfaces when they are not present. Because both the top and bottom cells of a T2 event are more likely to be slightly acute, the watershed segmentation is more likely to incorrectly detect a short T1 interface rather than accurately detecting a T2.

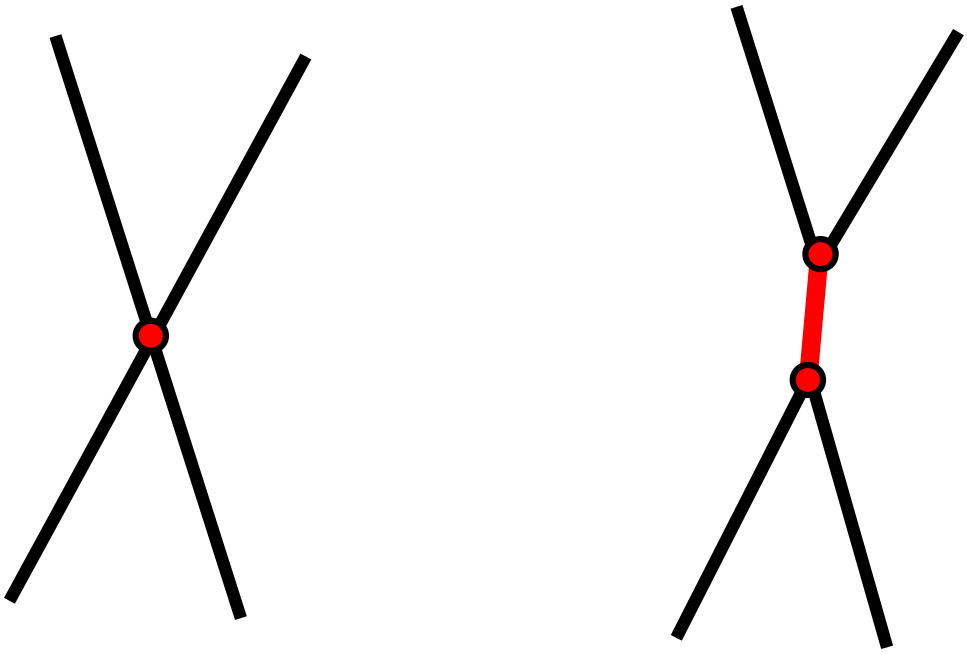

Figure 28: Illustration that represents the tendency of the watershed segmentation to incorrectly measure a T1 interface instead of a T2 event. If a T2 event is composed of two oppositely juxtaposed cells on top and bottom that are acute, the watershed segmentation will detect an artifactual vertical interface. Left image shows the true situation, a T2 event with four cells meeting at one vertex, top and bottom cells with acute internal angle measurements. The right image shows the output of the watershed segmentation incorrectly identifying a vertical interface. Black lines represent interfaces, red dots represent vertices, and red lines represent T1 interfaces.

The watershed segmentation systematically overestimates T1 interface lengths, and each vertex is segmented independently

When simulating T1 events where two cells were separated with a short vertical interface 3 pixels in length, a clear systematic overestimation of T1 lengths was identified. Although the true length of the interface was 3 pixels, the watershed segmentation overestimated this length by up to 11 pixels in the

most extreme case, again directly influenced by the acuteness of the opposing cells (Fig. 29).

The most accurate detection of interface length in the T1 simulations was not the isotropic case when all four cells measured 90 degrees, as was the case for the T2 simulations. The most accurate measurement was made when the two cells separated by the vertical T1 interface were 120 degrees. From this, independent segmentation of each vertex can be concluded. This situation allowed for equal angle measure around each vertex composing the vertical T1 interface, and its isotropy is hypothesized to be the contributing reason for such accuracy.

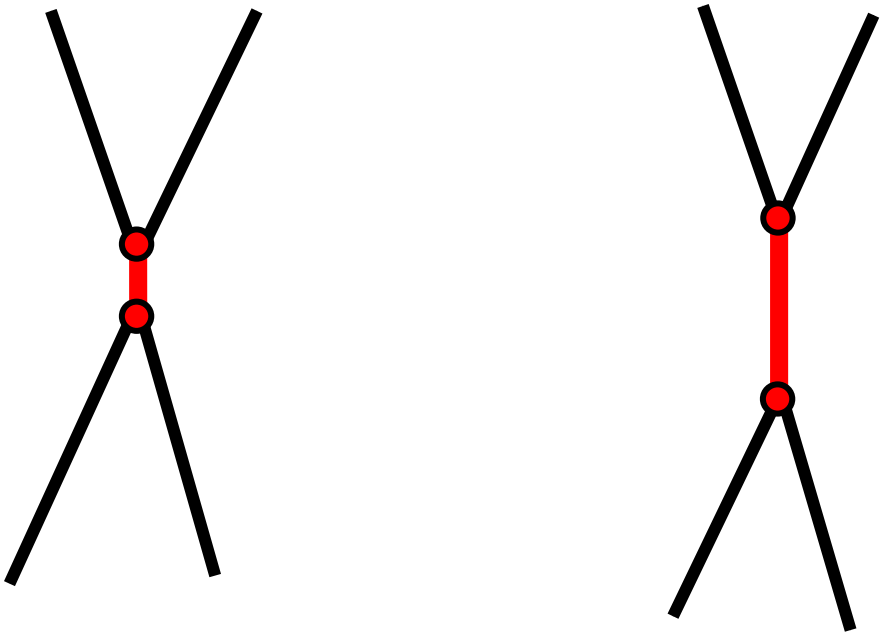

Figure 29: Illustration representing the watershed segmentation's tendency to overestimate T1 interface length. Left image shows the true situation, a short T1 interface separating the left and right cells, and the top and bottom cells

have an acute internal angle measurement. The right image shows the output of the watershed segmentation incorrectly identifying the T1 interface as much longer than it actually is. Black lines represent interfaces, red dots represent vertices, and red lines represent T1 interfaces.

Order three vertex simulations: measured centroids are biased toward most acute angle

Three cell vertex simulations provided evidence for why acuteness of opposing angles results in a likelihood of an artifactual interface to be detected. In cases where one cell was acute and the other two were obtuse, the watershed segmentation displaced the position of the central vertex into the area inside the acute angle the vast majority of the time regardless of tilt of the simulated image (Fig. 30). This corroborates previous findings that show a systematic overestimation of T1 lengths as a function of cell acuteness. Each small angle draws the position of each artifactual vertex further from each other, displaced into each acute cell and resulting in a measurement of interface length that is higher than the ground truth.

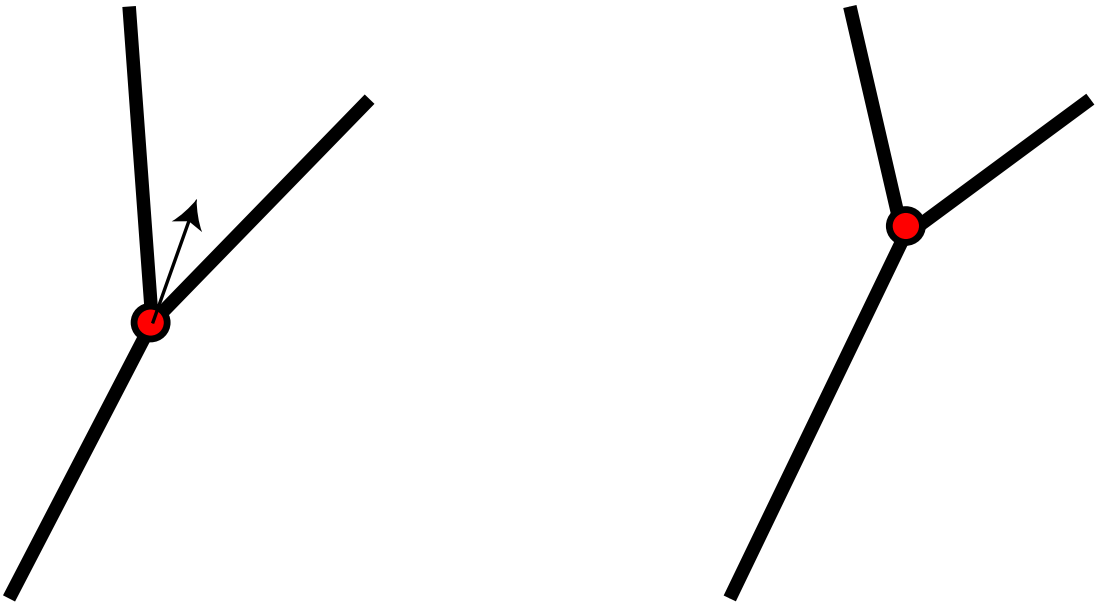

Figure 30: Illustration showing the watershed segmentation's tendency to displace vertices in three-cell junctions toward the cell with the smallest angle measure. Left image shows the true situation, three cells meeting at a single vertex, with one cell being much more acute than the other two. The right image shows the output of the watershed segmentation incorrectly identifying the vertex of the junction as displaced in the direction of the area of the most acute cell. Black lines represent interfaces, red dots represent vertices. Arrow represents direction of vertex displacement.

Rosette simulations: threshold of vertices and interfaces as definition of a rosette

Because the watershed segmentation does not allow for accurate visualization of vertices composed of five or more cells, the manner in which the watershed segmentation breaks down the central vertex of a rosette into a series of short interfaces was analyzed (Fig. 31). In a five-cell rosette, two additional interfaces were detected on top of the 5 interfaces separating each of the 5 cells, and were up to 5 pixels in length. The distribution of artifactual

vertices occurred within a radius of 8 pixels from the true central vertex of the rosette. The number and length of artifactual interfaces and radius of artifactual vertices increased as the number of cells in the rosette increased. These data may serve as a threshold of interface frequency and length that when detected in real data, is indistinguishable between a series of short interfaces and a true rosette.

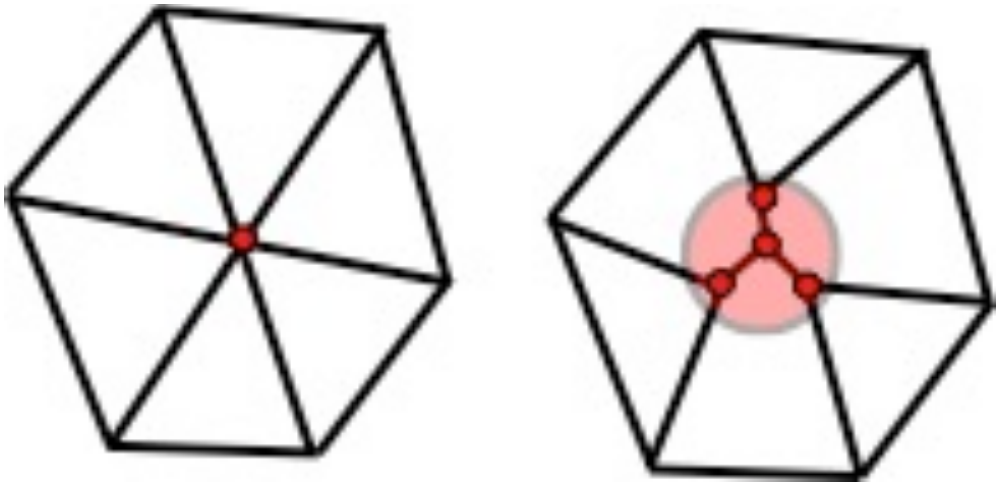

Figure 31: Illustration showing the watershed segmentation's inability to represent a vertex joining more than four cells. The left image shows the ground truth of a six-cell rosette, where six cells are joined at a single common vertex. The right image shows one way the watershed segmentation may breaks down this central vertex into a series of short interfaces with multiple vertices, all distributed in a certain radius, represented with the opaque red circle. Black lines represent interfaces, red dots represent vertices.

There are limits to these findings. By creating simulations and measuring the difference between the ground truth and the algorithm's measurements, the question of what parameters increase bias is answered, but when working with real data, the inverse problem is proposed. When measuring real data, we have

measurements done by an imperfect algorithm, but no absolute ground truth to compare the measurements to. Although my findings show what parameters may influence error on the part of the watershed segmentation's detection, it is unclear still whether conclusions made from watershed analysis are the truth or an artifact in real data. There are also alternative approaches for image processing and use of thresholding for boundary recognition that may be developed past the efficacy of the watershed segmentation (Farrell, et al., 2017; Rauzi, et al., 2008).

Additionally, more robust and detailed simulations could provide more insight into the intricacies of each parameter's influence on incorrect measurements by the algorithm that accounts for varying fluorescence levels across experimental conditions, and that incorporates image tilt that is distributed across experimentally relevant cell orientations. It would be interesting to see how different imaging techniques, fluorescent probes or noise reduction methods influence the accuracy of the watershed segmentation. It will also be interesting to see the advancement of artificial intelligence to accomplish segmentation and tracking with more efficacy, as recent applications of machine learning for this purpose have resulted in promising new findings (Wang, et al., 2017).

My results may complicate the claims made by many researchers in terms of interface and cell dynamics when using this algorithm as a means of detection, due to the identified sources of error in detection for these purposes. The errors detected may increase hesitance when making definitive claims about the length of T1 or T3 events and the additional criterion may aid in identification of rosettes to further study whether rosettes are functionally distinct or occur by random chance.

Bibliography

Bertet, C., Sulak, L. & Lecuit, T., 2004. Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation. *Nature,* 429(6992), pp. 667-671.

Beucher, S., 1992. The Watershed Transform Applied to Image Segmentation. *Scanning microscopy*.

Beucher, S. & Lantuéjoul, C., 1979. Uses of Watersheds in Contour Detection. *International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation,* Volume 132.

Blankenship, J. T., Backovic, S. T. S. J. S., Weitz, O. & Zallen, J. A. Z., 2006. Multicellular Rosette Formation Links Planar Cell Polarity to Tissue Morphogenesis. *Developmental Cell,* 11(4), pp. 459-470.

Butler, L. C. et al., 2009. Cell shape changes indicate a role for extrinsic tensile forces in Drosophila germ-band extension. *Nature Cell Biology,* 11(7), pp. 859-864.

Collinet, C., Rauzi, M., Lenne, P.-F. & Lecuit, T., 2015. Local and tissue-scale forces drive oriented junction growth during tissue extension. *Nature Cell Biology,* 17(10).

da Silva, S. M. & Vincent, J.-P., 2007. Oriented cell divisions in the extending germband of Drosophila. *Development,* 134(17), pp. 3049-3054.

De Smet, P. & Pires, R. L. V. P. M., 2000. Implementation and analysis of an optimized rainfalling watershed algorithm. *Proceedings of SPIE,* Volume 3974, pp. 759-766.

Farrell, D. L., Weitz, O., Magnasco, M. O. & Zallen, J. A., 2017. SEGGA: a toolset for rapid automated analysis of epithelial cell polarity and dynamics. *Development,* 144(9), pp. 1725-1734.

Fernandez-Gonzalez, R. & Zallen, J. A., 2011. Oscillatory behaviors and hierarchical assembly of contractile structures in intercalating cells. *Physical Biology,* 8(4).

Irvine, K. D. & Wieschaus, E., 1994. Cell intercalation during drosophila germband extension and its regulation by pair-rule segmentation genes. *Development,* 120(4), pp. 827-841.

Jessica, C. Y. & Fernandez-Gonzalez, R., 2016. Local mechanical forces promote polarized junctional assembly and axis elongation in drosophila. *Elife,* 5(e10757).

Jewett, C. E. et al., 2017. Planar polarized Rab35 functions as an oscillatory ratchet during cell intercalation in the Drosophila epithelium. *Nature Communications,* 8(476).

Leung, C. Y. B. & Fernandez-Gonzalez, R., 2015. Quantitative Image Analysis of Cell Behavior and Molecular Dynamics During Tissue Morphogenesis. *Methods in Molecular Biology,* Volume 1189, pp. 99-113.

Meijster, A. & Roerdink, J., 1995. A proposal for the implementation of a parallel watershed algorithm. In: V. Hlavac & R. Sara, eds. *Computer Analysis of Images and Patterns: 6th International Conference, CAIP '95 Prague, Czech Republic.* Berlin/Heidelberg, Germany: Springer, pp. 790-795.

Meyer, F., 1994. Topographic distance and watershed lines. *Signal Process,* Volume 38, pp. 113-125.

Rauzi, M., Verant, P., Lecuit, T. & Lenne, P.-F., 2008. Nature and anisotropy of cortical forces orienting Drosophila tissue morphogenesis. *Nature Cell Biology,* 10(12), pp. 1401-1410.

Serra, J., 1982. *Image analysis and mathematical morphology*. London: Academic Press.

Vichas, A. & Zallen, J. A., 2011. Translating cell polarity into tissue elongation. *Seminars in Cell & Developmental Biology,* Volume 22, pp. 858-864.

Vincent, L. & Soille, P., 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactoins on Pattern Analysis and Machine Intelligence,* 13(6), pp. 583-598.

Wang, M. F. Z. et al., 2017. Automated cell trackign identifies mechanically oriented cell divisions during Drosophila axis elongation. *Development,* Volume 144, pp. 1350-1361.

Weaire, D. & Hutzler, S., 2000. *The Physics of Foams.* Oxford, England: Oxford University Press.

Appendix

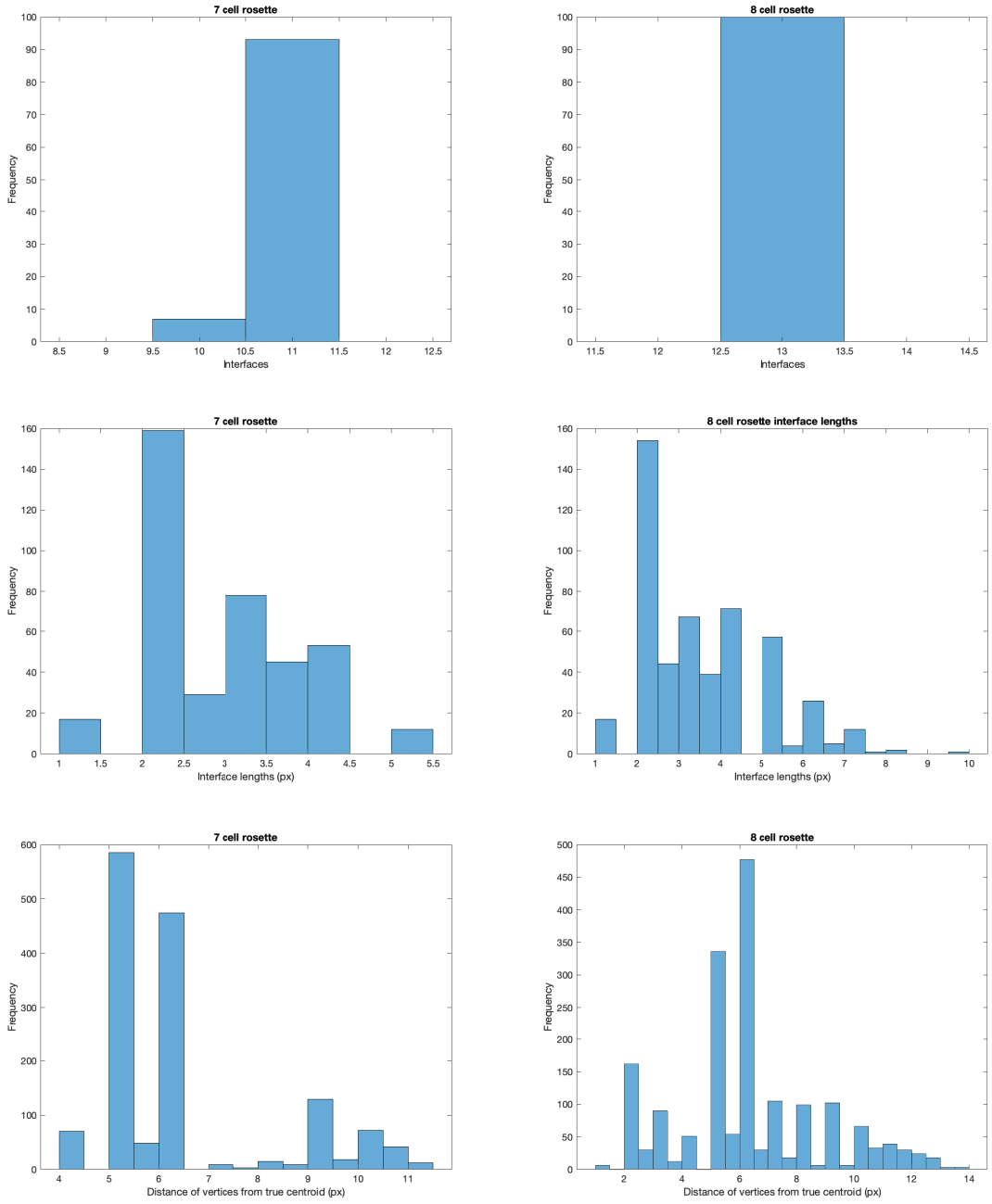

Figure 32: Distribution of detected number of interfaces, length of interfaces in pixels, and distance of detected vertices from true centroid in 7 and 8 cell rosette simulations. 100 simulations were done for each rosette.

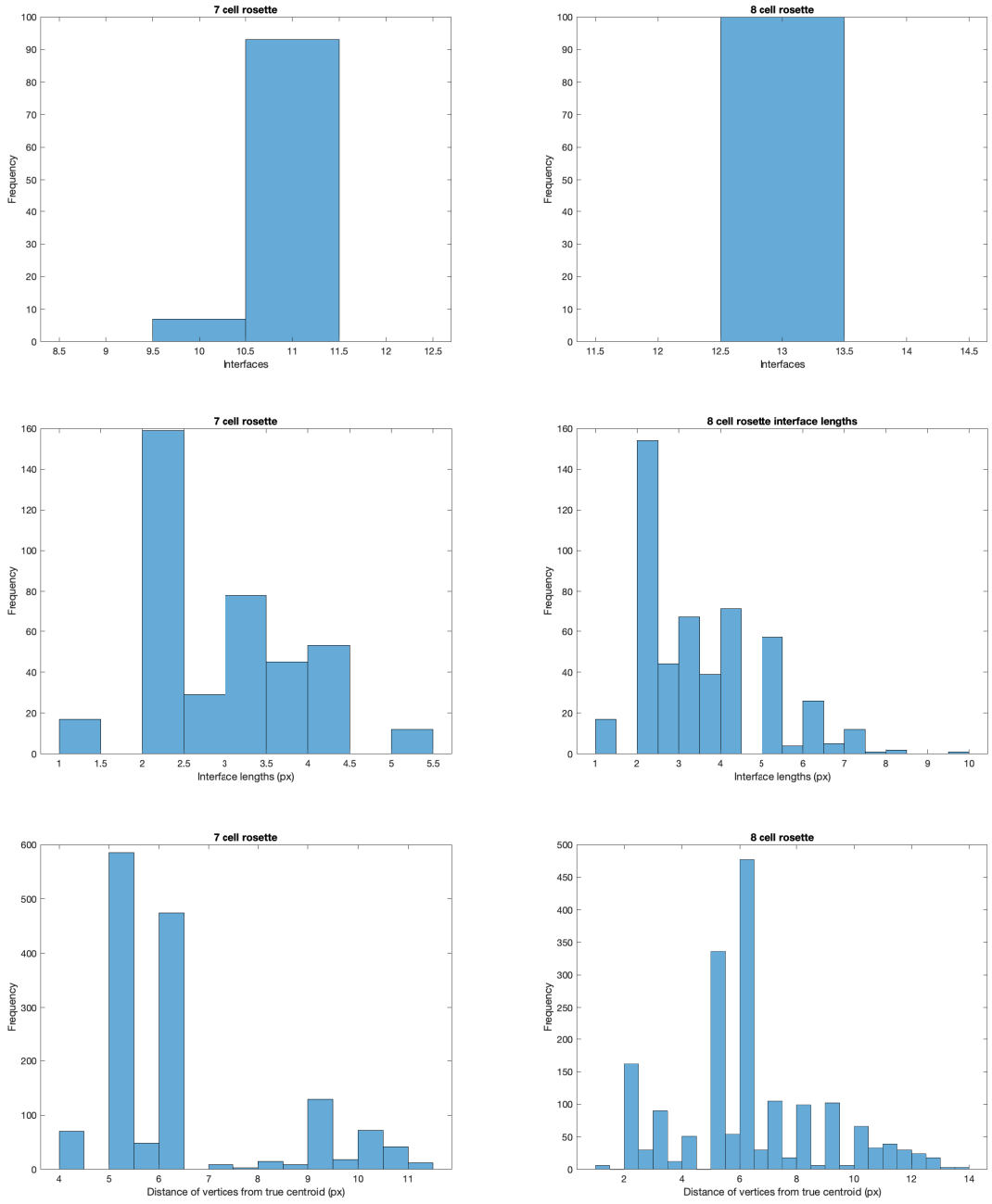

Figure 33: Distribution of detected number of interfaces, length of interfaces in pixels, and distance of detected vertices from true centroid in 9 and 10 cell rosette simulations. 100 simulations were done for each rosette.

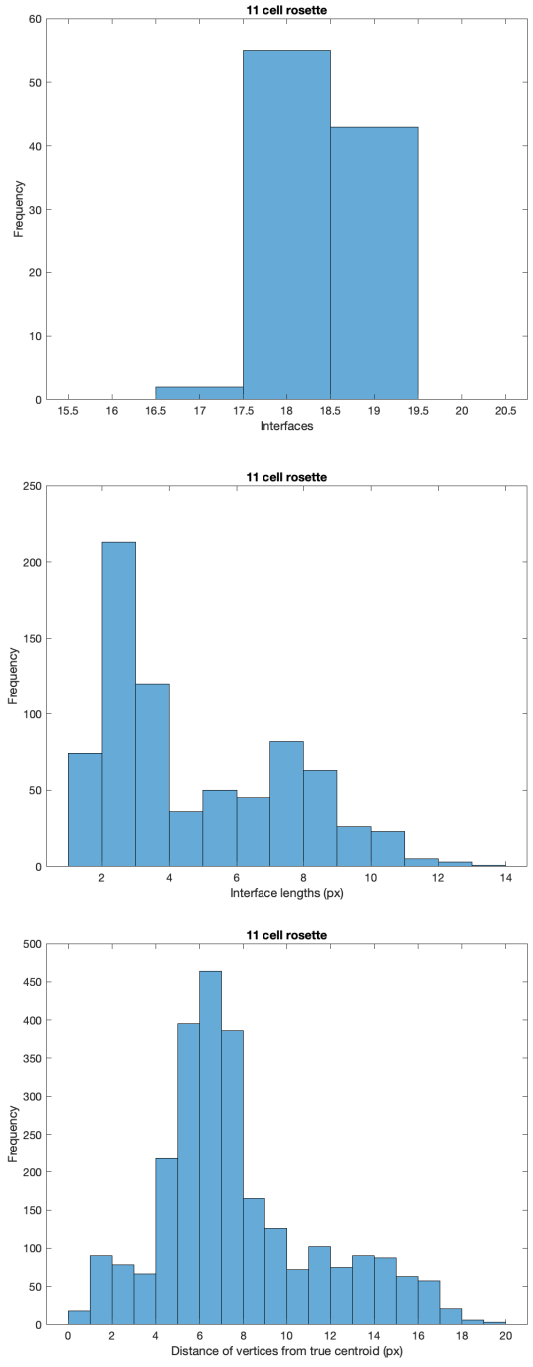

Figure 34 : Distribution of detected number of interfaces, length of interfaces in pixels, and distance of detected vertices from true centroid in 11 cell rosette simulations. 100 simulations were done for the 11 cell rosette.