Lord of the Wings: A Computational Toolbox for the Automated Processing and Geometric Analysis of Odonate Wings

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:38811459

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Abstract

Insect wings have interested a surprisingly broad range of disciplines. Engineers explore the mechanical properties of insect wings for applications in microrobotics and aeronautics, and biologists tend to focus on the development, evolutionary history, and scaling relationships amongst insect wings. In order to collect detailed quantitative measurements of wings, these researchers have primarily used manual methods. In this paper, we develop tools to automate the extraction and analysis of the shape and arrangements of the structural components of insect wings. Specifically, our tools compute statistics for the shape and size of the transparent “cells” that are formed by the rigid cuticular struts called “veins.” We calculate a range of characteristics, including edge number, size, and curvature for the wing cells of more than 10 wings from various species of dragonflies and damselflies. Our results provide insight into the mechanics of the wing and allow us to hypothesize rules that govern wing formation. Furthermore, our research also suggests that increased wing length tends to scale with increased number of cells and decreased average wing cell size, such that more of the wing’s surface area is dominated by smaller cells. Developmental principles have been studied in detail for *Drosophila melanogaster* wings, and it is unknown whether these principles apply to other insects with more complex wing structures. Our observations enable us to hypothesize that there are geometrically distinct regions in the wing that are laid down in sequence. We have used these findings to make mathematically precise, testable predictions about how the patterns in dragonfly and damselfly wings develop.
Acknowledgements

I was initially very intimidated to write a thesis, but I am so happy that I decided to pursue it. While this paper is a rewarding product to have, the most invaluable prize I earned from this thesis was being able to collaborate with many wonderful people who generously donated their time and wisdom; I truly could not have done it without their help.

I would first like to thank Chris Rycroft, who I have had the pleasure of knowing since sophomore year, when he first arrived at Harvard. Over the past three years as my advisor, he has taught me so much about the field of applied math, but more importantly, he has served as a funny, spirited, and caring person that I can easily talk to about anything (from Voronoi to polar bears, and beyond!). He is also the one who encouraged me to write a thesis in the first place, and I thank him so much for believing in me.

Secondly, I have to thank Jordan Hoffmann for his endless support and supply of coffee, intriguing ideas, strategic chess moves, and weirdly specific jokes that always seem to go over my head. I never thought I would find someone who enjoyed staying at Starbucks until 1 AM as I do. In all seriousness, I am extremely lucky to have gotten to know him
through this thesis and he made the experience more fun than I thought it could be. Hopefully I can repay him for his help someday by buying him a whale shark with the profits from my dragonfly wing coloring book.

This thesis also could not have been possible without the tremendous support of Seth Donoughe, who provided much of the biological inspiration behind this thesis and helped motivate many of the questions I ended up exploring. I always admired his plethora of different ideas, his boundless passion and enthusiasm, and his ability to wake up before 9 AM.

I also thank Cassandra Extavour and the Extavour Lab for their support and encouragement throughout this process, as well as for letting me use equipment and resources. It was amazing being able to actually image the wing data myself from start to finish. Additionally, thank you to the Harvard Center for Biological Imaging, who provided the microscope for imaging the wing data.

Of course, I am also grateful for my friends who kept me motivated throughout these past months. I am thankful especially for Keyon Vafa, who was always there for support when I needed it most. I also have to thank Ting Zhou, my BFF since freshman year, for keeping the most optimistic attitude and playing Facebook chess with me. And to my fellow thesis-er and roommate for the past three years, Stephanie Manivanh – I am so fortunate to have you as a friend, and I will never stop admiring your energy and spirit. Thank you for the encouragement not only throughout this process, but always.

As for my family – Cary Li, Charlie Li, Mom, Dad – they have always been the people in my life that I hold dearest, and as with any big project I have tackled over the years, I
could not have done it without them. I hope this thesis makes them proud and that it can (somewhat) live up to my Dad's PhD thesis from back in the day!

I have the utmost respect for everyone that I have worked with throughout this process, and I cannot properly express how grateful I feel for the opportunity to have embarked on this learning experience with them. I hope you enjoy reading the thesis as much as I enjoyed creating it.
3.2 Identifying Regions by Geometric Character .......................... 59
3.3 Hypothesized Pattern Formation ............................................. 64

4 Conclusion and Further Directions .................................................. 67

5 Appendix ................................................................................. 73
Chapter 1

Introduction and Background

Dragonflies and damselflies are insects of the order *Odonata*. As aerial predators, their ability to fly is crucial to survival; luckily for odonates, they have extraordinarily agile flight. This flight ability is thought to be due in part to the complex corrugated nature and delicate veining patterns seen in the wings. These veins contribute to the flexible quality of the wing and the local aerodynamic qualities found in different areas of the wing that cannot be replicated by something like an aerofoil [1]. Thus, understanding the mechanical characteristics of different odonate wings and the rules that govern the structural pattern formation could be essential to various applications, such as replicating the wings for micro-air-vehicles (MAVs) [2]. Furthermore, understanding both the formation and mechanical function of the wings could allow engineers to more easily adjust parameters of the wings to fit their needs. While engineers have already been able to design 3D-printed wings that are capable of stable flight [2], the wings on these flight machines are still faulty - they are often subjected to high mechanical wear from the constant clapping motion that allows them to fly and consequently lack the durability of real insect wings [3]. Some developments in the understanding of
wing mechanics have improved the engineering of these MAVs, such as the discovery that morphological spacing of wing veins usually prevents fracturing and deformations throughout the insect's lifetime \[4\]. However, there is still much room for exploration and understanding that caters to both biological and mechanical interest.

Figure 1.1: Compilation of different odonate (dragonfly and damselfly) wings. We see that there is geometric variation between the wings - no two wings look identical, but there are repeated patterns in the wings. This geometric complexity makes odonate wings difficult to study, and provides our motivation for being able to characterize geometric patterns in the wings. Seth Donoughe and Kathy Li obtained wing data.

In order to understand both the mechanical function of odonate wings and the wing formation process, it is necessary to first investigate the geometry of the wing. While regulation of wing formation in insects such as *Drosophila melanogaster* (fruit fly) is understood quite well \[6\], the study of pattern formation in odonate wings has been a mystifying question. This is largely due to the fact that odonates can be difficult to
cultivate and take years to fully form, unlike *Drosophila*, which reproduce quickly. Furthermore, whereas *Drosophila* have very stereotyped wings [7] with only a few critical veins, odonates have a much more varied vein structure that renders them more difficult to study. These differences are evident visually, as seen in Figure 1.1 and Figure 1.2. Figure 1.1 in particular showcases the geometric variation that is present between different odonate wings. This observed geometric variation led us to develop a computational set of tools that can be used not only for odonate wings, but in a variety of contexts to study and analyze complex images (i.e., ones of multicellular structures, like odonate wings) in a streamlined manner. These tools allow us to examine the geometric aspects of wing cells (where we define *wing cells* as regions surrounded by veins) such as the number of edges of a cell and the number of neighboring cells it touches, which will motivate our hypotheses for rules that govern odonate wing pattern formation.

The study of shape geometry and shape tiling has been conducted in various fields where specific arrangements of shapes have been observed. In an example drawn from biology, the observation that predominantly hexagonal patterns exist in epithelial layers inspired research into how these patterns form in *Drosophila* epithelia. This research suggested that these polygonal arrangements form not to maximize cell packing, but rather to satisfy a mathematical equilibrium that is reached between epithelial cells during cell proliferation [8]. This existing research is more focused around the process of cell division during epithelial tissue formation, which is examined by looking at the geometric organization of cells at various time steps. In this paper, we instead take a reverse approach: we begin with the fully formed wing to then deduce how the existing patterns can uncover geometrically distinct regions that characterize the wing formation.
Chapter 1. Introduction and Background

process. In the non-biological arena, observed geometric patterns are also widespread. One elucidating case is the study of the patterns of soap bubbles, which visually resemble the patterns we see in odonate wings. Research into the patterns formed by clusters of soap bubbles revealed that these structures were a consequence of energy minimization, and that no other structures would be possible [9]. For example, it was shown that a cluster of three bubbles could only intersect in one way - the three bubbles would meet along a curve, forming 120-degree angles with one another [[9], [10]]. Therefore, it was an examination of how bubbles interacted with their neighboring bubbles that motivated this resulting discovery. While bubbles and epithelial cells are two very different subject matters, in each case we see how an understanding of the object’s underlying geometric structure uncovers principles of how these arrangements form.

We focus on odonate wings as a case study for utilizing an understanding of geometric characteristics to hypothesize principles of formation. Unlike wings in birds or bats, which can be thought of as modified forelimbs, insect wings do not contain any muscle tissue beyond the axilla, which is the base of the wing that emerges from the insect body [11] and instead are composed of chitin. Thus, while birds can control their wing shape through muscle movement, insect wing movement is controlled by external forces; specifically, the forces are distally transmitted to the rigid parts of the wing, or the veins and the areas of the wing membrane (the clear material that forms the wing cells) that are thickened [11]. These rigid components are subjected to external aerodynamic forces that deform the three-dimensional shape of the wing. In addition to wing cells, veins are another important aspect of the insect wing that are studied, often for their ability to support the wing and because different venation patterns often distinguish insect
Chapter 1. Introduction and Background

orders [12]. The typical definition of a vein is a cuticular tube that contains fluids like hemolymph [11]. We distinguish between *longitudinal veins*, which in the insect wing are defined as veins in the wing that radiate from the base of the wing, and *cross-veins*, which are veins that are positioned perpendicularly to the longitudinal ones [5].

The genetic and cellular basis of wing development is known only for *Drosophila*. Genetically, venation in *Drosophila* has been widely studied - in contrast to odonates, venation is much more stereotyped, lending itself well to being altered by manipulating gene expression. The formation of *Drosophila* wings is largely regulated by the *vestigial* gene, whose expression is controlled by various signals like the *Su(H)* protein, which regulates the dorsal/ventral boundary enhancer of the *vestigial* gene to allow for symmetric expression of the gene on both sides of the dorsal/ventral boundary [6]. Other regulatory pathways such as the Hedgehog pathway and Notch pathway also contribute to the promotion of vein differentiation and formation in odonate wings through different signaling interactions [[13], [7]]. Thus, we see that regulatory pathways and interaction between signals play a crucial role in pattern formation in *Drosophila*. The interactions between these signals is well-modeled by a reaction-diffusion gradient model (i.e., interaction between different levels of regulatory proteins) [14]. Biological pattern formation as regulated by this type of interaction between inhibitory and stimulatory forces is seen in other tissues, such as in the pattern formation of chicken basilar papilla hair (auditory sensory organ) [15] or in pattern formation on shells [16]. Thus, we see that pattern formation is not only omnipresent in various specimens, but can also stem from a simple set of rules that result in a multitude of different patterns.

Other studies have attempted to see whether the reaction-diffusion model seen in
Drosophila wing formation is also present in other insects, such as sawflies, to determine if this type of formation model can be widely applied [17]. Therefore, while there has been some progress moving outside the realm of just Drosophila wing formation, there is still not much known about odonates. Understanding any of these methods of patterning first necessitates an investigation of the geometry of the specimen, and thus these prior studies on both geometry and pattern formation contextualize our interest in how we can apply a geometrical approach to studying wing formation in odonates. In this paper, we identify distinct regions in the odonate wings by first developing the capabilities to characterize these regions as geometrically distinct from others. After we are able to quantitatively identify patterns in this way, we then use this geometric insight to motivate a hypothesis about the mechanism of wing cell pattern formation.
Chapter 2

Methods

Biological Imaging

Our wing data was imaged in two different ways: the first, preliminary method imaged the wings by laying them on a flat scanner, and the second, more precise method imaged the wings with a microscope. For this second method, odonate wings were dissected off of specimens and mounted onto acrylic boxes that were cut from a sheet of 1.5 mm thick clear acrylic. We then imaged the wings with transmitted white light on a Zeiss AxioZoom.V16 microscope at the Harvard Center for Biological Imaging (HCBI). Using a 35× optical magnification, we tiled the wings (i.e., captured small sections of the wing to be stitched together later, allowing us to capture a higher resolution image) using Zen Pro software. We also captured up to 10 Z-stacks at each position in order to account
for the fact that the corrugated nature of the wings means that different sections of the wing require different levels of focus - at different Z-positions (heights) along the wing, some areas are more in focus than others. These slight differences in focus along the Z-positions of the wing can be seen in Figure 2.1. We then combine the most in-focus tiles of each Z-stack in order to obtain our final image.

The images were stitched together using Fiji image processing software, and the most in-focus tiles were combined using an extended depth of focus functionality. This flattens the wing so that we can examine it in two dimensions - while this removes the

---

1See: http://fiji.sc/

2http://imagej.net/Extended_Depth_of_Field
information about the corrugation in the wing, we expect that this reduction in dimensionality has little impact on the geometric calculations we are interested in. Although the complex corrugation in dragonfly wings is essential to their aerodynamic nature [1], and thus should not be disregarded when thinking about odonate wing functionality, in our case the differences in height created by the corrugation are negligible in relation to the length of the wing. This is evident in Figure 2.1, which shows that the necessity of adjusting focus along the height of the wing is minimal. We could likely image only with one Z-stack (i.e., take a picture of a flat image, rather than imaging along multiple heights), but we include multiple for increased image clarity. We also reduce the size of the image to 3600 pixels in the horizontal direction. Post-processing of the images consisted of manually fixing broken veins in damaged wings and turning stained cells (especially one stained cell that appears in the upper corner of the wing, called the pterostigma) in the image white. We were conservative in deciding when to fix damaged veins in order to prevent further issues with later segmentation processes. Figure 2.2 depicts the different images obtained through these methods, as well as an example of an edited wing. For each of the wings, we assign an abbreviated title based on the species identification and measure the wing length. This information is displayed in Figure 2.3.

Computational Segmentation

In order to calculate the statistics of interest, such as neighbor number for each cell, we transformed the raw data into a usable form. From looking at a small subsection of
Figure 2.2: This plot shows the different images obtained from the two imaging methods. The top image shows a raw image from the scanner imaging method. The middle image shows a raw image from the microscope imaging method, and the bottom image shows the same image after it has been manually edited. The main differences between the middle image (raw image) and the bottom image (edited image) are that we whiten the pterostigma (stained cell in upper right corner of the wing) and remove the background. These edits are not essential, but they aid in the speed and accuracy of the fast marching method.

the raw data obtained from the scanner imaging method in the left image of Figure 2.4, we see that determining the number of neighbors is an ambiguous question, largely due to the black veins that separate different cells from one another. In order to resolve
### 2.2. Computational Segmentation

#### Figure 2.3: For each wing in this table, we assign an abbreviated name for the insect. We also measure each wing's length so that we will be able to examine scaling relationships among wings.

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
<th>Abbreviated name</th>
<th>Left or right</th>
<th>Forewing or hindwing</th>
<th>Wing length (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aeshna</em></td>
<td><em>constricta</em></td>
<td>Aesh-con</td>
<td>R</td>
<td>F</td>
<td>47</td>
</tr>
<tr>
<td><em>Aeshna</em></td>
<td><em>verticalis</em></td>
<td>Aesh-vert</td>
<td>R</td>
<td>F</td>
<td>52</td>
</tr>
<tr>
<td><em>Calopteryx</em></td>
<td><em>augustipennis</em></td>
<td>Calop</td>
<td>R</td>
<td>F</td>
<td>40</td>
</tr>
<tr>
<td><em>Epitheca</em></td>
<td><em>cynosura</em></td>
<td>Epith</td>
<td>R</td>
<td>F</td>
<td>29</td>
</tr>
<tr>
<td><em>Erythemis</em></td>
<td><em>simplicicollis</em></td>
<td>Eryth</td>
<td>R</td>
<td>F</td>
<td>32</td>
</tr>
<tr>
<td><em>Ischnura</em></td>
<td><em>posita</em></td>
<td>Isch</td>
<td>R</td>
<td>F</td>
<td>17</td>
</tr>
<tr>
<td><em>Lestes</em></td>
<td><em>rectangularis</em></td>
<td>Lest</td>
<td>R</td>
<td>F</td>
<td>22</td>
</tr>
<tr>
<td><em>Somatochlorella</em></td>
<td><em>tenebrosa</em></td>
<td>Somat</td>
<td>L</td>
<td>F</td>
<td>37</td>
</tr>
<tr>
<td><em>Sympetrum</em></td>
<td><em>vicinum</em></td>
<td>Symp-vic</td>
<td>L</td>
<td>F</td>
<td>25</td>
</tr>
<tr>
<td><em>Sympetrum</em></td>
<td><em>rubicundulum</em></td>
<td>Symp-rub</td>
<td>R</td>
<td>F</td>
<td>26</td>
</tr>
</tbody>
</table>

#### Figure 2.4: The left image shows a section of the raw data from scanner imaging method, and the right image shows the same section of raw data that has had the fast marching method applied. We can see that the image on the right has much more easily identifiable direct contact between cells, whereas the image on the left does not have clear direct contact between cells, due to the presence of the black veins.
Figure 2.5: This figure depicts the start to finish image processing pipeline. We begin with the raw data and then binarize the image through either a local background subtraction process (for images obtained through scanner imaging method) or through a global thresholding process (for images obtained through microscope imaging method). We then segmented the image by identifying connected components in the image, which allows us to find the seed points for the fast marching method. Finally, we implement the fast marching method utilizing these seed points and a background velocity matrix that is dependent on proximity to a vein in the wing (the velocity slows down as the expanding contour approaches a vein, which mitigates the impact of improper seeding), which allows us to obtain a final segmented image.

This issue and facilitate the computation of statistics like size and edge number, we develop an image segmentation technique that utilizes a numerical method called the fast marching method. This method serves two main purposes: it removes the vein material by expanding out from the center of the cells, thereby establishing contact between neighboring cells; and it segments the cells into distinct objects by determining cell boundaries. This allows us to identify which cells are in direct contact with one another, as seen in the right image of Figure 2.4, which is the left image of the figure that has had the fast marching method applied. Figure 2.5 showcases the full pipeline from start to finish of image processing, beginning with the raw wing data and ending with a segmented image with a unique ID for each cell in the wing.
Introduction to the Fast Marching Method

The fast marching method is a numerical method to march or expand outward along level sets from a set of seed points at a given velocity to find the fastest possible arrival time at any point in a domain [18]. A byproduct of this calculation is that boundaries are formed when the level set contours collide (these boundary points take equally long to reach from at least two points). Computationally, this method finds numerical solutions to the Eikonal equation, which is defined as

$$F(x)|\nabla T_v(x)| = 1,$$  \hspace{1cm} (2.1)

where $F(x)$ represents the travel speed of each point $x$. The fast marching method therefore presents an algorithm for computing the solution to this equation, $T_v(x)$, which solves for the travel time for each propagating contour (i.e., every point $x$ within one of our wing cells) to reach a given point on the boundary, $v$, which in our case would represent a point on a vein.

These contours will continue to propagate outwards until they come into contact with another contour, at which point they will cease expansion. In addition to a set of seed points, the fast marching method can also utilize a specified velocity field that will allow the contours to expand at different speeds depending on their location [19]. We can alter
2.3. *Introduction to the Fast Marching Method*  

Chapter 2. *Methods*

this speed by adjusting $F(x)$ in the Eikonal equation. The fast marching method will then determine boundaries between the initial seed points by expanding outward from them at the specified speed until the contours collide.

In order to illustrate the idea of the fast march, we initialize a simple example that showcases how the method is able to create boundaries based on a seeding input and velocity matrix input. In Figure 2.6, we initialize a simple background image and subject it to two different velocity matrices: one that is a uniform velocity field and another that is position-dependent. The second column in Figure 2.6 depicts these two velocity fields, with lighter colors corresponding to higher velocities, such that velocities in the bottom right corner of the matrix are higher. This figure demonstrates how a different speed matrix can impact the result of the fast marching method.

We begin to march outwards from the seed points, which are shown in blue in the first column of Figure 2.6, depending on the speed matrix in the second column. The differential effects of this speed matrix are seen in the third column of Figure 2.6, which plots the travel time to reach each point in the image. While in the first row of Figure 2.6, this travel time appears to be consistent with expanding at a constant radius outwards from each initial point, in the second row we see that in the bottom right corner of the image, we travel a greater distance in a shorter period of time due to the increased speed of the expanding contours in this region. We can thus see that for the image that was subjected to a non-uniform speed map, some contours expanded out more quickly than others, resulting in different boundaries than the image subjected to the uniform speed map. This occurs because the contours will collide at different times and thus, meet at different locations. This differential velocity field not only more accurately
Chapter 2. Methods

2.3. Introduction to the Fast Marching Method

Figure 2.6: Demonstration of fast marching method on simple example. The first row showcases an example of fast marching method with a uniform speed map; the second row showcases an example with a non-uniform speed map. In the last two columns, lighter colors correspond to higher values - that is, in the second column, the yellow color represents the fastest speed, while the black color represents the slowest speed. Similarly, in the third column, the lighter the color is, the longer it takes to reach that point.

captures the distinction between veins and internal cell matter in our data, but also minimizes the detrimental effect of improperly identified seed points. That is, by varying the velocity, we combat any seeding errors in the data because when the fast march contour approaches a vein, it will travel a smaller distance (i.e., slow down), as depicted in Figure 2.7.
2.4. Seeding for Fast March

Figure 2.7: These contour plots for a sample cell showcase fast marching method contour propagation progression at each time step from a seed point (represented by a blue dot). We see that the effect of errors in seeding is mitigated - the resulting boundaries of the shape are captured well in both images, despite the left image having an off-center seed point. This improper seeding does not significantly impact the fast march because the velocity of the contours decreases as they approach a vein (boundary) of the cell.

Seeding for Fast March

Figure 2.8: Section of the raw data that showcases inconsistencies in lighting due to preliminary imaging errors. For example, along the upper edge of the wing, we can see some shadowing caused by the corrugated edge of the wing casting a shadow onto the scanner.

In order to find seed points for the fast marching method, we develop two different methods meant to handle images of variable quality. The methods both begin with transforming the raw wing data into a binary image and then finding the connected components of this image to locate all of the wing cells. We then find the centers of mass
for these cells to serve as our seed points for our fast march contours to expand from. For the higher quality images that we obtained from the AxioZoom.V16 microscope, we subject each pixel in the image to a global threshold to determine whether the pixel is black or white. Because our images were imaged using transmitted light, this method works well - there are no issues in our image where there is a very bright or very dark spot. However, for the lower quality images that we obtained from the scanner imaging method, our methodology for binarizing the image is slightly more involved - we utilize a local background subtraction method that allows us to remedy lighting inconsistencies in the images. That is, for each pixel in the image, we consider a local area around the pixel when determining whether that pixel should be black or white in the binary image.

**Seeding for Images from AxioZoom.V16 Microscope.** For images obtained through the AxioZoom.V16 microscope imaging method, we are able to directly transform the image into a reduced binary image by first translating the image to grayscale and then subjecting each pixel to a threshold. From there, we can find the connected components of the binary image to locate distinct wing cells. We cannot perform this direct, global thresholding method on the initial images due to the different intensity distributions found throughout the image, as demonstrated in Figure 2.8. Thus, our preliminary images require a local approach depending on the intensity distribution of a specific area in the wing.

**Seeding for Images from Scanner.** In contrast to the global thresholding method of binarizing the image utilized for the microscope images, our initial images obtained with a scanner require a more finessed approach to combat the poorer image quality. We face issues in this wing data with differential lighting depending on the section of
2.4. Seeding for Fast March

Chapter 2. Methods

the image, which largely stems from the fact that the scanner imaging method creates more shadows and bright spots in the image due to the corrugated nature of the wings, as seen in Figure 2.8. Though one approach would be to utilize a watershed algorithm to preliminarily segment the image, these lighting inconsistencies led us to instead take a local background subtraction method to binarize the image and identify distinct wing cells. This method ensures that our computational tools are robust enough to handle imperfect or damaged images.

Our background subtraction method involves first transforming our raw data into a grayscale image, then iterating through every pixel in the image and identifying some area window, \( w \), around the pixel. For this local window, we then calculate the mean and standard deviation of the intensities, with a completely black value representing 0 and a completely white value representing 1. For this pixel, we then determine whether it will be identified as black or white in our final binary image by calculating a normalized intensity score that we subject to a threshold intensity score, \( z \). This thresholding process is explained in Algorithm 1 and shown in Figure 2.9.

**Data:** Raw Image  
**Result:** Binarized Image  
Setup binarized matrix of 0’s \( M \) with dimensions of Raw Image;  
**for** Pixel \( i, j \) **in** Image **do**  
  Find pixels in window size \( w \) centered around pixel;  
  Compute mean, \( \mu \) and standard deviation \( \sigma \);  
  **if** \( (\text{Pixel} - \mu)/\sigma > z \) **then**  
    \( M[i, j] = 1 \);  
**end**

**Algorithm 1:** Binarizing Raw Data to Find Seeds for Fast March
Figure 2.9: For two different sections of a wing, we showcase the local background subtraction technique. In the top image, we showcase the small sections that we crop from the wing (a red box and a blue box demonstrate this). Each row corresponds to the local background subtraction for each section - we begin with the raw data, then plot a histogram of the pixel intensities for that data. The purple dotted line in the histogram represents the mean intensity for the given section. The next image in each row then depicts a heatmap plot of the image of the local score for each pixel (calculation for local score described in Algorithm 1). Finally, the last image demonstrates the binarized image from the background subtraction approach. We can see how despite issues with lighting (as demonstrated by the very different intensity histograms for the two sections), we are able to obtain a binarized image that captures the cell boundaries.
This results of the local binary thresholding procedure on two different subsections of the wing are shown in Figure 2.9, demonstrating how the local approach to translating the raw image into a binary one resolves differential lighting issues to successfully segment the wing. In the final image, we are able to clearly identify distinct cells.

From the binary image, we then identify the connected components, subject to a size cutoff that is dependent on the particular wing, as some wings contain larger cells on average. We then find the center of mass of each cell; these centroids are then utilized as our seed points for the fast marching method.

One consideration in our local threshold approach is the variability of the two parameters in our algorithm: local window size, \( w \), and an intensity cutoff, \( z \). Optimal parameters for each wing were determined through a parameter sweep, but these parameters were similar enough among wings that we could feasibly utilize a single parameter set and still obtain a successful segmentation. In Figure 2.10 we can see that there are slight differences in preliminary wing segmentation quality (i.e., identification of connected components from performing the thresholding process) depending on the \( w \) and \( z \) values that we choose, but these will not end up having a noticeable impact on the quality of our final segmentation from the fast marching method. Some potential errors that we may notice are seen in Figure 2.11. While incorrectly creating vein spots in the internal matter of the cell will not severely impact the segmentation accuracy, incorrectly connecting the cells that should be kept separate will be problematic, since we would treat the cells as one object rather than multiple. Therefore, we would only identify one seed point where there should be multiple. Due to this issue, we err on the side of utilizing a small \( z \) value, which prevents the scenario of connecting multiple cells seen
Figure 2.10: Table of images obtained from setting different \( w \) and \( z \) values for the local binary thresholding algorithm and finding the connected components of the binarized images. We see subtle differences in the success of the preliminary segmentation, but ultimately these differences will not drastically impact our ability to fully segment the image through our fast marching method approach, due to our inclusion of a differential velocity depending on proximity to veins. This velocity matrix will prevent errors in boundary determination due to mistakes in seeding that might occur from improper connected components analysis.

in Figure 2.11. As long as we are able to capture most of the internal cell matter through this thresholding approach, we will be able to find a reasonable seed point for the cell.
Figure 2.11: Potential errors resulting from connected components analysis: top image results from a $z$ value that is too large, thus connecting areas that should be kept separate because the intensity cutoff is not able to correctly capture the cell boundaries (i.e., venation structure); bottom image results from a $z$ value that is too small, thus creating black vein spots in the internal matter of the cell, since the intensity cutoff is not strict enough. Only the top image will cause issues in the segmentation, since we will not identify enough seed points (as we treat the three green objects as one cell, for example); thus, we err on the side of a smaller $z$ value to prevent this.

Velocity Matrix

Once we have obtained our seed points, we then determine our velocity matrix, where each position in the matrix corresponds to an appropriate speed for that position (closer to the center of the cell corresponds to a faster speed). That is, when expanding outwards from the center seed point, the contours expand more quickly, but when the contour becomes closer to the veins surrounding the cell, the speed of expansion slows down. This location-dependent expansion velocity is again shown in Figure 2.7, which depicts the expanding contour from the seed point of a cell at each time step. Although we are
expanding into and thereby eliminating the vein pixels in the raw data, we still aim to preserve the correct boundary information.

In order to obtain this, we begin with two speeds for the binarized image, with black corresponding to a slow speed and white corresponding to a fast speed. We then apply a Gaussian blur with a relatively small width of 2.25 pixels to the image, such that we obtain a larger range of speeds that scale with distance from the veins. Lower values in the velocity matrix correspond with slower speeds, thus resulting in a longer travel time to reach that pixel. Because our velocity matrix comes from a binary image of our wing, our velocity values are between 0 and 1, with a speed of 1 inside the cell region and a speed of $\epsilon$ in the vein material, where $\epsilon$ is a small value that we varied between 0.01 and 0.033. The tradeoff for the $\epsilon$ value was that a higher value (i.e., faster speed) was much faster to run, while a lower value (i.e., slower speed) would be expected to yield the most accurate boundary identification. We typically chose an $\epsilon$ value of 0.033 to maximize this tradeoff consideration, as it was able to identify the boundary correctly without being overly computationally slow. This method of determining a velocity matrix works well independent of the wing size - that is, since we base the velocities off of a binarized image of a wing, the speeds will naturally scale depending on the exact image. If, for example, one wing has a much thicker vein than another, then the speeds will adjust to be slower through these regions. Thus, the speed parameter $\epsilon$ stays constant, regardless of the wing we are considering.

Once we have determined the velocity matrix and seed points, we then implement the fast marching method by utilizing the scikit-fmm package.‡

‡https://pypi.python.org/pypi/scikit-fmm
2.6. **Mask Creation**

In order to remove unnecessary background information from the image, we create a mask that separates the wing from the background in the data. We create this mask by beginning with a binarized version of the raw data and expanding outwards from a background point in the data until we reach a point on the outside of the wing. In a separate matrix that we initialize to ones, we turn these positions into zeroes, and this serves as our preliminary mask. We also then refine the mask by ensuring small objects around the edge of the image are completely eliminated. In order to do this refinement, we find the IDs of pixels that are removed in the preliminary mask creation and ensure that all pixels with any of these IDs are assigned to zero in the mask. We can then multiply this masking matrix with our data matrix, resulting in a clean image that shows only the wing against a black background.

This process of mask creation and refinement is shown in Figure 2.12 and explained in Algorithm 2.
Figure 2.12: Top to bottom: Segmented image without mask applied; segmented image with preliminary mask applied, prior to refinement; zoomed in image of segmented image with preliminary mask, showcasing how there are small objects on the edge that are not completely removed; segmented image with refined mask applied. The reason we see many small objects in the initial image (top image) is that in order to speed up our fast marching method and capture the outside boundary of the wing, we initialize many seed points outside of the wing. This results in the identification of many small objects that are not actually cells of the wing; thus, we find the mask in order to remove these objects and create a visually clean image.
2.7. Statistic Calculation

Algorithm 2: Mask Creation and Refinement

<table>
<thead>
<tr>
<th>Data: Binarized Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result: Mask</td>
</tr>
<tr>
<td>Setup matrix of 1’s $M$ with dimensions of Raw Image;</td>
</tr>
<tr>
<td><strong>for</strong> Pixel $i$, $j$ in Binarized Image <strong>do</strong></td>
</tr>
<tr>
<td>Find pixels in the background using Breadth-First Search;</td>
</tr>
<tr>
<td>Store positions of background pixels, Background Positions;</td>
</tr>
<tr>
<td><strong>for</strong> Position in Background Positions <strong>do</strong></td>
</tr>
<tr>
<td>$M[\text{Position}] = 0$;</td>
</tr>
<tr>
<td><strong>end</strong></td>
</tr>
<tr>
<td><strong>end</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data: Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result: Refined Mask</td>
</tr>
<tr>
<td><strong>for</strong> Position in Mask <strong>do</strong></td>
</tr>
<tr>
<td>Find ID of Position, i.e. find Data[Position] where Data is the original Image;</td>
</tr>
<tr>
<td>Find all positions in Data with that ID, Data Positions;</td>
</tr>
<tr>
<td><strong>for</strong> Position in Data Positions <strong>do</strong></td>
</tr>
<tr>
<td>$M[\text{Position}] = 0$;</td>
</tr>
<tr>
<td><strong>end</strong></td>
</tr>
<tr>
<td><strong>end</strong></td>
</tr>
</tbody>
</table>

Statistic Calculation

In the segmented images, each cell has a distinct identification number. Thus, we are able to straightforwardly calculate some statistics by iterating through each wing's cells. This allows us to answer the questions that motivate this paper: namely, how a geometric analysis of the cells in odonate wings can help us to understand the process of wing
2.7. Statistic Calculation

Figure 2.13: Depiction of two shapes that would have differing $\sigma/\mu$ values; the shape in question is colored blue. Each red length represents a length to center calculation, which is computed for each of the cell’s neighbors and is the distance between the cell’s center of mass and a neighboring cell’s center of mass. The figure on the left would have a high $\sigma/\mu$ value due to the disordered packing of the blue shape in the context of its neighboring shapes, whereas the figure to the right would have a low $\sigma/\mu$ value, since it is very ordered - it is a hexagon surrounded by other hexagons on all sides.

formation. In order to address this aim, we first must be able to consider many geometric characteristics of the wing cells.

**SIZE.** Since each cell in the segmented image has a unique identification number, we are able to calculate size by counting how many appearances of each ID are in the matrix.

**SHAPE ANISOTROPY MEASURE.** For each cell, we also calculate a statistic of the shape's uniformity, or its anisotropy. In order to do so, we calculate the distance from the center of cell $i$ to the centers of each of its $j$ neighboring cells, $\ell_{c(i,j)}$. We then calculate the mean and standard deviation of these distances to find $\sigma/\mu$ for cell $i$, which serves as
our anisotropy measure,

$$\frac{\sigma}{\mu}_i = \sum_{j \in \text{neighbors}} \frac{\sqrt{\left(\ell_{c(i,j)} - \langle \ell_{c(i,j)} \rangle \right)^2}}{\langle \ell_{c(i,j)} \rangle \sqrt{\text{number of lengths}}}.$$  \hspace{1cm} (2.2)

The shape to the left in Figure 2.13 would have a relatively high $\sigma/\mu$ measure as compared to the shape to the right, which has sides of more uniform length. This is important not only for determining the nature of the shape of each cell, but also for examining the packing structure of cells with one another.

This $\sigma/\mu$ statistic is scale-free because we divide by the mean length, ensuring that the statistic is not inflated for one cell because its $\ell_{c(i,j)}$ measures are larger on average than another cell’s $\ell_{c(i,j)}$ measures. For example, for a large cell that has large surrounding neighbors, we would not want the $\sigma/\mu$ statistic to be larger than a small cell that has small surrounding neighbors if the packing structure is of the same uniformity; we would want these two instances to be represented by the same value of $\sigma/\mu$.

**Shape Calculation.** Since our fast marching method removes vein boundary information between cells, we recapture boundary information by walking (i.e., tracing) directly around the outside of the shape. This will allow us to identify various characteristics of the cell, such as its neighbor number and its vertices. Additionally, identifying the pertinent boundary points will eventually allow us to determine how many edges a cell has. This method of walking around is especially vital as it ensures sequential ordering of points. As seen in Figure 2.14, which depicts a particular cell from a wing that has the boundary plotted around it, the walking around method precisely captures the boundary.
Figure 2.14: Heatmap plot of walking around a shape’s immediate outer boundary, such that we begin at the top left at the red point on the perimeter of the shape and end at the white point directly to the left of the starting point (a lighter color corresponds to a higher value, thus the lightest color (white) represents the end of the sequence). We see that our walking around method obtains sequential identification of boundary points.

information one pixel away from the cell. Furthermore, we see that the ordering is also successfully captured – the walking around begins at the top left with the red point and ends at the white point directly to the left of the starting point.

In order to begin walking around the boundary of a shape, we first locate a point directly outside of the shape and choose our starting direction to be upwards. We then cycle through four different directions – up, right, down, and left in a clockwise motion around the perimeter of the shape following the right-hand rule. That is, we trace along the outside of the shape like one would if one was walking around a building with one’s right hand tracing around the outside of the wall. Once we have returned back to our start position and our starting direction, we are then finished walking around the entire shape. This walking around procedure is described in further detail in Algorithm 3.
Data: Wing Data for Cell ID
Result: Ordered Boundary Points for Cell

Find point immediately outside of cell to begin walking around, start position;
Initialize direction, up;
Move according to procedure below until we have returned to start position and our direction is up;

while True do
    if direction is up then
        if position above current position is outside of cell then
            move up;
        if position right of current position is outside of cell then
            change direction to right;
        else
            direction is left
    if direction is right then
        if position right of current position is outside of cell then
            move right;
        if position below current position is outside of cell then
            change direction to down;
        else
            direction is up
    if direction is down then
        if position below current position is outside of cell then
            move down;
        if position left of current position is outside of cell then
            change direction to left;
        else
            direction is right
    if direction is left then
        if position left of current position is outside of cell then
            move left;
        if position above current position is outside of cell then
            change direction to up;
        else
            direction is down
end

Algorithm 3: Walking Around Method for Cell Boundary Identification
2.7. Statistic Calculation

**Edge Number.** Once we have obtained our boundary positions for each cell, we can utilize these positions to determine a measure of edge number. This statistic is crucial not only because it is a defining geometric characteristic, but also because it is currently very tedious to calculate manually. While interpretations on how many edges a shape has can be ambiguous, a reasonable edge number calculation still represents a visually recognizable and intuitive measure that can allow us to quantify visually identifiable patterns in the wings, such as those that involve packing of cells. Thus, while edge number is not a completely precise statistic (i.e., since the interpretation of how many edges a shape has can vary widely), it is a useful tool for grouping the cells.

**Elimination Method Based on Angle and Length Cutoff.** A polygonal shape can be defined by its vertices, or points on the shape where two straight lines meet. Since these points can be connected in sequential order to redraw the shape, our first step in finding the number of edges for a shape involves first locating the vertices for the shape. In order to do so, we find locations on the boundary where the pixel ID changes from one position to the next; that is, the vertex in our image can be represented by a point where three different cells meet, as seen in Figure 2.15. We then connect the vertices in order to obtain a preliminary polygon like the one shown in Figure 2.15 that allows us to obtain an initial edge number count. From there, we consider two different situations: one in which a vertex is nearly collinear with adjacent points, and another in which the edge that is defined between two vertices (i.e., the line drawn between two vertices) is of insufficient length.

To determine where vertex points are collinear with surrounding vertex points, we iterate over all triplet sets of points and calculate the associated angle. For a set of three vertex
Figure 2.15: This plot depicts vertex identification based on boundary positions located during walking around process. Vertices are identified as points on the immediate outer boundary that change IDs as they move to the next position in the sequential list, as shown in the zoomed in section of the plot. In this plot, the vertex points for the dark blue shape are in red. We use these vertices to represent the shape.

points, \([A, B, C]\), we find \(\vec{v}_1 = A - B\), \(\vec{v}_2 = C - B\); then we calculate the associated angle as

\[
\theta = \cos^{-1}\left(\frac{\vec{v}_1 \cdot \vec{v}_2}{||\vec{v}_1|| ||\vec{v}_2||}\right).
\]  

(2.3)

If this angle \(\theta\) is greater than a specified length cutoff, which is close to \(\pi\), then we exclude
this vertex from our shape definition, because it is not essential to the creation of an edge.

From our reduced set of vertices, we then redraw the shape and look at the second scenario of removing edges that are of insufficient length to be considered in our edge number calculation. In order to do so, we iterate over each pair of vertices and calculate the distance between them. We then subject the distances to a specified length cutoff, removing the distances that are too short. We then count the number of distances that do meet the length cutoff, and this represents our edge number - put simply, it is the number of distinct edges in the shape that are long enough to be classified as edges. This process of first truncating edge number based on internal angles and then on edge length is showcased in Figure 2.16.

One potential concern with our vertex elimination method of subjecting the internal angles of the shape to an angle threshold is that we could over-eliminate vertices if there are two angles next to one another in a shape that exceed the angle threshold (i.e., the midpoints of both of these angles would be eliminated). Instead, one would likely want to first eliminate one vertex, then redraw the shape, instead of removing multiple vertices at once. This would tend to occur along edges that appear to be more curved in nature. We examined the possibility of this scenario occurring in our dataset, subject to different angle thresholds, and found that only between ~ 0.419% and 1.94% of all cells would be impacted. Thus, we decide not to consider this a consequential issue and utilize the vertex elimination method described above.
2.7. Statistic Calculation

Figure 2.16: Process of determining edge number based on set of vertices. The left image shows the initial set of vertices, followed by the process of checking angles to determine which vertices are collinear. We then show the reduced set of vertices redrawn with collinear vertices removed, and the right image shows the final edge number calculated based on the number of lengths that meet length threshold.

**Optimization of Parameters.** Our edge number determination methodology is largely reliant on two inputs: an angle cutoff and a length cutoff. Determining edge number is not always an obvious problem, but our aim was to best represent how an average human would identify the number of edges for a given shape. Our goal is to simplify the wing geometry enough to create a reduced geometrical map that can help us in formulating hypotheses about wing pattern formation and biological variation between wings without losing too much information about the intricacies of the wing.

Thus, we chose to optimize three parameters: $a$, $b$, and $d$. These optimized parameters were then used to calculate the angle and length cutoffs referenced in the previous section, such that
2.7. Statistic Calculation

\[
\text{length cutoff} = a \times (\text{cell size})^b 
\]  

(2.4)

and

\[
\text{Angle cutoff} = d .
\]  

(2.5)

We propose that the length cutoff is a function of the cell size, since larger cells will have longer sides on average, and that the angle cutoff is not size-dependent.

In order to optimize these parameters, we utilized human identification for the number of edges of shapes to serve as our ground truth values. This involved two parts: a collaborator first classified over 100 wing cell shapes, and then we also collected shape identification data from 20 different participants in a survey of 40 different shapes in order to make sure that our initial identification was reasonable. As seen in Figure 2.17, the interpretation of number of edges is a more complex problem than it may appear. See Appendix for full survey data results.

We then optimize our parameters by minimizing a cost function that calculates the total difference between our true values and the predicted values. Our predicted values are obtained through the number of edge elimination algorithm described in the previous section, using the length and angle cutoffs defined in Equations 2.4 and 2.5. That is, we minimize cost where
Figure 2.17: Sample of data collection for human shape identification. We created a survey of different shapes, a subset of which is seen in the first column. The next two columns depict a breakdown of the responses that we received; we can see that there is a high degree of variation in responses, indicating that human interpretation of shapes is not uniform. Full survey data results available in Appendix.

\[
cost = \sum_{\text{cells}} |\text{predicted edge number} - \text{true edge number}| \quad (2.6)
\]

We optimize these parameters by utilizing the Nelder–Mead method; this method allows us to minimize our non-differentiable cost function.\(^\text{6}\)

\(^\text{6}\)http://scholarpedia.org/article/Nelder-Mead_algorithm
Due to the stochastic nature of the Nelder–Mead algorithm, we initialize 30 random seeds to begin the optimization, and we find that the $a$, $b$, and $d$ parameters converge to a small range of values. This is somewhat expected; for example, if the shape was a square, we might expect that $b$ would converge to a value around 0.5 (making $\text{size}^b$ the length of one edge of the shape) and that $a$ would converge to some value $< 1$, making $a \times \text{size}^b$ some small subsection of the length of one edge. Furthermore, since we hypothesize that the angle parameter, $d$, is likely independent of the specific shape in question and is also likely to be close to $\pi$, this convergence is also a somewhat expected result. This intuition is largely confirmed – $a$ converges to between $\sim 0.12$ and $0.23$, $b$ converges to a range between $\sim 0.47$ and $0.53$, and $d$ converges to around $2.7$, or $\sim 0.86 \times \pi$. For our edge number determination algorithm, we utilize the parameter set that best minimized the cost function, $[0.23, 0.49, 2.71]$.

As seen in Figure 2.18, our cost function tends to converge to values around 50, which is reasonable considering the degree of variation we see in human identification. For reference, our training set has over 100 shapes, with an average edge number of about five per shape. Thus, we have over 500 edges to classify. In looking at Figure 2.17, we can already see that there is often not consensus in shape identification - oftentimes if the mode response is $x$, there will be at least one respondent each who identify the shape as having $x-1$ or $x+1$ edges. In the context of this human variability, we see that our automated shape identification performs reasonably well, with slightly less variation than we might expect for human shape identification.

**Neighbor Number.** In addition to edge number, we can also utilize our boundary positions to calculate a neighbor number statistic. We define a cell’s neighbor to be a...
Figure 2.18: This plot shows the value of the cost function over each iteration for each of the 30 initial seeds, which tends to converge to values around 50. The minimization of the cost function allows us to optimize our parameter values, $a$, $b$, and $d$, for our edge number determination algorithm.

Figure 2.19: Two examples of identifying neighbor number for a cell. The cell in question in each image is colored black. In the left image, the cell has four distinct neighbors, marked by white x's. In the right image, the cell has at least six distinct neighbors, marked by white x's, but it is unclear whether the cell marked by a yellow question mark can be considered a neighbor, indicating that identifying neighbor number can be an ambiguous problem.

cell that shares an edge with that cell. For example, the cell colored in black in the left image in Figure 2.19 appears to have four distinct neighbors, as marked by the white x's. However, determining whether a cell is considered to share an edge with another is often not well defined, as evidenced by the right image in Figure 2.19. The black cell appears
Chapter 2. Methods

2.7. Statistic Calculation

Figure 2.20: Our interest in calculating a difference metric of neighbor number minus edge number is partly rooted in a belief that this misalignment may indicate that the associated cell has a structural importance in the wing, much like the misaligned pattern of bricks lends increased strength to the structure of a wall. In the left image, we show a brick wall with aligned bricks laid one on top of another, while in the right image we show a brick wall with misaligned bricks that are offset in their placement. We can see that in the misaligned brick image, the bricks would have a nonzero neighbor number minus edge number metric, while the aligned brick image would have a zero metric.

to have at least six neighbors, as indicated by the white x’s, but it is arguable whether the cell marked with the yellow question mark would be considered a neighbor. We will not consider this edge long enough to be called a neighbor since we are most concerned with being able to geometrically classify different cells. Thus, in order to combat this issue of neighbor ambiguity, we set a length cutoff to eliminate neighboring cells whose edges are not considered long enough to be sharing an edge with the cell in question. In contrast to our length cutoff explained in Equation 2.4 from the edge number determination method, which is dependent on object size, we take this length cutoff for neighbor determination to be a constant value because we do not believe neighbor determination to be size-dependent. We utilize the boundary position information to find the cell IDs at each boundary position and then take the unique list of cell IDs that contain the minimum number of pixels required to be considered a long enough edge of contact with the cell to find the number of neighbors. Computing neighbor number also allows
2.7. **Statistic Calculation**

us to calculate a difference metric of neighbor number minus edge number - neighbor number and edge number are not necessarily equal, and a distinction between these statistics may indicate that a cell plays an important functional role for the wing. That is, when cells have a non-zero difference metric, this indicates they are misaligned. An elucidating example is a brick wall: rather than one brick being laid directly on top of another, bricks are laid in a misaligned pattern, as shown in Figure 2.20. In a brick wall, this offers additional structural support to the form; it is possible that in odonate wings, these cells may perform a similar function by stabilizing and strengthening the wing.

![Figure 2.20: An example of a brick wall showing misaligned bricks.](image-url)

Figure 2.20: This image depicts an example of a brick wall showing misaligned bricks. Bricks are laid in a pattern where one brick is not directly on top of another, offering additional structural support to the form. This may represent a similar function in odonate wings, where cells with a non-zero difference metric perform a stabilizing and strengthening role.

**CURVATURE.** Although we largely consider the wing cells to be well-estimated by polygons, it is also important to consider the curved nature of the cells, especially because a high degree of curvature may allude to an important mechanical function in the structure of the wing - for example, they may allow the wing to bend more easily. In particular, we find that there are some cells that are conserved among many wings, and these cells are often characterized by having one or more highly curved edges. In Figure 2.21, for example, we can see that the orange cell marked by a white x has a very curved edge.

![Figure 2.21: An example of a highly curved cell.](image-url)

Figure 2.21: This image depicts an example of a highly curved cell in a wing. This cell, which is orange and marked by a white x, is often seen in other wings in a similar form, showcasing how curvature can be a useful measure for mapping cells to one another across wings.
cells similar to this one are often seen in a similar position in other wings. Due to the conserved nature of these curved cells and their potential contribution to the mechanical function of the wing, it is important to be able to identify them and compare them across wings.

Thus, in order to move beyond simply reducing the wing cells to polygons, we also calculate a curvature measure for each edge of the wing. The first step in performing this calculation is to locate the points along each edge; we achieve this by utilizing our list of vertices in conjunction with our set of boundary points for each cell. Since our boundary points are ordered sequentially, we are able to identify the boundary points that fall between two vertices, thus allowing us to segment the full set of boundary points into sets of points along each edge. In order to find curvature along each edge, we then fit a circle to the points that define each edge with \texttt{scipy.optimize.leastsq},\footnote{http://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.leastsq.} which

Figure 2.22: This plot showcases the circle fitting process: on the left, we see an image of the fitted least square circle, which is calculated from a set of points along an edge of a cell; on the right, we see a zoomed in image of this circle, which allows us to see more clearly the ability of the least square circle fitting to the points. This circle fitting process allows us to calculate a curvature measure along each edge of each cell.
allows us to minimize the total distance between the data points and the fitted circle by optimizing over three parameters: center \((x_c, y_c)\) and radius \(r\). Figure 2.22 depicts one example of this fitting procedure for a set of points, showcasing that this least squares circle approximates the points well.

After obtaining this radius parameter value, \(R\), we then compute curvature, which is \(\kappa = 1/R\). This allows us to identify regions that have higher curvature than others and thus, gives us the ability to identify the curved shapes that are apparent among many wings.

Beyond just its applications to odonate wings, a streamlined process for calculating curvature along edges of a shape can be utilized in many different contexts. In particular, one can imagine it being very useful for instances where researchers may be interested in tracking change of shape of an object as it moves.
Chapter 3

Results and Discussion

Researchers from a wide range of disciplines have studied pattern formation. While the scope and nature of the questions they asked vary widely, the data that they utilized was similar: it resembled the form of data we have in our odonate wing images. With the capabilities that we have developed for examining geometric statistics, we can therefore begin to address many different questions related to pattern formation, beyond just what we see in odonate wings. Developmentally, being able to identify distinctions in neighbor number and edge number may tell us something about the temporality of wing formation - cells that have no difference between their neighbor number and edge number, for example, likely form in tandem, in contrast with cells that are misaligned with their neighboring cells. Mechanically, being able to assess the packing structure of cells by looking at neighboring geometric statistics throughout the wing may aid robotics research for areas like insect flight simulation; it may allow us to engineer wings that not only look generally like an insect wing, but actually emulate the root properties of wing geometry that exist in nature and echo how insect wings function as elegantly as they do. Finally, since our dataset covers a wide range of different odonate wings, we are also
able to provide insight into the scaling relationships between wing length and various geometric statistics that may tell us how functional differences that exist as a variant of wing length may also be reflected in geometric trends.

With these goals in mind, we begin by examining different geometric properties in our wing data that reveal several patterns that allow us to identify characteristics of distinct regions in wings. The ability to recognize these different regions leads us to hypothesize about the process of odonate wing formation, such as the order in which venation occurs. We also begin to explore the question of how different geometric properties may scale with wing length. While the aim of this thesis is mainly centered around creating a viable method for analysis of complex images like those of odonate wings, this preliminary geometric examination allows us to develop hypotheses that can be explored in the future with the tools that we have already developed.

**Geometric Pattern Exploration**

With the computational capabilities that we have developed, we begin by examining different geometric characteristics of the wings. The exploratory plots of these characteristics then motivate our belief that distinct regions in the odonate wings can be identified based on their geometric nature.

One characteristic that we examine is cell size. In Figure 3.1, where we color each cell based on its size, we can observe some trends in the wing based solely on cell size. For
Figure 3.1: We color the cells in each of the wings of a full set of dragonfly wings (left forewing, right forewing, left hindwing, right hindwing) based on their size, scaled out of 1.0 (i.e., 1.0 represents the maximum size, whereas 0.0 represents the minimum size). Here, we can see that there are some distinctions between the left and right wings (i.e., in comparing the left and right forewings, while the general trend of sizes appears similar, there are slight differences, as indicated by a difference in color). This plot also indicates a general trend of decreasing cell size as we move towards the outer bottom perimeter of the wing, which may be a useful metric when considering mechanical aspects of the wing, such as flexibility.

For example, we see that as we move towards the bottom edge of each wing, the cells tend to decrease in size, as indicated by the cooler colors. Furthermore, we see that there are certain cells that are anomalies - a few cells that are very large. These larger than average cells will be examined later as possibly belonging to the conserved regions in the wing, and it is likely that they may provide some structural or mechanical function for the wing. By conserved, we mean that they are seen across different odonate wings, both within and between species.

In addition to observing size trends within a specific specimen’s wings, we also examined trends in average cell size across different specimens. Because our dataset encompassed
3.1. Geometric Pattern Exploration

Chapter 3. Results and Discussion

Figure 3.2: In this plot, we examine the relationship between wing length and average cell size for the wing. The data points for average size are plotted in blue. We also plot the standard deviation of the sizes for each wing, which are depicted in red. From this plot, we can see a general decreasing average size as the length of the wing increases. See Figure 2.3 for abbreviated species names.

wings of many different lengths, ranging from 17 mm to 52 mm, as depicted in Figure 3.6, we were able to compare trends across many different odonates on the basis of the overall wing length. As seen in Figure 3.2 we find that an increased wing length appears to be associated with a decrease in the average cell size. Similarly, Figure 3.5 depicts a decreasing trend in average cell perimeter against increasing wing length. This relationship between average cell size and wing length may also be best illustrated by comparing two of our wings that depict this trend - in Figure 3.4, for example, we compare surface area distributions for various wings, such that we can determine what size cells tend to occupy most of the wing surface area. Here, we find that a shorter wing, such as that of 46
Chapter 3. Results and Discussion  

3.1. Geometric Pattern Exploration

Figure 3.3: In this plot, we examine the relationship between the length of the wing and the number of objects (i.e., cells) in the wing. From this plot, we can see that as the wing length increases, the number of objects tends to follow an increasing trend as well. This, in conjunction with the decreasing average cell size trend seen in Figure 3.2, appears to support our hypothesis that an increased number of smaller cells (i.e., the fingerprint region) is seen in larger wings. See Figure 2.3 for abbreviated species names.

Epith, is not as dominated by small cells as a large wing like Calop is. This supports our hypothesis that as the size of the wing increases, the associated decrease in average cell size may be in part due to the increased size of what we call the “fingerprint region.” We define the fingerprint region to be the region in the bottom right quadrant of the right hand wings and the bottom left quadrant of the left hand wings that is dominated by smaller cells. The nomenclature for the fingerprint region stems from our belief that these areas are largely randomized or unique for each wing, whereas other areas appear to be more conserved. Our comparison of the number of cells in each wing against the length of the wing, as seen in Figure 3.3, indicates that an increase in wing length is associated with an increase in number of cells. This could also imply an increased...
Figure 3.4: This plot shows the wing surface area distribution by normalized object size for various wings. In other words, along the x-axis we have the normalized object size, where 1.0 is the maximum object size, and along the y-axis we have the proportion of the wing that this size object accounts for. For example, Calop (which is also a larger wing, as can be seen in earlier plots since it has a long length) is peaked at an object size of 0.2, indicating that most of its surface area is dominated by smaller cells. In contrast, a wing like that of Epith, which happens to have a shorter wing length than Calop, is peaked at about 0.3, indicating that its surface area is not as dominated by small cells as Calop’s is. This further supports our hypothesis that larger wings have a larger fingerprint region that is composed of mainly small cells. See Figure 2.3 for abbreviated species names.

fingerprint region area in larger wings, since this would mean an increased number of smaller cells. Previous research into scaling relationships of insect wings has shown that the size of the wing accounts for about 95% of the variability of flexural stiffness
Chapter 3. Results and Discussion

3.1. Geometric Pattern Exploration

Figure 3.5: We plot the wing length of different wings against their corresponding average cell perimeter. As we might expect, considering the result of Figure 3.2 that showcases a decreasing average cell size with increased wing length, we also see a general trend of decreasing average cell perimeter as the wing length increases.

In wings [20]. Thus, in future work, comparing across wings based on wing length may prove to be useful in examining how different mechanical characteristics of the wing (how much weight they can withstand, their flapping speed) scale with wing size.

Closely related to the statistic of cell size, we also calculate the perimeter value for each cell, as seen in Figure 3.7. Compared to Figure 3.1, there appear to be even more similarities across wings - namely, between the left/right forewings and the left/right hindwings. For example, the set of two cells that form a V-shape are visible in both sets of wings and are colored the same. This higher degree of similarity, as opposed to another measure like size, may be due to the fact that the calculation of perimeter is of linear order (i.e., the units would be in pixels) rather than of quadratic order (where area would be measured in squared pixels).
### 3.1. Geometric Pattern Exploration

#### Chapter 3. Results and Discussion

<table>
<thead>
<tr>
<th>Wing length (mm)</th>
<th>Wing image</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td><img src="image1.png" alt="Image" /></td>
</tr>
<tr>
<td>22</td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td>25</td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>26</td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>29</td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td>32</td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>37</td>
<td><img src="image7.png" alt="Image" /></td>
</tr>
<tr>
<td>40</td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td>47</td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
<tr>
<td>52</td>
<td><img src="image10.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Figure 3.6: In this table, we showcase the range of different wing lengths that we have in our dataset. The shortest length is 17 mm, while the longest length is 52 mm - we thus are able to examine scaling relationships among wings on the order of about three times magnitude of our shortest length.

Another characteristic we are interested in is the edge number of each cell, which helps in characterizing the structural organization of the wing. As seen in Figure 3.8, where we color each cell based on the number of edges it has, a preliminary plot of this geometric statistic already begins to uncover a distinction between different areas of the wing. The dark green colored areas in Figure 3.8 represent four-sided shapes; we can see that in the
Chapter 3. *Results and Discussion*

3.1. *Geometric Pattern Exploration*

Figure 3.7: We color the cells based on their perimeter value, scaled out of 1.0 (i.e., 1.0 represents the maximum perimeter, whereas 0.0 represents the minimum perimeter). We can see similarities across the forewings and the hindwings, such as the two reddish colored cells that form a V-shape in the forewings. This also exists in the hindwings, where we see a set of two yellow-green colored cells that form a V-shape. Thus, we see that perimeter may be a statistic that is largely conserved across fore and hindwings.

...wings, there are stretches of these colored regions that extend along the upper edge of the wing and along a curve within the wing interior. This is particularly evident in the left hindwing. These predominantly four-sided regions are visually distinguishable from the regions lower in the wing that have mainly five (light green) and six-sided (orange) shapes. Furthermore, we also observe that this area dominated by five and six-sided shapes has a smaller average cell size than the regions composed of the rectangular shapes - this is confirmed by the plot in Figure 3.1, since the lower regions in the wing are colored darker blue and purple, indicating a smaller cell size. Consequently, we already begin to see the utility in examining the geometric nature of the wing: by combining information about different statistics like size and edge number, we can glean how
3.1. Geometric Pattern Exploration

Chapter 3. Results and Discussion

Figure 3.8: We color the cells in each of the wings of a full set of dragonfly wings based on their edge number. We can see that there are already some emerging patterns based on this statistic - there are some long stretches of four-sided shapes that appear along curves in the wings and along the top edge of each wing, as indicated by a dark green color. Again, we also observe that while the general edge number trend is similar among left/right wings, there is not a direct mapping of one cell to another (i.e., the left/right wings are not a perfect mirror image). This result may be somewhat surprising, as we might expect that the left/right wings may appear almost identical in terms of edge number for reasons like balance, since edge number is fundamental to the structure of the wings.

different regions may be distinct from one another.

In addition to examining the edge number of each cell, we are also interested in learning about the edge number of each cell relative to its neighbors - that is, one can imagine that if a cell that is four-sided is surrounded by, on average, four-sided cells, this may say something about the ordered nature of that region as opposed to another that has large discrepancies between a given cell's edge number and the edge numbers of its neighbors. We examine this idea of looking at neighboring edge number statistics in two different ways: 1) by finding the mean edge number of each cell's neighbors and 2) by finding the previous statistic (the mean edge number of each cell's neighbors) and taking the
3.1. Geometric Pattern Exploration

In order to examine the structure of each cell relative to its neighboring cells, we also calculate neighboring statistics (i.e., calculate a statistic of each cell's neighbors). Here, we color each cell by the mean edge number value of its neighbors - for example, if a cell has four neighbors with five edges each, then that cell will be assigned a value of five. This allows us to determine how cells are organized relative to one another, and we can see in this plot that cells in the bottom region of the wing tend to have a higher mean edge number of its neighboring cells, as compared to the upper regions.

Figure 3.9: In order to examine the structure of each cell relative to its neighboring cells, we also calculate neighboring statistics (i.e., calculate a statistic of each cell's neighbors). Here, we color each cell by the mean edge number value of its neighbors - for example, if a cell has four neighbors with five edges each, then that cell will be assigned a value of five. This allows us to determine how cells are organized relative to one another, and we can see in this plot that cells in the bottom region of the wing tend to have a higher mean edge number of its neighboring cells, as compared to the upper regions.

The first method gives us an idea of how the cells are packed together - for example, if all of the cells ended up having the same mean edge number of its neighbors, then that would tell us that the wing was very ordered and uniform in terms of the number of edges of each polygon. The second method is related, but also is able to indicate how each cell's mean edge number of its neighbors is related to that particular cell - for example, if a cell itself has four edges, but its mean edge number of its neighbors is also four, we would be able to distinguish that from a cell that has four edges but has a mean edge number of its neighbors of five. Figure 3.9 showcases the first method, while Figure 3.10 showcases the second method. Examining these measures uncovers a distinction between the areas.
3.1. Geometric Pattern Exploration

Chapter 3. Results and Discussion

Figure 3.10: Similarly to in Figure 3.9, we plot a measure of each cell's neighboring cells. Here, we color each cell by the mean edge of its neighboring cells minus its own edge number. For example, if a cell has four edges and its neighboring cells have a mean edge number of four, it would be assigned a value of zero. From this plot, we can see that most values oscillate around zero, though in general the lower regions of the wings tend to have more extreme values than in the upper regions. This is in line with our belief that the lower regions of the wing tend to be less ordered, in contrast with the top edge of the wing which has mainly four-sided objects.

We believe are conserved - the upper edges of the wings and the curve that appears to bisect each wing - and the other regions of the wing. In particular, these conserved areas tend to have less extreme values in both of these plots. In Figure 3.9, these conserved regions are a darker green color, as compared to the regions below the conserved regions, which tend to be brighter green or yellow. In Figure 3.10, the conserved regions are near-white, whereas the regions below them tend to have more extreme values of darker pink or green. While these observations may not be enough to conclusively state that the stretches of four-sided shapes that we see in Figure 3.8 are conserved, they do uncover some distinct patterns that differentiate these regions from the rest of the wing.
Figure 3.11: We calculate neighbor number for each cell (i.e., the number of cells that share an edge with the cell of interest) and color each cell by its neighbor number, where black represents 0 neighbors and white represents 10 neighbors.

Figure 3.12: For each cell, we calculate the mean neighbor number of its neighbors - for example, if the cell we are looking at has four neighbors with five neighbors each, then it would be assigned a value of five.
3.1. Geometric Pattern Exploration

Chapter 3. Results and Discussion

Figure 3.13: For each cell, we find the neighbor number and the edge number. We then find the difference metric, which we take to be neighbor number minus edge number. We then color each cell by this difference metric in order to find whether there are regions that appear to have largely misaligned cells (i.e., a nonzero difference metric) and vice versa. We find that there are some regions in the wing that have a difference metric of zero, such as the regions below the stretch of four-sided cells. Furthermore, we see that many of the cells that are large have a high difference metric, indicating that they may play an important structural role in the wing.

Another measure that we believe will help quantify patterns in the wing is that of neighbor number, which allows us to showcase the difference between neighbor number and edge number. In addition to cells with a high difference metric playing a functional role in the wing, this misalignment of neighbor number and edge number can also aid us in determining the order in which cells are formed, since we presume that cells that have no difference between neighbor number and edge number are likely laid down concurrently. In Figure 3.11, we color the cells by neighbor number; in Figure 3.12, we then employ a similar method of examining the behavior of each cell’s neighboring cells by coloring the wing cells by the average number of neighbors that its neighbors
has. Here, it appears that there is not much of a clear trend as far as certain regions having homogeneous coloring. However, when we plot the difference metric of each wing cell (i.e., number of neighbors minus the number of edges), we begin to see some outliers, as seen in Figure 3.13 - while most of the differences are around zero, there are some cells that have a high difference value of four or five, as indicated by red and gray cells, respectively. These cells also happen to be some of the largest cells in the wing. Thus, it may be likely that these cells play an important role in the function of the wing. Furthermore, plotting the difference metric allows us to see that there are regions that are dominated by a zero value, which may help us to test our belief that these regions are laid down concurrently. These regions are particularly evident in the forewings, in the section below what we believe to be the conserved stretch of four-sided cells.

Cell edges colored by curvature measure

Figure 3.14: For each cell edge, we calculate a curvature measure. Here we color each edge by its curvature measure, and we can begin to see that some cells have highly curved edges.
Curvature is another measure that we are interested in examining, particularly because it moves beyond simply reducing all of the cell shapes to polygonal ones. Furthermore, we also utilize curvature to identify cells that we believe to be conserved across wings, as visually it appears that certain cells that seem to appear across wings within a specimen are unique because of their highly curved edges. Thus, these cells are compelling not only because of their consistent presence in the wings of the odonate, but also because of their somewhat unusual appearance - in contrast with the generally regular, polygonal structure of the other cells, these cells stand out visually as being curved, suggesting that they may play an important role developmentally or structurally for the insect. For example, the fact that the edges of these recurring cells have high associated measures of curvature may suggest that they play a mechanical role in the wing, such as allowing it to fold or flap along those edges more easily. In Figure 3.14, we color each edge of the cell
based on its curvature metric. One particular example where this measure may prove useful is shown in Figure 3.15, where we highlight two cells that are conserved across the forewings - and which also happen to have high measures of curvature along their edges. Thus, we see how this metric can be utilized to identify cells that we believe to be conserved; in the case of these two cells, it would have been difficult to differentiate them from other cells in the wing by another measure like size or edge number, but by utilizing the curvature measure, we are able to locate and examine them.

**Identifying Regions by Geometric Character**

Through our investigation of different geometric properties of odonate wings, we hypothesize some of the laws that may govern pattern formation. Namely, we propose that through geometric trends that we observe, we are able to identify distinct regions in the wing that we can show are distinct in quantitative terms. This idea of identifying distinct regions is not altogether new - it has been suggested that different regions of the wing can be classified based on the corrugation of that area (hollow, flat, or ridged) [21]. However, in our case we focus solely on the two-dimensional nature of the wing to deduce regions based on a flat geometrical map. First, we summarize the findings from our geometric examination: from looking at cell size across wings, we see a general trend of decreasing cell size as we approach the outer bottom perimeter of the wing; from looking at number of edges, it seems that there are stretches of four-sided shapes that appear to be conserved among wings, and that regions outside of these predominantly
four-sided stretches tend to have higher edge number variability; and from looking at a difference metric of number of neighbors minus number of edges, it appears that there are certain regions that have a zero metric, which indicates that they may be formed at the same time. These observations of geometric trends motivate our hypothesis that we can parse different regions in the wing.

From this preliminary presumption of different regions, working with Seth Donoughe, a PhD student in the Harvard Organismic and Evolutionary Biology Department, we hypothesized that the stretches of four-sided shapes seen among the wings - one along the upper edge and one that bisects the wing along a curve - lie between longitudinal veins in the wing that are highly conserved. Our collaborator also suggested more veins that are also believed to be conserved. Thus, since these rectangular shapes lie between two major longitudinal veins, we suggest that the region separation criteria may be partially dependent on the number of longitudinal veins that a cell touches. This conclusion is also partly based on the hypothesis that shapes that touch more longitudinal veins are more restricted; that is, they have fewer degrees of freedom for number of edges, as they are not bound on any given side by a longitudinal vein. This would fit our observations that there appears to be more shape variability in the regions outside of these conserved four-sided stretches. To illustrate this hypothesis that regions can be categorized on the basis of how many conserved longitudinal veins that a cell touches, in Figure 3.16 we highlight shapes based on whether they touch 0, 1, or 2 longitudinal veins. We are able to see similarities between these regions and some of the regions in Figure 3.13 that had a value of 0 (a teal color), suggesting an order in which some of the regions in Figure 3.16 are laid down.
Figure 3.16: We hypothesize that we can identify three different regions in the wing, depending on how many conserved longitudinal veins each cell touches. That is, a cell can either touch 0, 1, or 2 conserved veins. In this figure, we depict what we would consider these regions would be, according to our hypothesized conserved vein identification. The dotted lines indicate the location of conserved veins, the cells highlighted in purple touch no longitudinal veins, the cells highlighted in blue touch one longitudinal vein, and the cells highlighted in green touch two longitudinal veins. Upon identifying these potential regions, we then aim to quantify how these regions are in fact distinguishable on the basis of geometric statistics such as the distribution of cell sizes and edge numbers in these regions.

After we categorized cells into these three categories based on the number of longitudinal veins each one is in contact with, we are able to compare statistics of each region's geometry to support our hypothesis that these regions can be considered distinct.
3.2. Identifying Regions by Geometric Character

Chapter 3. Results and Discussion

Figure 3.17: This plot showcases the edge number distribution of each of the identified regions in Figure 3.16. We see that the region of cells that touch two longitudinal veins is peaked at an edge number of four, which supports our observation that these regions are mainly dominated by quadrilateral cells. Furthermore, we see that these distributions are peaked at five for the remaining regions, which also supports our observation that these regions contain cells with more sides than in the previously discussed region.

is a particularly compelling statistic that distinguishes these regions - in Figure 3.18, we see that the distribution of sizes for cells that touch two major veins is peaked at a higher value than for cells that only touch one or no veins. These size values also fall within a smaller range than in the other two categories, demonstrating that there is some measure of uniformity between cells that is more present in cells that fall between two longitudinal veins. This may aid our belief that these areas that fall between two conserved veins are also conserved because there is some existing consistency that makes the cells in these regions more similar to each other than to other cells. Furthermore, we also examine
Chapter 3. Results and Discussion  3.2. Identifying Regions by Geometric Character

Figure 3.18: This plot showcases the size density plot for each of the identified regions. We see that the curve highlighted in red, which is the density plot for the region that has cells touching two longitudinal veins, is peaked at a higher value than in the other two regions. Furthermore, we see that the region that touches no longitudinal veins, which is represented by the blue curve, has the lowest peaked value of the three regions. These two observations support our belief that the region of cells that touch two longitudinal veins is generally composed of larger cells, while the region of cells that touch no longitudinal veins tends to contain smaller cells, some of which are part of the fingerprint region. The size density plot is computed by fitting a univariate kernel density estimate to the size data.

the distribution of edge numbers for the different categories in order to test our belief that cells that touch two longitudinal veins tend to be mostly four-sided, while cells that touch one or none have more variability in edge number. As seen in Figure 3.17, we find that the distribution of edge numbers for the cells that touch two longitudinal veins is peaked at four with a relatively small variance, as compared to the other two categories, which display a higher degree of variation. Consequently, this plot matches our prior belief that cells that touch two longitudinal veins may be more restricted in terms of edge
number because they are bound on two sides. Finally, as stated earlier, our presumption that cells that do not have a discrepancy between neighbor number and edge number are laid down concurrently would also strengthen our belief that these regions are distinct in a temporal sense. For instance, the regions that touch no longitudinal veins are similar to regions in Figure 3.13 in terms of position in the wing. Hence, perhaps the cells in this region (that touch no longitudinal veins) are laid down at the same time, which would distinguish them from the other regions on the basis of time of formation.

While these stipulations of region identification are only preliminary and require further investigation, we can already begin to see proof of concept in terms of the importance of geometric examination in determining patterns in odonate wings.

**Hypothesized Pattern Formation**

Our observations of the geometric nature of different regions of the wing lead us to hypothesize that the wing formation process involves an ordering; that is, there is some sequential nature to the forming of the different regions. Since the cells that fall between two longitudinal veins appear to be distinct from the other cells, we propose that first the longitudinal veins are laid down, followed by the cells that lie in between these cells, which form through segmentation of the regions between the veins.

In the fingerprint region of the wing, it appears that visually, the Voronoi diagram may be an appropriate model for the wing cells. We test this theory by taking the seeds to be the
3.3. Hypothesized Pattern Formation

Figure 3.19: This image depicts our hypothesized pattern formation mechanism. We employ a Voronoi segmentation, utilizing the centroids of the cells that either touch one or no longitudinal veins. As we can see in the figure, the Voronoi diagram does not approximate the wing pattern as well as one might expect - it appears that the cells need to be stretched in a non-uniform way (i.e., depends on the particular cell) in order to fit the shapes of the wing cells.

centroids of each cell in the region between the longitudinal veins and performing the Voronoi diagram to find cell boundaries using Voro++, a C++ library [22]. As depicted in Figure 3.19, where we apply the Voronoi segmentation to the seed points in the regions of the wing that touch no longitudinal veins, we find that this theory is actually flawed - while the general shapes produced generally resemble those observed in an odonate wing, it appears that they need to be subjected to some non-uniform stretching in order to fill the full region in the wing. Consequently, the Voronoi cells do not appear to be
an accurate approximation for the wing cells, since the Voronoi approximation is not conserved under a stretch transformation.
Understanding the geometric character of odonate wings is useful for a variety of reasons: developmentally, it may help to explain the process of venation and wing formation by showcasing patterns in the wing; mechanically, the geometry of the wing may contribute to how quickly an odonate is able to fly, how flexible the wing is, or how much breakage it is able to take before being nonfunctional; and computationally, automated image processing of an odonate wing can be applied in a multitude of biological contexts to understand the geometry of different specimens. Beyond just biological imaging applications, this set of computational tools can be used more broadly in any field where analysis of images is vital, and especially where that analysis is currently performed in a largely manual way. In fact, the original motivation for developing these tools was not for geometric examination of odonate wings, but for research in gastrulation of *Gryllus bimaculatus* (cricket) by studying images from embryo development. We are motivated not only by questions related specifically to odonate wings, but also by the prospect of developing a computational software toolkit that is able to handle automatic processing of complex images, even ones that are not of the highest image quality.
By developing computational tools and examining odonate wings in a geometrical context, we were able to not only theorize about wing formation processes, but were also able to showcase some of the flaws in our hypothesis for wing cell patterning. Since we are now able to compute geometric characteristics of a given image in a fast and automated way, we can test future hypotheses by comparing the geometric statistics of a wing formed from our hypothesized method to an actual specimen wing. That is, these geometric computation capabilities allow us to test whether a wing created from our hypotheses of pattern formation truly emulates the geometric trends seen in odonate wings. In the future, we hope to implement an anisotropic growth factor to the wing cell formation in the fingerprint region as well as account for the tiling of regular polygons within a curved boundary. Similar problems have been encountered by a group studying bubbles, which led to the development of the Voronoi Implicit Interface Method (VIIM) [10]. Furthermore, we aim to examine other aspects of the materiality of the wing that we do not consider in this paper, such as the vein and membrane thickness, a metric that has been preliminarily studied as it relates to flight aerodynamics [23]. While this is only the beginning of our understanding of the rules that govern odonate wing formation, we have established a foundation that will prove very useful moving forward.

Thus, what is perhaps most notable about the research conducted to write this paper is that we have developed a robust set of tools that have widespread applications. While here we have offered a preliminary example as applied to odonate wings, the tool was developed with a plethora of different contexts in mind, and further, what has been shown is only a portion of what the tool can do - while things like size and edge number
are important in themselves, we can also combine these measures in different ways to find new metrics that may yield powerful results. All of the possibilities for how these tools can be utilized and what discoveries they may help uncover are yet unknown, but with these computational abilities now developed, we will be able to tackle many unanswered research questions.
Bibliography


1997.


Appendix

This table shows the full survey data results for human shape identification.* For each of the 40 shapes, we recorded the response for the number of edges for that shape; here, we also compute the mean, minimum, and maximum response as well as the variance in responses.

<table>
<thead>
<tr>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
<th>R8</th>
<th>R9</th>
<th>R10</th>
<th>R11</th>
<th>R12</th>
<th>R13</th>
<th>R14</th>
<th>R15</th>
<th>R16</th>
<th>R17</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>4.70588</td>
<td>4.</td>
<td>6.</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5.41176</td>
<td>5.</td>
<td>6.</td>
<td>0.257353</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>4.94118</td>
<td>4.</td>
<td>6.</td>
<td>0.183824</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.05882</td>
<td>5.</td>
<td>6.</td>
<td>0.0588235</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.23529</td>
<td>5.</td>
<td>8.</td>
<td>0.566176</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4.</td>
<td>4.</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.23529</td>
<td>5.</td>
<td>8.</td>
<td>0.566176</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.05882</td>
<td>5.</td>
<td>6.</td>
<td>0.0588235</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5.34118</td>
<td>5.</td>
<td>7.</td>
<td>0.308824</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.11765</td>
<td>4.</td>
<td>8.</td>
<td>1.73529</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4.</td>
<td>4.</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>5.17658</td>
<td>4.</td>
<td>8.</td>
<td>0.764706</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.</td>
<td>5.</td>
<td>0.</td>
<td></td>
</tr>
</tbody>
</table>

*http://tinyurl.com/jzv2887
<table>
<thead>
<tr>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
<th>R8</th>
<th>R9</th>
<th>R10</th>
<th>R11</th>
<th>R12</th>
<th>R13</th>
<th>R14</th>
<th>R15</th>
<th>R16</th>
<th>R17</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>4.41176</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>4.70588</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4.47059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5.52941</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>4.41176</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4.47059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5.76471</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6.05882</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6.11765</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5.35294</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.28412</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.64706</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6.05882</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6.11765</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4.41176</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.17647</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5.17647</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5.17647</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>