STRAWBERRY REPORT

1. Abstract

This research applies shape analysis to images of strawberries for the purpose of acquiring information about the target strawberry's spatial orientation. The research goal is to locate the stem of a strawberry within a digital image obtained from a strawberry harvesting robot for the purpose of picking the strawberry. Intuitively, one can look at an image of a strawberry and locate the stem without a problem. The challenge is in determining the steps involved and automating that decision process. A tool called the Medial Axis is central to our approach.

2. Introduction

In general, digital images contain a lot of information as a set of pixel values. For humans it is relatively easy to interpret this information and recognize objects, but in order for a computer to analyze the shape of the objects contained in an image we have to extract the key data points through a segmentation process. The result of image processing will produce an outline of the image's object of interest. One way of further analyzing this shape is to look at the features of its medial axis. Looking at the medial axis of a shape can be a useful tool for many different applications. In our research on predicting the location of the stem of a strawberry given a 2D image, the medial axis will provide information on the strawberry's shape and orientation and thus help to determine a prediction for the stem location. The procedure developed uses an approximation of the medial axis of the strawberries, since a digital image will give a discrete set of points representing the boundary of a shape, rather than the continuous function needed to compute the actual medial axis.

This report will begin by giving the background information on the image segmentation technique called K-Means Clustering, as well as the definition of the Medial Axis and related concepts. The second portion of the report will outline the algorithm applied to each input strawberry image. The process these images go through begins with a segmentation process which leads to finding the border of the strawberry. That is followed by computing the medial axis of the strawberry shape, simplifying the medial axis, and finally using the extracted key information to draw conclusions about the orientation of the strawberry. This will lead to determining the position of the stem within the image. An example of a strawberry image which was taken through the process described in the remainder of this report is provided and the resulting prediction for the stem location shown.

3. Background

3.1. **K-Means Clustering.** K-Means clustering is a digital imaging process that results in a segmented image obtained by grouping the pixel's based on their RGB values into similarly colored groups. The 'K' in K-Means clustering stands for the variable number of clusters produced, which is chosen depending on how many different segments we want to obtain from the input image. The process begins by selecting K random starting pixels from the image. Those individual pixels represent the initial mean of their clusters. Next, more pixels are examined and assigned to the cluster whose mean is nearest to their value. When pixels are added to a cluster, the mean is recalculated with the new value included. This assignment continues for all the pixels in the image,

Figure 1. Sample Image of a Strawberry to be picked

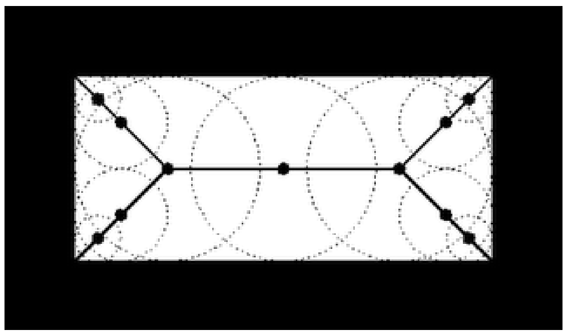

Figure 2. Medial Axis of a Rectangle [1] The connected black line is the medial axis and the dotted circles show how that medial axis line was computed. This example shows a sample of the maximal circles whose center points form the medial axis. We can see that the medial axis is giving us key information about the structure of the shape. From it, we obtain data about the bends in the boundary (look at the behaviour near the corners), as well as the symmetry and orientation.

associating each with one of the K clusters. It is an iterative process, terminating when clusters are unaffected by pixel reassignments. At this point, the image has been segmented into K parts with each cluster of pixels as one group of similarly colored pixels. This technique is applicable to shape analysis when you know the object whose shape you want to consider is of itself similarly colored, yet relatively dissimilar from the other pixels in the image.

3.2. **Medial Axis.** The medial axis is an important tool to be used in shape analysis. It can be thought of as the skeleton of a shape, representative of the overall structure.

Definition 1. *The Medial Axis of a shape is composed of the collection of centers of all the maximum-sized circles that are contained within the shape and are tangent to the boundary of the shape in at least two places.*

Because the medial axis is constructed from maximal circles, any given point on the medial axis has the property of being equidistant from the closest points on the boundary. Given a continuous function of a boundary curve, we can calculate the exact medial axis by using methods involving tangent and normal vectors and angles between them. The precise computation requires knowing the continuous boundary function from which you can derive these attributes. In cases where such

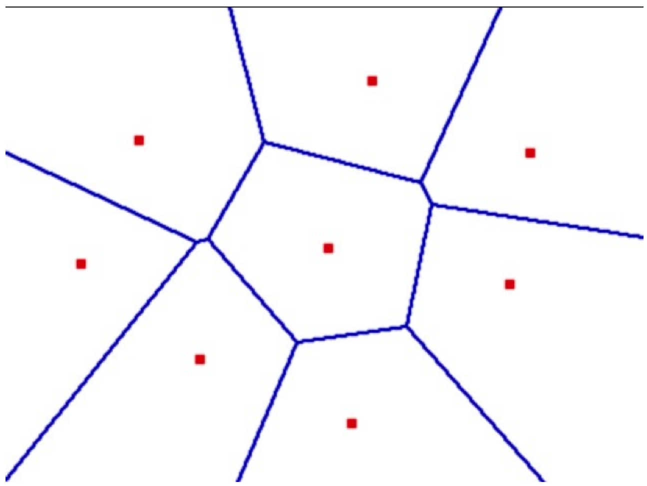

Figure 3. In this sample Voronoi Diagram, the collection of red points generate the blue equidistant lines called Voronoi edges.

a function cannot be obtained, approximation techniques are employed. Whether an approximation or an exact computation, the medial axis gives more information about a shape than the border does alone. Thus, it is helpful to use when you are curious about properties of a shape.

3.3. **Voronoi Diagram.** The Voronoi Diagram offers a way to approximate the medial axis of a shape. The reason this is useful is because when a shape does not have a continuous function representing its boundary curve, the exact medial axis cannot be directly calculated. Such is the case with a shape in a digital image, having a set of points, namely pixels, identifying the border of the shape. A Voronoi Diagram can be created from a given set of points such as these as follows:

Definition 2. *The Voronoi Diagram consists of a set of points on a plane together with the lines constituting the set of all points that are equidistant from the two nearest original points.*

Where two edges created from two different pairs of points meet is a Voronoi vertex. The link between the Voronoi Diagram and the Medial Axis can be found in the property of equidistance. Since the points on the medial axis are centers of maximal circles, their consistent radii force them to be equidistant from the nearest points on the boundary. The points that lie on the Voronoi edges are equidistant from the nearest points as well. So when the Voronoi diagram is found for an object's boundary points, the medial axis is generated. This connection is crucial and can be further stated in this Theorem:

Theorem 1. *The Voronoi vertices created in the Voronoi diagram approximate the medial axis of a curve in 2D. Furthermore, if the sample density approaches infinity, the Voronoi vertices in this case converge to the medial axis.* [2]

Since the number of boundary points extracted from an image of the object is finite, the Voronoi Diagram will be an approximation. As the number of boundary points increases and the points become more and more dense, the Voronoi Diagram gives a better and better approximation of the Medial Axis.

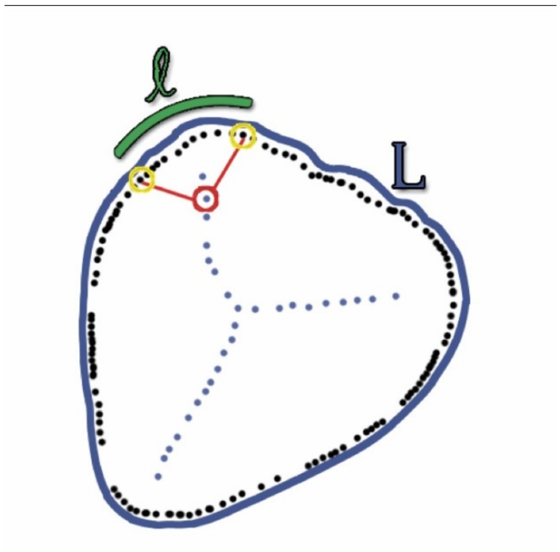

Figure 4. Shown here is a sample shape's border and associated medial points to be used in demonstrating the calculation of the Contour Ratio.

3.4. **Contour Ratio.** The contour ratio is a value that can be calculated for every point on the medial axis and it is used as a simplification technique to trim medial points. Consider the below figure as a depiction of how a Contour Ratio value can be obtained and what it really signifies.

Recall that the points on the medial axis are centers of circles contained within the shape and tangent to the boundary in at least two places. Thus, each medial point has at least two boundary points associated with it, namely its maximum circle's points of tangency. Looking at a specific point on the medial axis (such as the one circled in red), we can assign a value based on how evenly the two or more corresponding boundary points cut the boundary curve into sections. The second largest section of arc is the one we are concerned with labeled $l(P)$. Calling the total length of the boundary curve L, we can develop a ratio $r(P)$ for evaluating a medial point's significance:

$$r(P) = \frac{l(P)}{L} \tag{1}$$

The higher the value of $r(P)$ the more central we can assume that point is to the main structure of the medial axis since its points of tangency are far apart from each other breaking the boundary into even sections. A medial point that has associated border points that lie close together and do not represent the overall shape will have a low value of $r(P)$. We can use this value as a method of eliminating extraneous medial points. The way it is done is to throw out points whose Contour Ratio value is below a threshold chosen to give the desired amount of pruning. The higher the threshold value, the more points on the medial axis that will be removed, resulting in a more simplified medial axis representation of the original shape. The threshold is chosen depending on how much of the information about the boundary needs to be preserved or based on other factors specific to an image.

Figure 5. K-Means clustering: Divides the pixels into three clusters, grouping the most similarly colored pixels together and returns the separated clusters. Here is an example showing the three resulting segments for the sample strawberry image.

4. Procedure

We are going to be working with images taken by a strawberry harvesting robot for the purpose of finding the location of the stems to be picked. As already mentioned, the digital images must first be segmented to find the target object's border. Once that data is acquired, we can begin our shape analysis by computing the Medial Axis approximation from the set of boundary points using a technique such as the Voronoi Diagram. The information from the Medial Axis is useful because it allows us to see the skeleton of the shape and identify the global features. In order to capture this vital information, the medial axis must first go through a simplification process to eliminate unnecessary data points. For an ideal strawberry, once the Medial Axis has been appropriately simplified, it should consist of three branches meeting near the center. One branch leads to the bottom of the strawberry and the other two branches approach the top corners. Given the endpoints of these branches we can find the distances between each pair and compare to determine which point is at the bottom of the strawberry. If we look at the triangle formed by connecting these three key points we see that for an ideal shape strawberry it would yield a near isosceles triangle. Since the strawberry is assumed to be fairly symmetric, the distance from the bottom point of the strawberry to each point at the top of the berry will be very similar. Once the bottom point is identified, the orientation of the strawberry has been determined and thus the region where the strawberry's stem will be can be known as well. The process is explained and the steps in the procedure are visually represented with an example strawberry.

4.1. **Segmentation.** Before the strawberry shape can be analyzed, we need to segment the given image to find the strawberries and then find the exact pixels most representing their borders. First, the image is sent through the process of K-Means Clustering which groups similarly colored pixels together and separates them from those pixels that are dissimilar in color. For these images, we will let $k = 3$ because we found that we require three clusters: one for the red of the strawberries and two for the other background colors. When the process is complete and all the pixels are assigned, we should be left with a cluster of red pixels from the strawberries and two clusters of mostly green and brown pixels surrounding them.

Now the image is segmented and in order to analyze the strawberry shape, we want to find which of the three clusters contains the strawberries. The RGB pixel values from each cluster are compared and the cluster with the proportionately highest red values are chosen to be the strawberries. The selected image is then converted to binary. If there are multiple strawberries, we find the largest strawberry in the image. The other berries are ignored for now, but the remaining berries will be photographed again once the current berry being anaylzed is picked and removed. The border is found by shrinking the strawberry around the border by one pixel and removing the reduced sized strawberry from the original, leaving a border of single pixels.

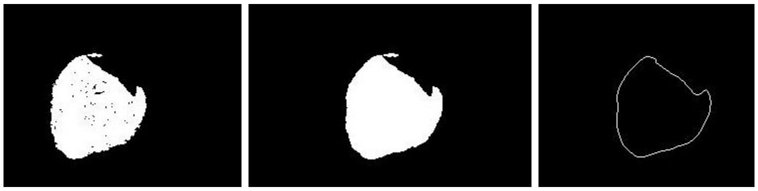

Figure 6. Binary image of strawberry, filled in and smoothed, and the border

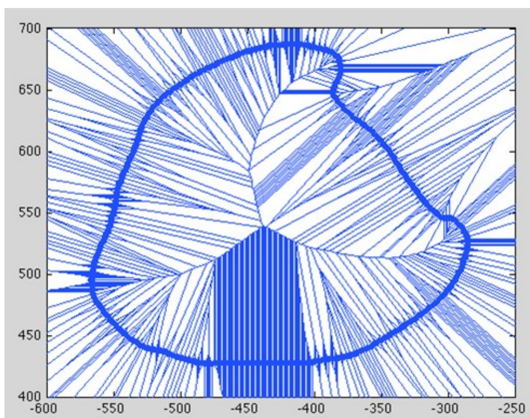

Figure 7. Voronoi Diagram produced from boundary points

From the original input image we have simplified the information, extracting the boundary of our shape as a discrete set of pixels as our points, and we can proceed to calculate the approximation of our medial axis.

4.2. **The Medial Axis.** The medial axis can be approximated using the original border of the strawberry found from the image. The aforementioned Voronoi diagram assists in this process by plotting points that are equidistant from each pair of boundary points.

The result of this technique is a lot of Voronoi points that lie outside the boundary of the shape. These points are dismissed and we are left with only interior points. We can see when we look at the plot of the medial axis that there are a lot of excess medial points that need to be trimmed down into a more simplified shape before we can further analyze its structure. The first simplification technique is smoothing the boundary by only working with the points on the convex hull. The points on the convex hull have the property that when a line is drawn between any two of the points, it is completely contained inside of the shape. This process eliminates a lot of medial points whose creation were due to boundary fluctuations and leaves us with fewer points. We can

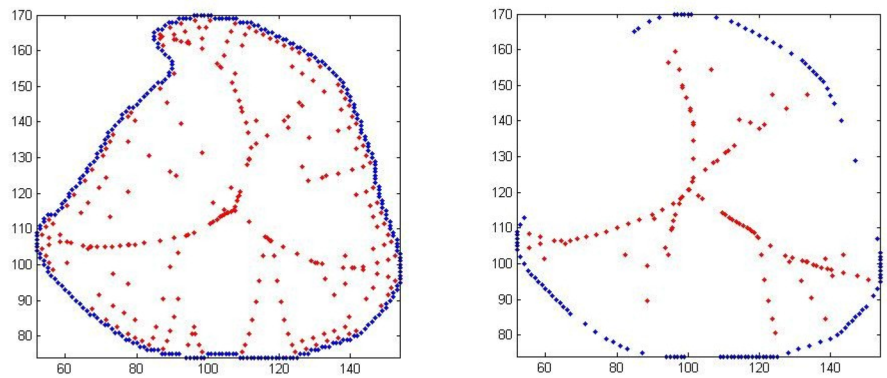

Figure 8. Medial Axis with Original boundary and after taking the Convex Hull

begin to see that this reduced set of medial points are more representative of the berry's overall shape.

The medial axis still requires simplification and we will now use the Contour Ratio as that additional simplification. Unfortunately, there is not one threshold value that produces the desired amount of trimming down of points for every strawberry image anayzed. There were six different thresholds chosen for pruning the Medial Axis and the values are 0, 0.01, 0.15, 0.3, 0.5, and 0.6. In each case, medial points with Contour Ratio values less than the selected threshold are removed from the medial axis representation. Once the Medial Axis has been calculated for each of the six threshold values, we must use that information to choose one of the simplifications to use in the stem prediction.

4.3. **Predicting the Stem.** Once the medial axis has been simplified to six different levels, we can examine the medial axis and see what information we can extract about the original shape. As stated, the pattern of the medial points should form three branches, coming together in the middle and resembling a 'Y' with the bottom point of the 'Y' at the bottom of the strawberry. To find the endpoints of the branches, the medial points are plotted and the rightmost, leftmost, highest, and lowest points on the plane are nominated to most likely represent the ends of the Y shape. If the result of this step yields four potential points, one has to be eliminated. Of those four points the two that are closest together are found and of those two the one that creates the triangle with the other two points of larger area is kept. Now using this information, the goal is to locate the top of the Y, where the stem will be, and a prediction box can be generated with the size dependant on the size of the strawberry. The determination is made by checking all the distances between each pair of endpoints, and then finding the minimum difference between all these distances. The sides of the triangle that have the most similar length will be the sides from the bottom to the two top points. The idea is that whether the berry is short or tall, as long as it is fairly symmetric the selection will be correct.

It was mentioned earlier that six threshold values were chosen to give six different degrees of simplification to the medial axis since not all berries require the same amount of pruning. However, we have to choose among these values which threshold produces the correctly simplified medial axis for each individual strawberry image so that the best stem prediction is made. The distance between

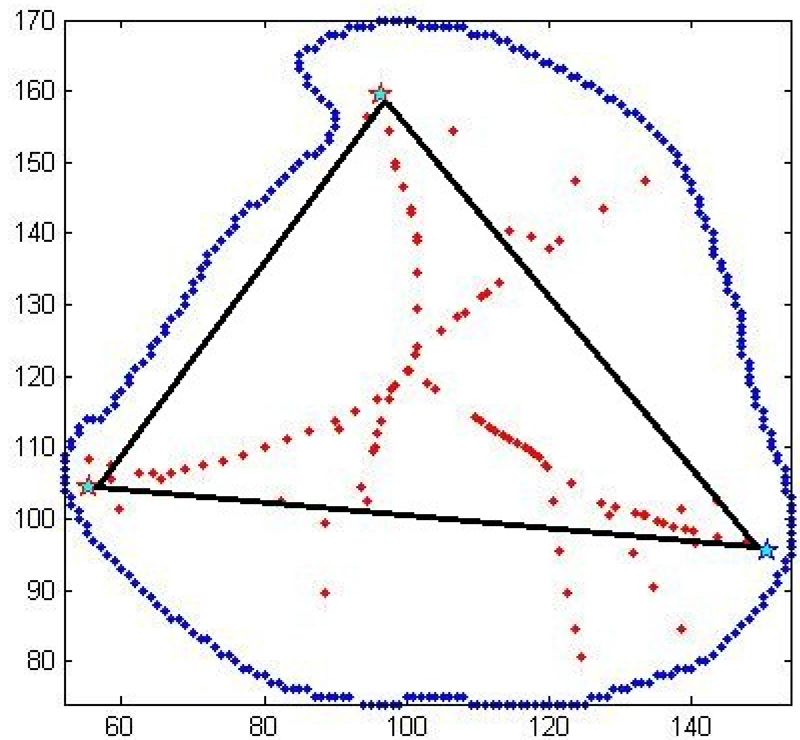

Figure 9. Simplified Medial Axis with Key Points. The blue stars with red border are the points correctly chosen as the top of the berry. The associated triangle is drawn to show how checking which are the most similar sides indicate where the top of the strawberry is.

the most similar lengths of the triangle is found using each of the contour ratio threshold values and the one selected is the minimum of the six. Once we have the orientation of the berry, we have all the information we need to draw a bounding box which will contain the stem of the strawberry. The dimensions we want for the stem box will depend on the size of the berry so the major and minor axis is computed and the length and width of the box is sized accordingly. Once the box is draw, the image inside the box is kept and the remainder of the image is blacked out.

5. Results

After testing this algorithm on a database of strawberry images, we had a success rate of 73%. The images for which the strawberry's stem location did not get properly predicted in general had certain characteristics that caused the failure. The algorithm cannot produce accurate results when

Figure 10. Prediction region for the stem of the strawberry output by our algorithm.

the input strawberry has a round or rectangular shape and there are no clear three branches, or it has an equilateral triangle shape, so the sides cannot be distinguished from the top. Finally, if the berry is obstructed in any way by a leaf or another berry, the outline of the shape cannot be extracted and thus the shape analysis ineffective.

References

[1] http://homepages.inf.ed.ac.uk/rbf/HIPR2/skeleton.htm

[2] Dey, Tamal K. Zhao, Wulue. Approximating the Medial Axis from the Voronoi Diagram with a Convergence Guarantee. 2002.