For a binay image B with structuring element s, the Top hat operation is given as:

Where the operator includes an erosion followed by a dilation. In most pictures, a stem removal can be completed using a disk shaped morphological structuring element with a radius of 12 pixels. It was observed that from the top-hat operation, the stem is consistently the largest shape, which simplified the selection of undesired shape.

Inner Distance Shape Context

The inner distance shape context method is one of the latest shape context method that can more accurately describe the characteristics of the part structures than the traditional shape context. The inner-distance is defined as the length of the shortest path between landmark points within the shape silhouette. It had been proven to be articulation insensitive.

The workflow of inner distance shape context is as followed:

After segmentation, a number of even spaced sampled points were picked from the shape contour, the connectivity between points were then tested to ensure that only the connection within shapes are allowed. The run time for this process grows exponentially as the number of sampled points grow. The connectable matches were saved as nodes, with the distance between them saved as the weight in a sparse matrix. MATLAB includes built in functions such as distances() and graphshortestpath() that can retrieve the shortest distance between nodes.

For illustration, the inner distance map has been created for an acer class leaf from the database Northen American dataset.

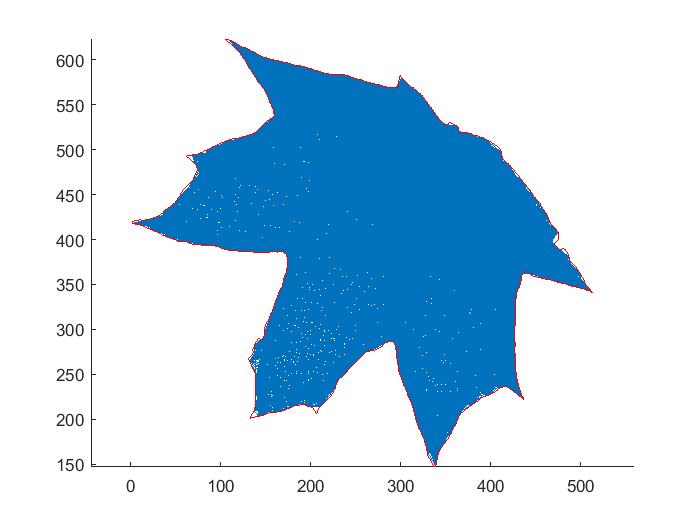
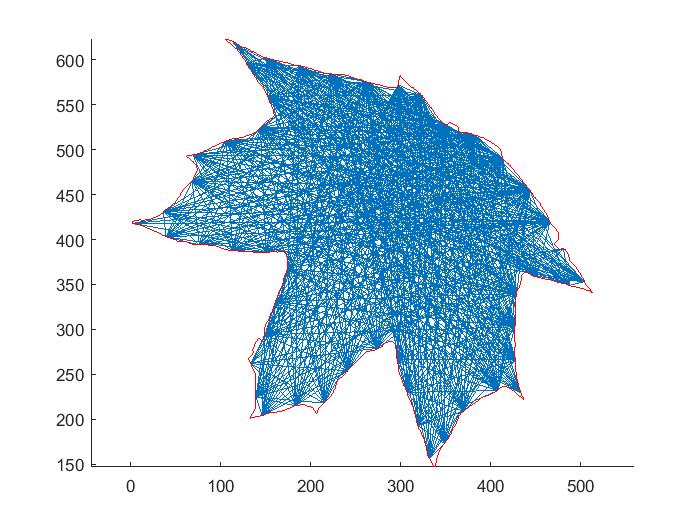
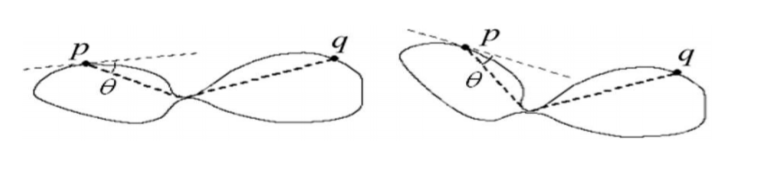


Figure , Using 50 samples (left) ,Using 100 samples (right) The blue lines are the connectable lines between sample points.

As opposed to the traditional shape context, the inner-angles, defined as the angle between the contour tangent and the inner-distance path were used in orientation bins. The contour tangent can be computed discretely using the connecting pixels.



Figure, An illustration of inner-angles

Unfortunately, a number of leaf shapes require a large number of sampled points to be categorized accurately. This leads to intolerable run time (30 to 40 seconds including segmentation and histogram building process) that rendered the operation unable to identify within a desirable time frame. For this reason, the method was not implemented in the final product.