Cell Clustering using Shape and Image Intensity Information

Kaushik Shankar

Mentors: Dilan Görür, Fred Park
University of California, Irvine

August 2011
Goal

Given an image containing the shape and concentrations of cells, cluster them based on shape and intensity information.
My Approach

Use the Multi-Scale Triangle Area Representation (TAR) of the shape to get signatures for the shapes contained in the data.
Multi-Scale TAR

Computing areas of triangles made by points on the boundary

Developed by Naif Alajlan, Ibrahim El Rube, Mohamed S. Kamel and George Freeman in August 2006
Simple Example

Original Shape

TAR
Simple Example

Original Shape

TAR
Simple Example

Original Shape

TAR
Simple Example

Original Shape

TAR
Computational Complexity for TAR

Complexity: Quadratic, in the number of points

<table>
<thead>
<tr>
<th>number of points</th>
<th>time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.13 ± 0.01</td>
</tr>
<tr>
<td>100</td>
<td>0.36 ± 0.01</td>
</tr>
<tr>
<td>200</td>
<td>1.50 ± 0.05</td>
</tr>
<tr>
<td>500</td>
<td>13.7 ± 0.4</td>
</tr>
</tbody>
</table>
Computational Complexity for TAR

Optimizing by vectorization of iterative steps allowed parallelization of computation, and thus increases in speed were achieved.

<table>
<thead>
<tr>
<th>number of points</th>
<th>time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.0075 ± 0.0001</td>
</tr>
<tr>
<td>100</td>
<td>0.015 ± 0.002</td>
</tr>
<tr>
<td>200</td>
<td>0.030 ± 0.003</td>
</tr>
<tr>
<td>500</td>
<td>0.097 ± 0.01</td>
</tr>
</tbody>
</table>
## Comparison

<table>
<thead>
<tr>
<th>number of points</th>
<th>original time (s)</th>
<th>optimized time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.13 ± 0.01</td>
<td>0.0075 ± 0.0001</td>
</tr>
<tr>
<td>100</td>
<td>0.36 ± 0.01</td>
<td>0.015 ± 0.002</td>
</tr>
<tr>
<td>200</td>
<td>1.50 ± 0.05</td>
<td>0.030 ± 0.003</td>
</tr>
<tr>
<td>500</td>
<td>13.7 ± 0.4</td>
<td>0.097 ± 0.01</td>
</tr>
</tbody>
</table>
Given Data Set

- Images of concentrations taken at regular intervals
- Interesting shapes form as time progresses
- Q: Can the shapes be clustered into specific groups?
Results
Clustering $k=2$
Clustering k=3
Dendrogram
Shape 0
TAR for Shape 0
TAR for Shape 1
TAR for Shape 2
Shape 3
TAR for Shape 2
Shape 0 and 1 Similarities
Shape 1 and 2 Similarities
What about the inside?

For each of the triangles in the multi-scale TAR, compute the sum of the intensities inside that triangle, and store those intensities.
Intensity-based clustering
Current Problems and Further Work

Downsampling is necessary to use the multi-scale TAR in a decent amount of time.

Downsampling causes the intensities within a triangle area to vary even at the same scale.

The intensities do not strongly influence the results of clustering.