Shape Contexts

Newton Petersen

4/25/2008

"Shape Matching and Object Recognition Using Shape Contexts",
Belongie et al. PAMI April 2002
Agenda

- Study Matlab code for computing shape context
- Look at limitations of descriptor
- Explore effect of noise
- Explore rotation invariance
- Explore effect of locality
- Explore Thin Plate Spline

"Shape Matching and Object Recognition Using Shape Contexts", Belongie et al. PAMI April 2002
Problem: How can we tell these are same shape?
Shape Context – Step 1 - Distance

Coordinates on shape:
(1) 0.2000 0.5000
(2) 0.4000 0.5000
(3) 0.3000 0.4000
(4) 0.1500 0.3000
(5) 0.3000 0.2000
(6) 0.4500 0.3000

Compute Euclidean distance from each point to all others:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>0.2000</th>
<th>0.1414</th>
<th>0.2062</th>
<th>0.3162</th>
<th>0.3202</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.2000</td>
<td>0.1414</td>
<td>0.2062</td>
<td>0.3162</td>
<td>0.3202</td>
</tr>
<tr>
<td>0.2000</td>
<td>0.2000</td>
<td>0</td>
<td>0.1414</td>
<td>0.3202</td>
<td>0.3162</td>
<td>0.2062</td>
</tr>
<tr>
<td>0.1414</td>
<td>0.1414</td>
<td>0</td>
<td>0.1803</td>
<td>0.2000</td>
<td>0.1803</td>
<td></td>
</tr>
<tr>
<td>0.2062</td>
<td>0.3202</td>
<td>0.1803</td>
<td>0</td>
<td>0.1803</td>
<td>0.3000</td>
<td></td>
</tr>
<tr>
<td>0.3162</td>
<td>0.3162</td>
<td>0.2000</td>
<td>0.1803</td>
<td>0</td>
<td>0.1803</td>
<td></td>
</tr>
<tr>
<td>0.3202</td>
<td>0.2062</td>
<td>0.1803</td>
<td>0.3000</td>
<td>0.1803</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Then normalize by mean distance…
Shape Context – Step 2 – Bin Distances

Normalized distances between each point:

0 1.0623 0.7511 1.0949 1.6796 1.7004
1.0623 0 0.7511 1.7004 1.6796 1.0949
0.7511 0.7511 0 0.9575 1.0623 0.9575
1.0949 1.7004 0.9575 0 0.9575 1.5934
1.6796 1.6796 1.0623 0.9575 0 0.9575
1.7004 1.0949 0.9575 1.5934 0.9575 0

Create log distance scale for normalized distances (closer = more discriminate):

0.1250 0.2500 0.5000 1.0000 2.0000

Create distance histogram: Iterate for each scale incrementing bins when dist <

1 0 0 0 0 0
0 1 0 0 0 0
0 0 1 0 0 0
0 0 0 1 0 0
0 0 0 0 1 0
0 0 0 0 0 1
5 1 2 1 1 1
1 5 2 1 1 1
1 1 2 5 2 1
1 1 1 2 5 2
1 1 2 1 2 5

Bottom Line: Bins with higher numbers describe points closer together
Shape Context – Step 3 - Angles

Coordinates on shape:
(1) 0.2000 0.5000
(2) 0.4000 0.5000
(3) 0.3000 0.4000
(4) 0.1500 0.3000
(5) 0.3000 0.2000
(6) 0.4500 0.3000

Compute angle between all points (0 to 2π):

<table>
<thead>
<tr>
<th>Angle</th>
<th>0</th>
<th>3.1416</th>
<th>2.3562</th>
<th>1.3258</th>
<th>1.8925</th>
<th>2.4669</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>5.4978</td>
<td>4.4674</td>
<td>5.0341</td>
<td>5.6084</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.1416</td>
<td>0</td>
<td>3.9270</td>
<td>3.8163</td>
<td>4.3906</td>
<td>4.9574</td>
</tr>
<tr>
<td></td>
<td>2.3562</td>
<td>0.7854</td>
<td>0</td>
<td>3.7296</td>
<td>4.7124</td>
<td>5.6952</td>
</tr>
<tr>
<td></td>
<td>1.3258</td>
<td>0.6747</td>
<td>0.5880</td>
<td>0</td>
<td>5.6952</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1.8925</td>
<td>1.2490</td>
<td>1.5708</td>
<td>2.5536</td>
<td>0</td>
<td>0.5880</td>
</tr>
<tr>
<td></td>
<td>2.4669</td>
<td>1.8158</td>
<td>2.5536</td>
<td>3.1416</td>
<td>3.7296</td>
<td>0</td>
</tr>
</tbody>
</table>
Shape Context – Step 4 – Quantize Angles

Binning angles is slightly different than distance:

\[
\begin{array}{cccccc}
0 & 0 & 5.4978 & 4.4674 & 5.0341 & 5.6084 \\
3.1416 & 0 & 3.9270 & 3.8163 & 4.3906 & 4.9574 \\
2.3562 & 0.7854 & 0 & 3.7296 & 4.7124 & 5.6952 \\
1.3258 & 0.6747 & 0.5880 & 0 & 5.6952 & 0 \\
1.8925 & 1.2490 & 1.5708 & 2.5536 & 0 & 0.5880 \\
2.4669 & 1.8158 & 2.5536 & 3.1416 & 3.7296 & 0 \\
\end{array}
\]

Simple Quantization:

\[
\text{theta\_array\_q = 1+floor(theta\_array\_2/(2*pi/nbins\_theta))}
\]

\[
\begin{array}{cccccc}
1 & 1 & 6 & 5 & 5 & 6 \\
4 & 1 & 4 & 4 & 5 & 5 \\
3 & 1 & 1 & 4 & 5 & 6 \\
2 & 1 & 1 & 1 & 6 & 1 \\
2 & 2 & 2 & 3 & 1 & 1 \\
3 & 2 & 3 & 4 & 4 & 1 \\
\end{array}
\]
Shape Context – Step 5 – Combine

- R and theta numbers are combined to one descriptor (slightly tricky Matlab code)
- Captures number of points in each R, theta bin
- Effectively turned N points into $N \times \text{NumRadialBins} \times \text{NumThetaBins} = \text{Rich Descriptor}$

```
1 0 0 0 2 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
```

... for each point

... relative to each point and not a global origin
Matching – Cost Matrix

- Calculate ‘cost’ of matching each point to every other point
- Cost of matching point i to point j = Chi-squared similarity between row i and row j in shape context descriptor

\[
C_{ij} \equiv C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}
\]

- All histogram bins in one row
- Bin values normalized by total number of points
Matching – Additional Cost Terms

- Easy to add in other terms
- For ‘real’ images, possible to add in other measures of difference between point i and j
  - Surrounding Color Difference
  - Surrounding Texture Difference
  - Surrounding Brightness Difference
  - Tangent Angle Difference

"Shape Matching and Object Recognition Using Shape Contexts", Belongie et al. PAMI April 2002
Matching

- Find pairing of points that leads to least total cost
- Hungarian Method
  - $O(n^3)$

Cost of matching point 1 of shape 1 to point 2 of shape 2

\[
H(\pi) = \sum_i C(p_i, q_{\pi(i)})
\]
So what Happened Here?

- Inexact rotation applied
Much better…
Systematic Rotation Experiment

- Rotate through $2\pi/40$ increments
- Quite sensitive to rotation
- Even if ‘shape context distance’ low
Providing Rotation Invariance

- Relation between tangent angles stays the same as points rotate

"Shape Matching and Object Recognition Using Shape Contexts", Belongie et al. PAMI April 2002
Rotation Invariance

- Use tangent angle as positive x axis for each point (as suggested in paper)
Rotation Invariance

- Do you really want 6 and 9 matched?
- Depends on the shape…
Locality issues - Matching Example

What happened here?

"Shape Matching and Object Recognition Using Shape Contexts",
Belongie et al. PAMI April 2002
What could produce ‘incorrect’ descriptors?

- As we just saw,
  - Rotation that puts points in different relative bins
  - Different numbers of points in different regions of shapes
- Any important distinction that ends up in the same bin is effectively lost
  - Chance of happening increases with distance
- Conversely any nearby feature relation that is unimportant is granted a distinction in the descriptor

"Shape Matching and Object Recognition Using Shape Contexts",
Belongie et al. PAMI April 2002
More realistic locality example

- Smaller radius creates more outliers that can match with points far away if nothing available locally
Effects of noise

- Not really all that good at dealing with noise (at least not this much noise)
Thin Plate Spline Warping

\[ I_f = \int \int_{\mathbb{R}^2} \left( \frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 f}{\partial y^2} \right)^2 \, dx \, dy \]

- Meant to model transformations that happen when bending metal
- Picks a warp that minimizes the ‘bending energy’ above and minimizes shape distance

"Shape Matching and Object Recognition Using Shape Contexts", Belongie et al. PAMI April 2002
Bend a fish?

"Shape Matching and Object Recognition Using Shape Contexts", Belongie et al. PAMI April 2002
TPS

Added Noise Points

- Helps absorb small local differences by having smoothing effect (regularization parameter)
- Helps smooth edge sampling jitter
- Provides small degree of rotation invariance
- Helps provide some immunity to noise by bunching noisy points together

"Shape Matching and Object Recognition Using Shape Contexts",
Belongie et al. PAMI April 2002
Conclusion

- Shape context => binning of spatial relationships between points
- Good for ‘clean’ shapes
  - Examples from paper => handwriting, trademarks
- Struggles with clutter noise
  - Thin Plate Spline helps quite a bit
Discussion

- How does this compare to other descriptors?
- What would work better with Maysam’s viruses?
- Any ideas for making descriptor know what geometrical relationships are most important? (like active appearance models)
- Any ideas for improving runtime