Poselets: Body Part Detectors Trained Using 3D Human Pose Annotations

*Lubomir Bourdev and Jitendra Malik*

*Experiments*

*Presented by Randall Smith*
Outline

• Introduction
• Dataset
• Overview
• Annotations
• Distance Function
• Segmentation
• Experiments
• Conclusion
Dataset

- Humans in 3D (H3D)
- 2480 annotations
  - (1500 train / 500 test / 240 validate)
- Java3D annotation tool
Overview: Training

Random Seed Patches → Select Nearby Patches → Train SVMs → Prune

Residual Error:
- Random Seed Patches: 0.15
- Select Nearby Patches: 0.20
- Train SVMs: 0.10
- Prune: 0.85
- Residual Error: 0.15, 0.35
Overview: Detection

- Run All Poselets
- Mean Shift Cluster
- Compute Activations
- Max Margin Hough
Annotations

• Bounding box placed over annotated figure
Annotations

• Live Demo
Annotations

- Live Demo
Annotations : Skeleton

- Annotated skeleton
Annotatons: Keypoints

- 20 manually annotated keypoints
- 15 manually annotated segments
Annotations : Query

Mean

Examples

Distance

• Query at green box
Annotations: Query

Mean

Examples

Distance

<table>
<thead>
<tr>
<th></th>
<th>0.0017</th>
<th>0.0020</th>
<th>0.0020</th>
<th>0.0028</th>
<th>0.0031</th>
<th>0.0032</th>
<th>0.0035</th>
<th>0.0041</th>
<th>0.0043</th>
</tr>
</thead>
</table>

Friday, November 2, 12
Annotations : Query

Mean

Examples

Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>0.0019</th>
<th>0.0079</th>
<th>0.0093</th>
<th>0.0096</th>
<th>0.0104</th>
<th>0.0134</th>
<th>0.0139</th>
<th>0.0154</th>
<th>0.0169</th>
</tr>
</thead>
</table>

Annotations: Query

Mean

Examples

Distance

<table>
<thead>
<tr>
<th></th>
<th>0.0067</th>
<th>0.0086</th>
<th>0.0167</th>
<th>0.0176</th>
<th>0.0178</th>
<th>0.0180</th>
<th>0.0183</th>
<th>0.0186</th>
<th>0.0198</th>
</tr>
</thead>
</table>

Friday, November 2, 12
Distance Function

- Paper: computes a weighted sum of Euclidean distances with additive penalty.
- Implementation: Procrustes distance plus penalty.
- What is the Procrustes distance?
Distance Function

\[ D_{\text{proc}}(x_1, x_2) = \min_{s, R, t} \| x_1 - (sRx_2 + t) \| \]

- Scale so that RMS is 1.0, translate to origin, and solve for rotation matrix \( R \).
- Non visible key points ignored
Distance Function

\[ D(x_s, x_r) = D_{\text{proc}}(x_s, s_r) + \text{Penalty} \]

- Need to compute linear least squares / SVD to solve.
- Is this very expensive?
Distance Function

- Live Demo: 2D Toy Example
Segments: UpperClothes
Segments : LowerClothes
Segments : Faces
Detection

• A simple, occlusion free test
## Detection

<table>
<thead>
<tr>
<th>Task</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features</td>
<td>0.69s</td>
</tr>
<tr>
<td>Detect Poselets</td>
<td>0.56s</td>
</tr>
<tr>
<td>Score</td>
<td>0.82s</td>
</tr>
<tr>
<td>Cluster</td>
<td>0.61s</td>
</tr>
<tr>
<td>Localize</td>
<td>0.11s</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2.49s</strong></td>
</tr>
</tbody>
</table>
Detection : Example

- Score: 14.10. How did the clusters vote?
Detection: Votes

- Inspect top hits
- Inspect bottom hits
Detection : Best Votes
Detection : Best Votes
Detection: Best Votes
Detection : Best Votes
Detection: Worst Votes
Detection : Worst Votes
Detection: Worst Votes
Detection : Worst Votes
Detection : Tests

- Some samples from PASCAL VOC2007
- With varying degrees of occlusion
- Comparison with *Discriminatively Trained Deformable Part Models* (DPM)
- Some pictures taken from my iPhone 4S
- Increasingly difficult in terms of occlusion
Detection : Tests

• PASCAL VOC2007
Detection Tests

• Some more difficult occlusion cases
Detection: Comparison

- Scores: 11.50 and 2.03.
- DPM failed.
Detection: Comparison

- Score: 5.21. 31 poselet clusters contributed.
- DPM HOG parts and bounding box shown.
Detection: Comparison

- Scores 12.48, 9.67, and 5.40.
- DPM failed.
Detection: Comparison

- Score 12.01.
- HOG parts and bounding box shown.
Detection:

Comparison

• Score: 5.21.

• HOG parts and bounding box shown.
Detection: Comparison

- Both fail.
Detection: Comparison

• When both succeed, DPM seems to get better bounding boxes.

• The poselet algorithm always tries to get the best bounding box it can.

• DPM has no way of degrading gracefully.
Detection: Occlusion

- Score: **22.4**. 54 poselet clusters contributed.
Detection : Occlusion

- Score: 0.29. 3 poselet clusters contributed.
Detection: Occlusion

- Score: 0.38. 2 poselet clusters contributed.
Score: 0.27. 1 poselet cluster contributed.
Detection : Occlusion

- Score: 0.21. 2 poselet clusters contributed.
Detection: Occlusion

- DPM fails on all of these.
- Poselets do pretty poorly, but it still computes a bounding box.
- Poselets have the chance of getting it right.
Conclusions

• Poselets are intuitive to find in an image.
• If a body part is exposed, a poselet might match it.
• Poselet ranking and scoring can be understood in an intuitive way.
• Can handle some occlusion
• Will always try to compute a bounding box.
Conclusions

• Sometimes poselet activations can be misleading.

• Sometimes, some poselets should have higher scores than others.

• This is sort of like getting the right answer for the wrong reasons.

• The dataset is very labor intensive.