Overview

Objects within general categories vary significantly from one instance to the next. We can group the numerous factors contributing to these changes into two broad categories:

- **Pose**
- **Appearance**

Mixture models treat each factor independently which requires lots of data to train, and may lead to overfitting.

**Contribution**: We propose using one appearance model that deforms according to pose.

**DPMs**

- Each part placement gives a new template
- **Exponentially many** templates that share parameters

- Can generalize to unseen poses
- **May not capture** all relevant deformations
- **Are mostly implausible**:

Can a small data-driven deformation dictionary capture most plausible deformations?

Our model

Each component has:

1. A coarse template \(w_c\)
2. A single fine template \(w_f\)
3. A dictionary of deformations \(D\)

Deformed templates \(Dw_f\)

Generating the dictionary

- Whitens all exemplar HOGs
- Estimate deformation from mean to exemplars
- Use block matching with small HOG patches
- Do PCA and project to top principal components
- Cluster using k-means

Deformed appearance templates:

Results

**Example detections**:

The two detections in each pair share the template but have different deformations.

**Parameter sharing**:

- \(n\)-comp: Each component has just a low-resolution root filter.
- \(nm\)-comp: A standard mixture model with as many root filters as we have high-resolution templates.
- \(n\)-comp x \(m\)-fine: Each component has a root filter and \(m\) high-res templates trained without parameter sharing.

Full experiments on PASCAL VOC 2012:

**Precision recall curves** for car (left) and person (right). We do well in the high precision regime.

**Sharing deformation dictionaries**:

<table>
<thead>
<tr>
<th>Method</th>
<th>Bike (100 samples)</th>
<th>Bike (full data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n-comp</td>
<td>bike (100 samples)</td>
<td>bike (full data)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>n-comp</th>
<th>nm-comp</th>
<th>n-comp x m-fine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bicycle</td>
<td>46.5</td>
<td>46.9</td>
<td>47</td>
</tr>
<tr>
<td>Monitor</td>
<td>31.1</td>
<td>30.8</td>
<td>31.6</td>
</tr>
<tr>
<td>Cow</td>
<td>13.4</td>
<td>12.3</td>
<td>10.3</td>
</tr>
<tr>
<td>Sheep</td>
<td>27</td>
<td>27</td>
<td>26.9</td>
</tr>
</tbody>
</table>

**Conclusion**

We have proposed using a discrete set of deformations. However, we can also search for the optimal deformation within the space defined by our set of 5 PCA bases. Using a greedy search technique, we were able to obtain similar results to that of our discrete model. While the discrete approach is more computationally efficient, it may prove beneficial to search in a continuous space of deformations for some object categories.