A2 (Inpainting) and Pictorial Structure


Department of Computer Science

Micha Livne
Assignment 2

Tips:

- Don’t waste time on setting the required libraries on your own machine - use CDF!
  
  `ssh -Y <CDF User Name>@cdf.toronto.edu`

- Again, code must work on CDF, so make sure it does.

- Start early.

Any questions?
Pictorial Structure - Overview

- A part-based modeling and recognition of objects.
- A seminal paper of 2D model recognition.

Figure 8: Input image, binary image obtained by background subtraction, and matching result superimposed on both images.

Articulated bodies are imprecise rather than being accurate generative models.

6.1 Parts

For simplicity, we assume that the image of an object is generated by a scaled orthographic projection, so that parallel features in the model remain parallel in the image. For images of human forms this is generally a reasonable assumption. We further assume that the scale factor of the projection is known. We can easily add an extra parameter to our search space in order to relax this latter assumption.

Suppose that objects are composed of a number of rigid parts, connected by flexible joints. If a rigid part is more or less cylindrical, its projection can be approximated by a rectangle. The width of the rectangle comes from the diameter of the cylinder and is fixed, while the length of the rectangle depends on the length of the cylinder but can vary due to foreshortening. We model the projection of a part as a rectangle parameterized by \((x, y, s, \theta)\). The center of the rectangle is given in image coordinates \((x, y)\), the length is defined by the amount of foreshortening \(s \in [0, 1]\), and the orientation is given by \(\theta\). So we have a four-dimensional pose space for each part.
Formulation

$L^* = \arg\min_L \left( \sum_{i=1}^{n} m_i(l_i) + \sum_{(v_i, v_j) \in E} d_{ij}(l_i, l_j) \right)$
Mismatch Potential $m_i$

- $m_i(l_i)$ - the mismatch of part $i$ in position $l_i$ given an image.
- Felzenswab used iconic representation - response of Gaussian derivative filters of different orders, orientations and scales.

$\begin{align*}
    p(l|l_i, u_i) &\propto \mathcal{N}(\alpha(l_i), \mu_i, \Sigma_i) \\
    \text{What other representation of parts can we use? (hint: from lecture)}
\end{align*}$
Mismatch Potential $m_i$

- $m_i(l_i)$ - the mismatch of part $i$ in position $l_i$ given an image.
- Felzenswab used iconic representation - response of Gaussian derivative filters of different orders, orientations and scales.

$p(l|l_i, u_i) \propto \mathcal{N}(\alpha(l_i), \mu_i, \Sigma_i)$

What other representation of parts can we use? (hint: from lecture)

- HoG.
Mismatch Potential $m_i$

- $m_i(l_i)$ - the mismatch of part $i$ in position $l_i$ given an image.
- Felzenswab used iconic representation - response of Gaussian derivative filters of different orders, orientations and scales.

![Gaussian derivative basis functions](image)

- $p(l | l_i, u_i) \propto \mathcal{N} (\alpha(l_i), \mu_i, \Sigma_i)$
- What other representation of parts can we use? (hint: from lecture)
  - HoG.
  - Also SIFT, SURF, etc.
Deformation Potential \( d_{ij} (l_i, l_j) \)

- \( d_{ij} (l_i, l_j) \) - the deformation distance between current part position and model part position.
- Felzenswab wants to allow deformation of the model by using a Gaussian distribution over model position.
- \( p(l_i, l_j | c_{ij}) = \mathcal{N}(l_i - l_j, s_{ij}, \Sigma_{ij}) \)
Matching a pictorial structure model to an image does not involve making any initial decisions about locations of individual parts, but rather an overall decision is made based on both the part match costs and the deformation costs together.
References