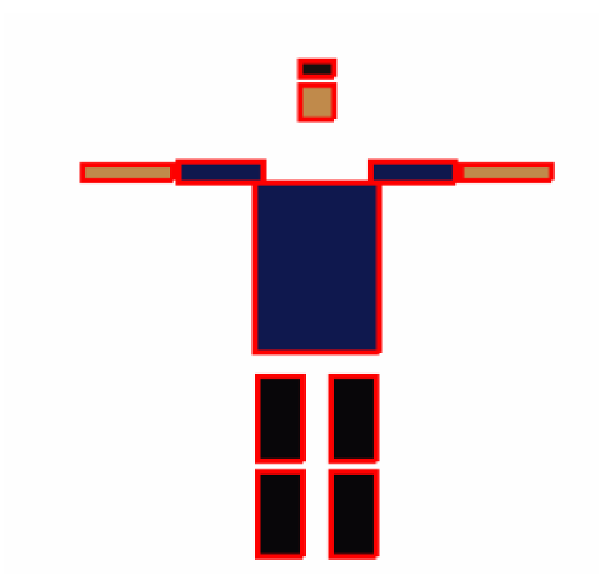
Clustering appearance and shape by learning jigsaws

A. Kannan, J. Winn, and C. Rother

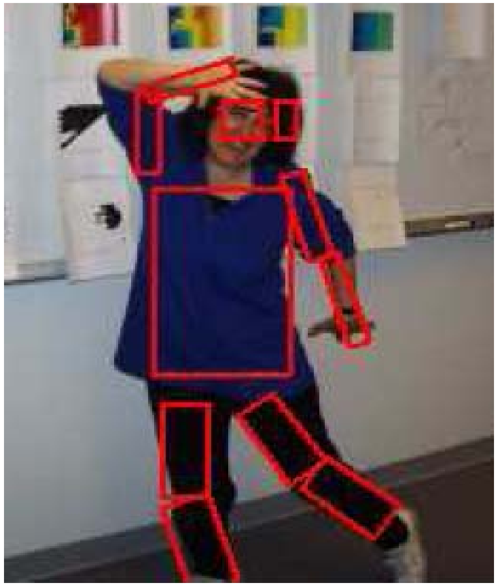
Presented by Michael Ryoo

Motivation

- Goal:
  - Unsupervised learning of appearance of 'parts'.
  - Part-based model



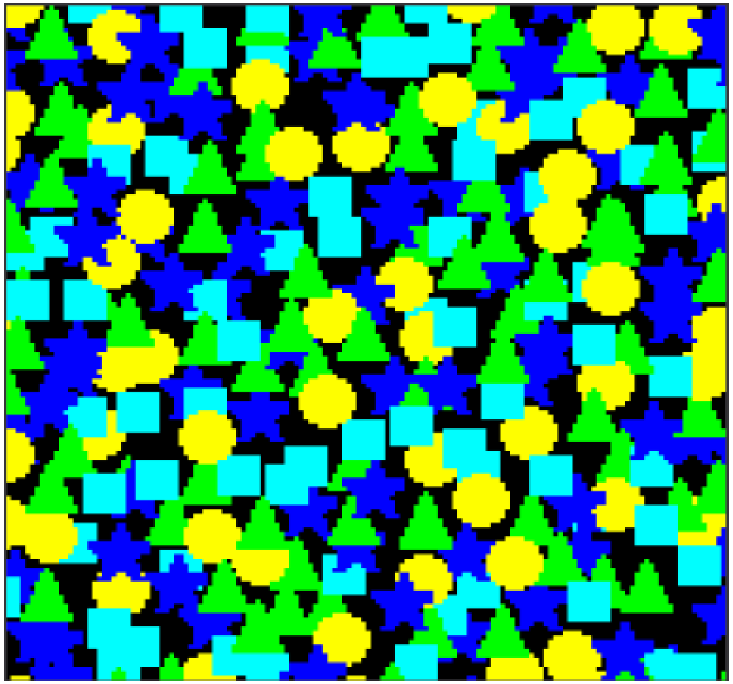
We want to learn this.

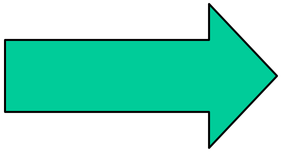


For this (recognition).

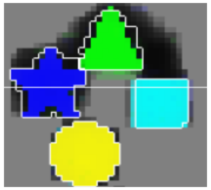
Approach

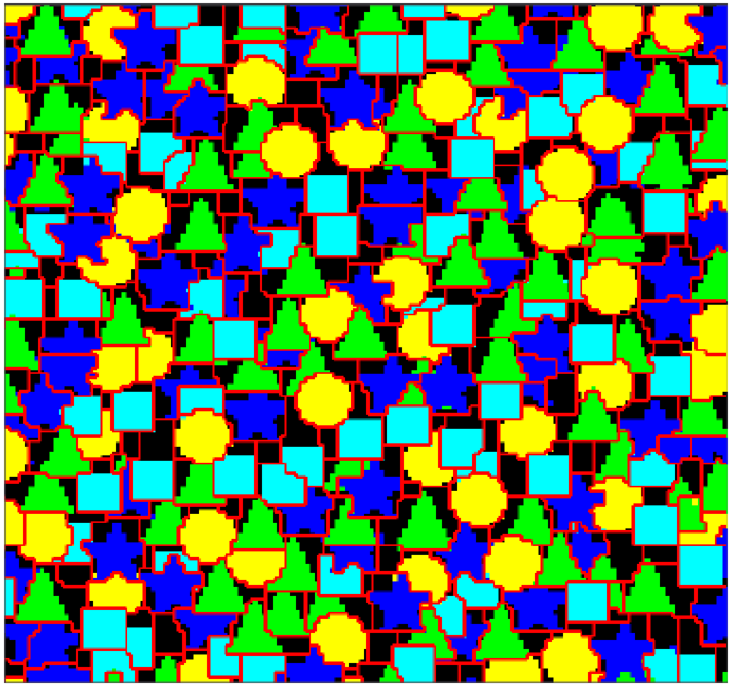
- From an image, extract a jigsaw that enables the reconstruction of the image.





mapping using

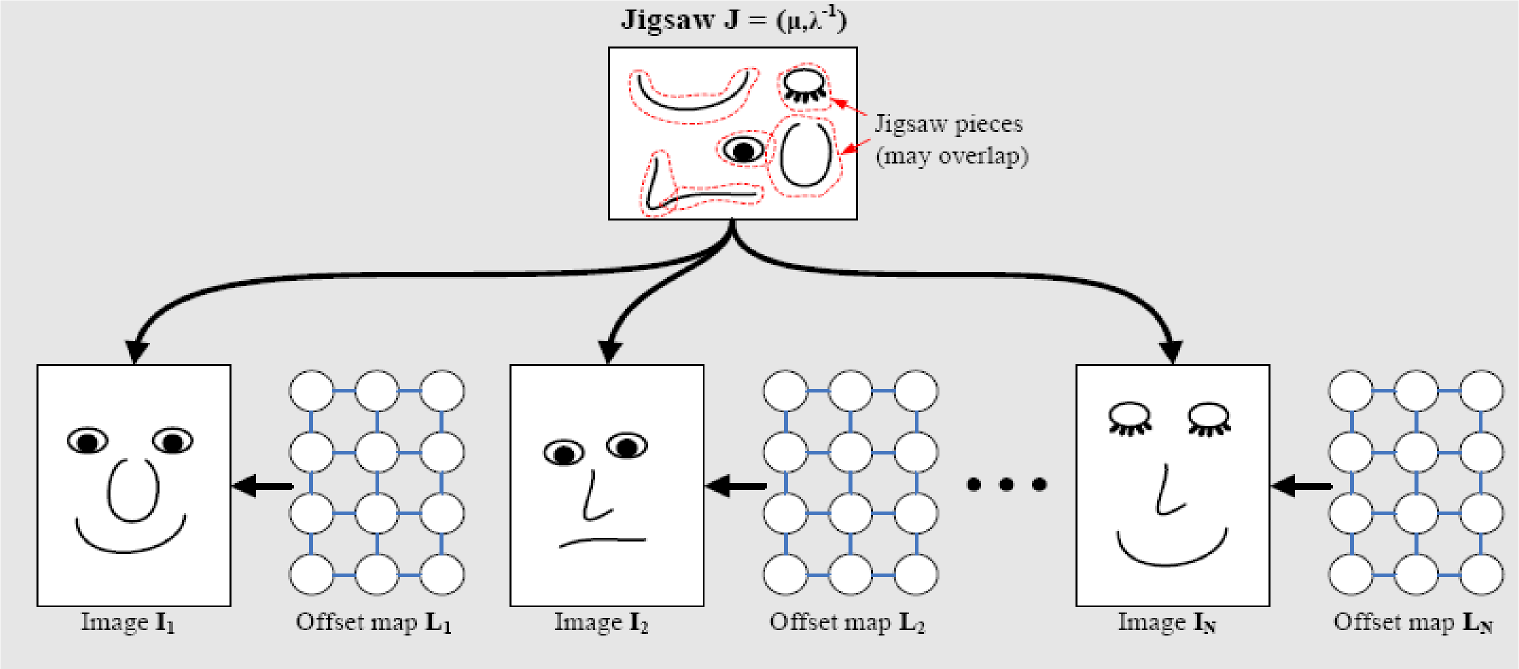




Method

- Mapping.

$$P(\mathbf{I} \mid \mathbf{J}, \mathbf{L}) = \prod_i \mathcal{N}(I(i); \mu(i - \mathbf{l}_i), \lambda(i - \mathbf{l}_i)^{-1})$$



Learning

- Find P(J) that maximizes

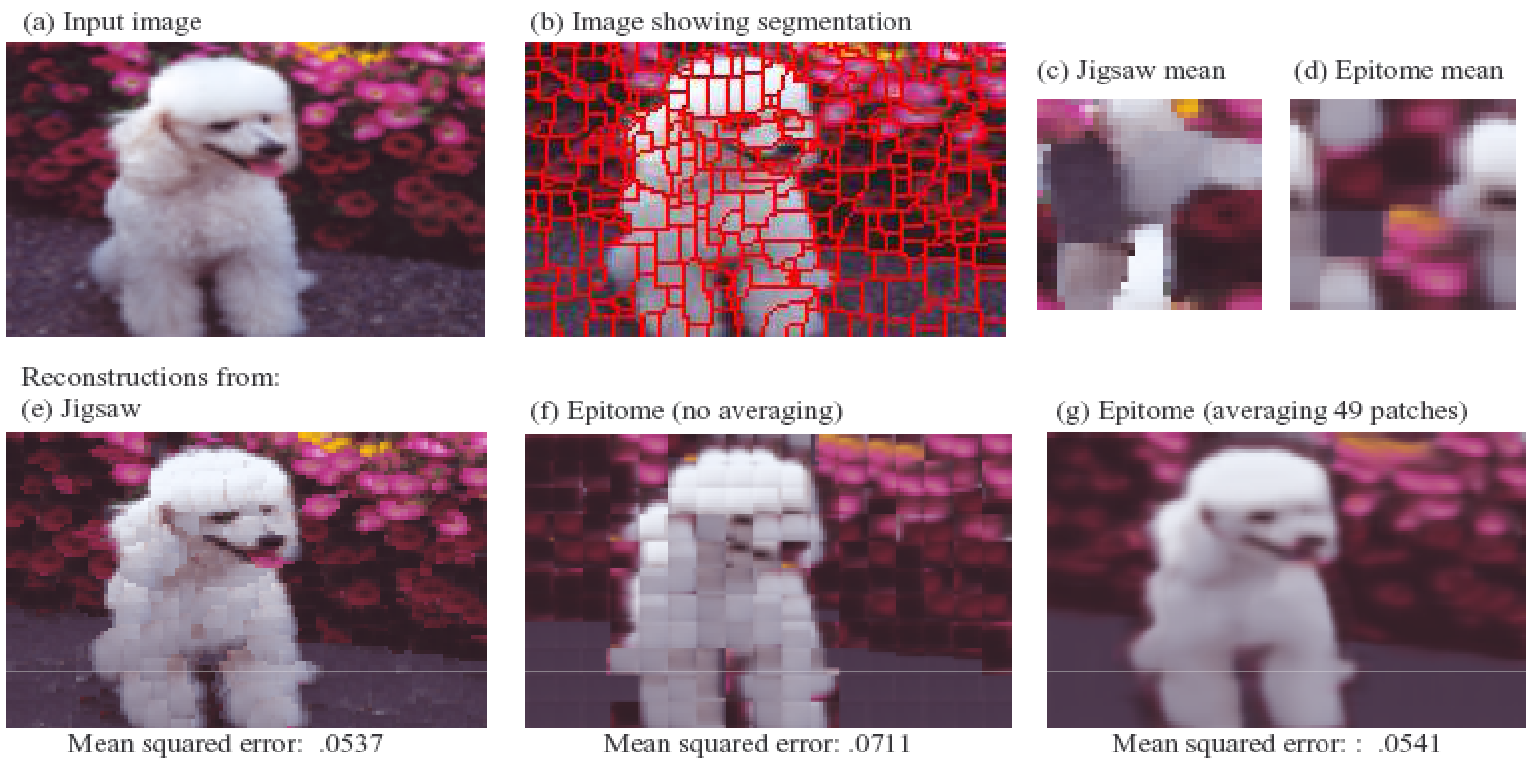
$$P\left(\mathbf{J}, \{\mathbf{I}, \mathbf{L}\}_1^N\right) = P(\mathbf{J}) \prod_{n=1}^{N} P(\mathbf{I}_n|\mathbf{J}, \mathbf{L}_n) P(\mathbf{L}).$$

  – Alpha extension graph-cut algorithm

$$\mu^\star = \frac{\beta\mu_0 + \sum_{\mathbf{x}\in X(\mathbf{z})} I(\mathbf{x})}{\beta + |X(\mathbf{z})|}$$

$$\lambda^{-1\star} = \frac{b + \beta\mu_0^2 - (\beta + |X(\mathbf{z})|)(\mu^\star)^2 + \sum_{\mathbf{x}\in X(\mathbf{z})} I(\mathbf{x})^2}{a + |X(\mathbf{z})|}$$

Discussion


(a) Input image
(b) Image showing segmentation
(c) Jigsaw mean
(d) Epitome mean
Reconstructions from:
(e) Jigsaw
(f) Epitome (no averaging)
(g) Epitome (averaging 49 patches)
Mean squared error: .0537
Mean squared error: .0711
Mean squared error: : .0541