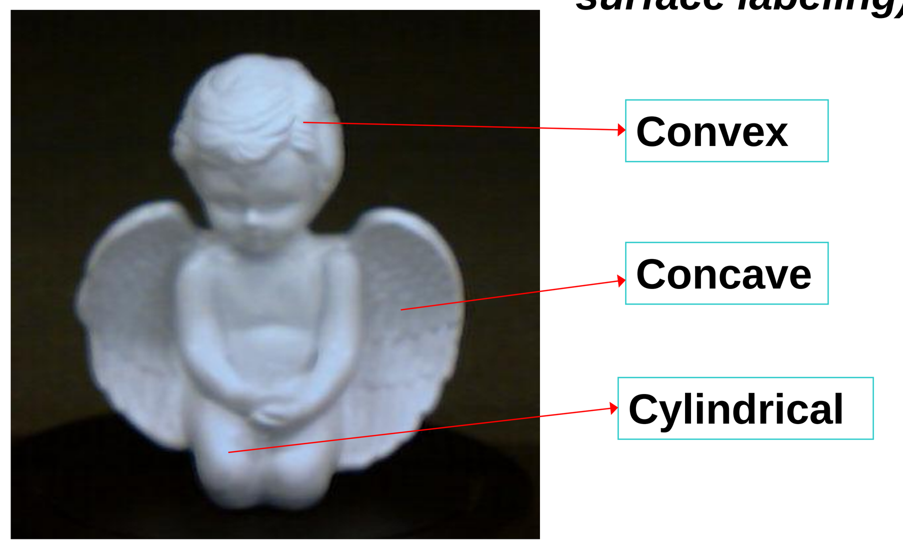
Qualitative 3D Surface Reconstruction From Images

Alex Levinshtein[1], Cristian Sminchisescu[2] and Sven Dickinson[1], [1]*University of Toronto*, [2]*University of Bonn*

Qualitative 3D from an image

Goal: Qualitative 3D from a single image

Input (image) — *Output (piecewise surface labeling)*



Approach

Learn a mapping from images to qualitative surface shape

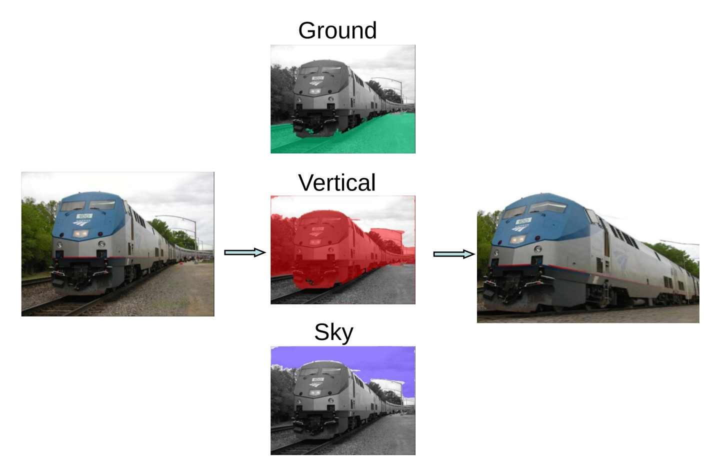
Motivation

- A mid-level object description
- Generic: coarser than exact reconstruction, not object specific, however
- Useful for indexing, even if sparse
- Useful for categorical recognition

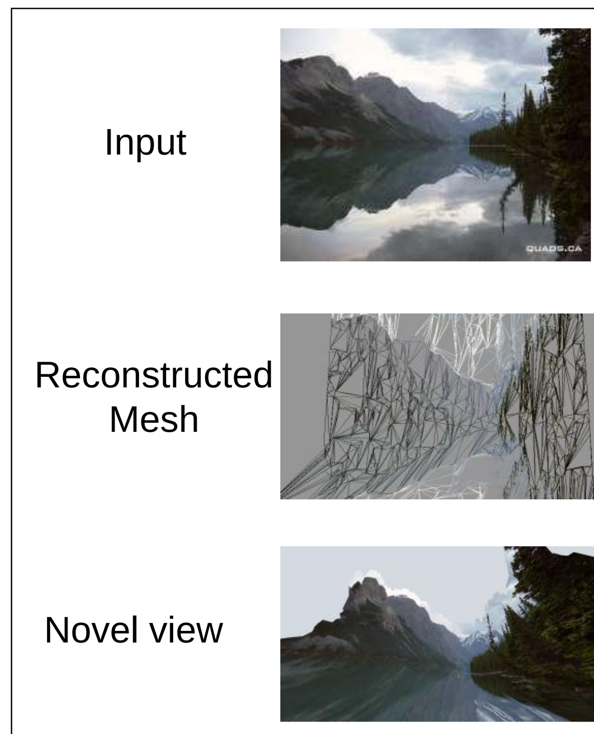
Execution

Train a Conditional Random Field (CRF) to recover piecewise labeling from image superpixel graph, based on (2d,3d) training set

Related Work

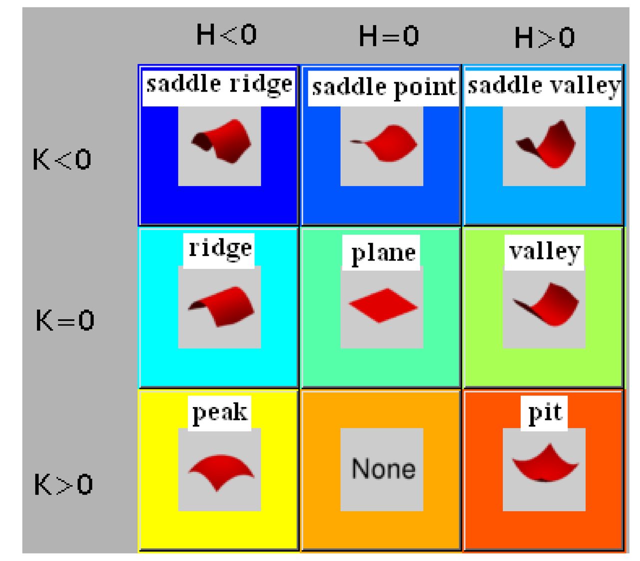


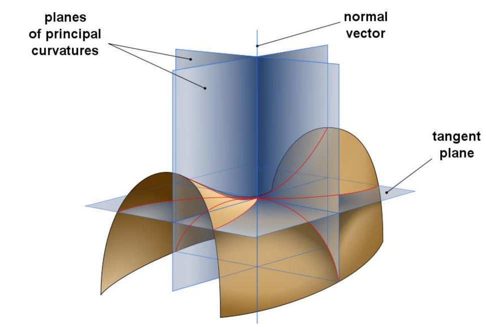
Hoeim et al., SIGGRAPH 2005



Saxena et al., 3DRR-07

Qualitative Surface Labels



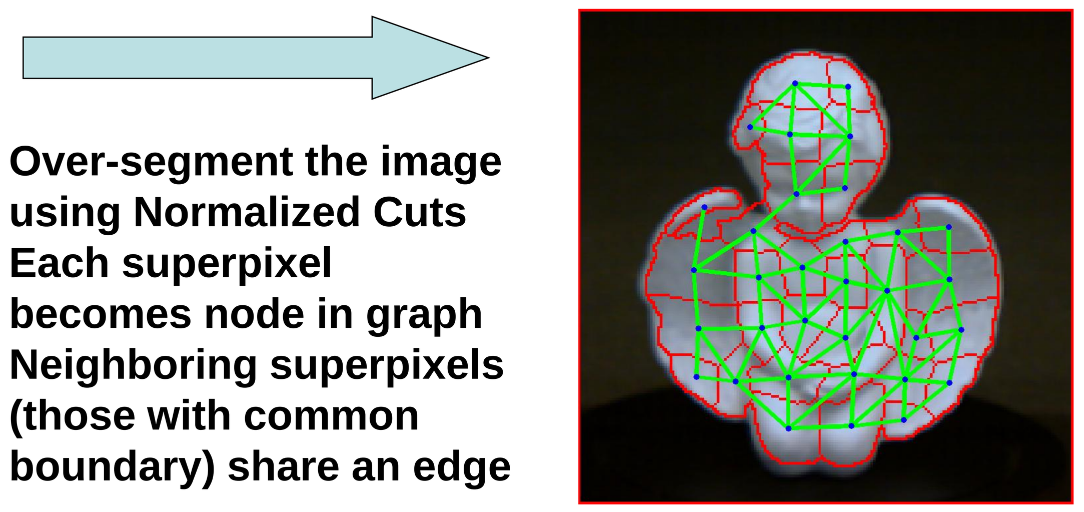


- Inspired by Besl and Jain's (PAMI 1988) representation previously used for segmentation of 3d range data

k_1 and k_2 are the surface (local) principal curvatures:

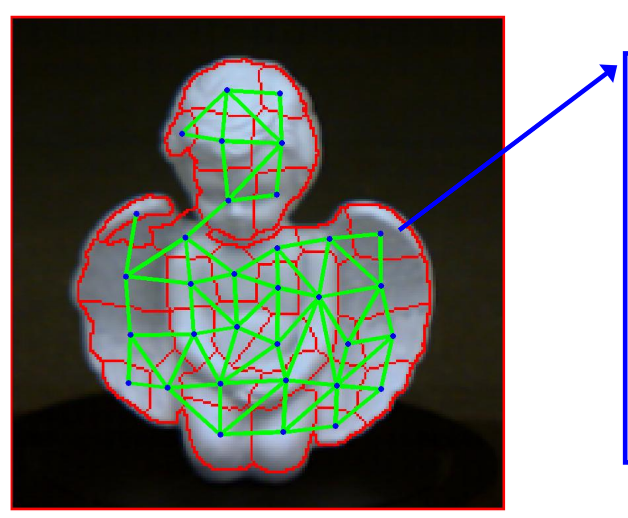
$$H = \frac{k_1 + k_2}{2}, \quad K = k_1 \cdot k_2$$

1. Superpixel graph



- **Over-segment the image using Normalized Cuts**
- **Each superpixel becomes node in graph**
- **Neighboring superpixels (those with common boundary) share an edge**

2. Feature Design



36 dimensional vector (F_i) encodes both contour and internal superpixel appearance information

- **Kernelize by computing covariance features for each superpixel**
  - ➢**Estimate** $F_i \cdot F_i^T$
  - ➢**Vectorize lower triangular the matrix as feature set**

3. CRF formulation

- X **– set of superpixel labels**
- x_i **– label for superpixel** i
- F_i **– feature vector for superpixel i (after the kernelization)**
- **K – number of possible labels (8 for HK encoding)**
- **S – number of superpixels**

$$p(X \mid F, w, v) = \frac{1}{Z(w,v)} \exp\left\{ \sum_i w_i^T F_i + \sum_{ij} v(i,j) \right\}$$

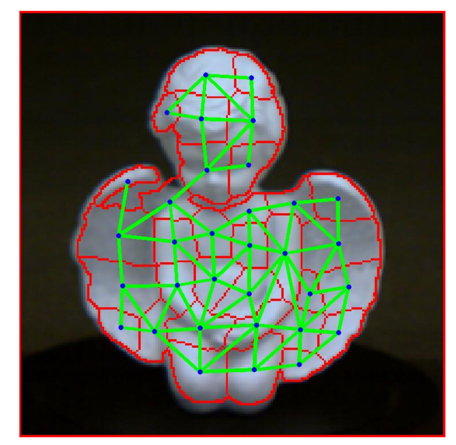
CRF parameters:

- w **– KxS matrix of coefficients**
- v **– KxK symmetric matrix of binary potentials (label compatibilities)**

Overview of the Method

1. Construct superpixel graph

- **Oversegment the image**
- **Construct a graph were segments are nodes and edges are neighborhood relations between nodes**

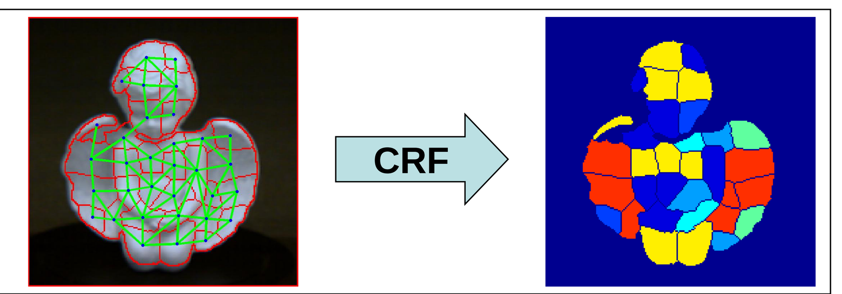


2. Extract features

- **Each superpixel** ***i*** **is associated feature vector** F_i
- **Features capture**
  - ➢**Contour strength and curvature**
  - ➢**Moments of intensity inside superpixel, local jets, etc**

3. Train CRF

- **Find the maximum likelihood parameters of the CRF**



3a. Bethe Free Energy Approximations for CRF

- Use Conditional Maximum likelihood learning
- Compute log-likelihood of the data under the model
  - ➢Use Bethe Free Energy to approximate log Z

$$\log Z = -F_\beta$$

$$F_\beta = -\sum_{ij}\sum_{x_i,x_j} b_{ij}(x_i,x_j)\log\psi_{ij}(x_i,x_j) - \sum_i\sum_{x_i} b_i(x_i)\log\psi_i(x_i)$$

$$+\sum_{ij}\sum_{x_i,x_j} b_{ij}(x_i,x_j)\log b_{ij}(x_i,x_j) - \sum_i (q_i - 1)\sum_{x_i} b_i(x_i)\log b_i(x_i)$$

Use standard optimization package to finding optimal parameters, w and v

- Compute gradients of log likelihood
  - ➢Derivative of the partition function involves intractable expectation
  - ➢Use loopy belief propagation to approximate necessary marginals

3b. CRF inference

Use Loopy Belief propagation for inference

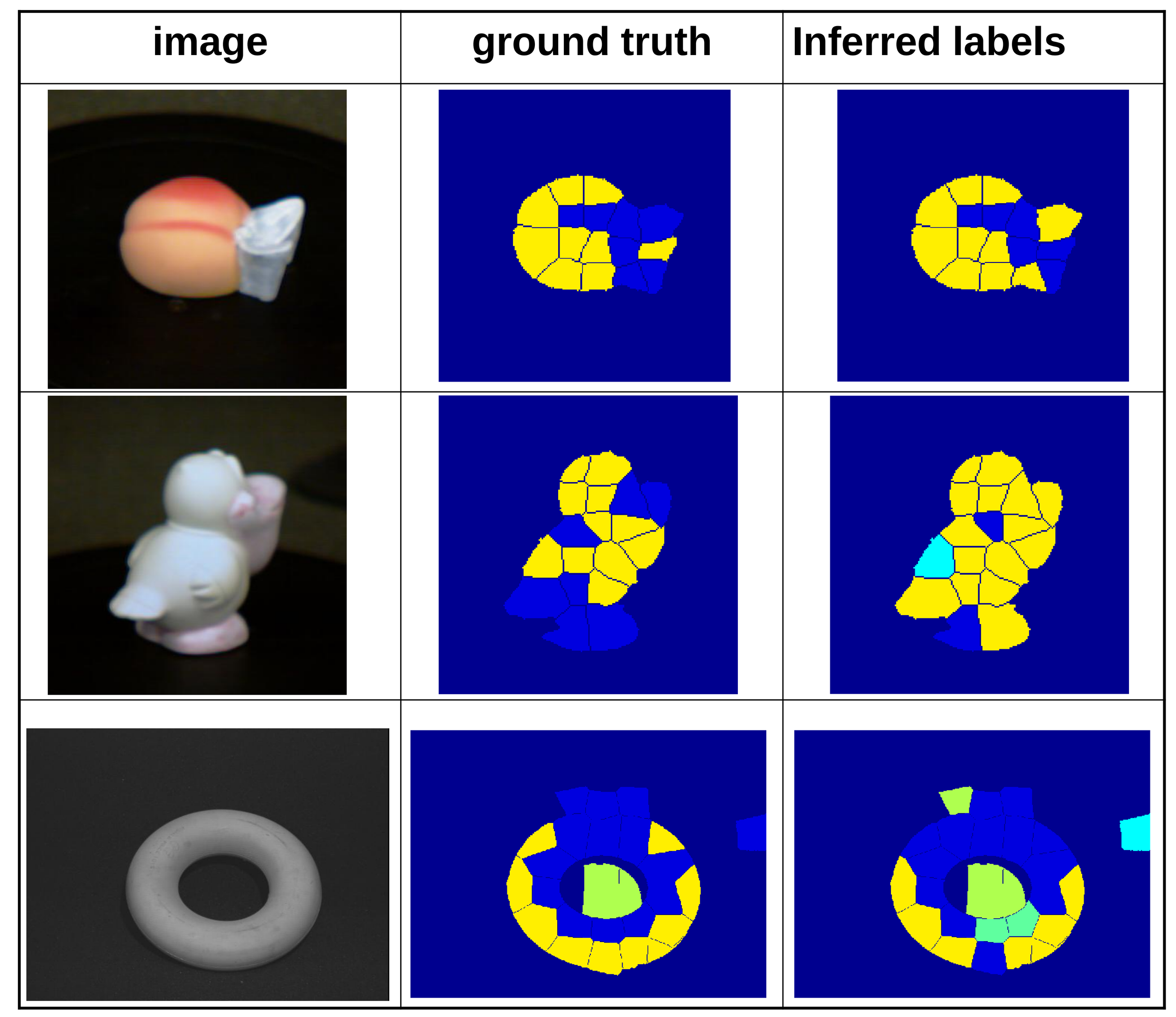
- **At every iteration update messages in parallel**
- **Use damping for stability**

$$\mu_{new} = (1-\alpha)\mu_{new} + \alpha\mu_{old}$$

4. Results

Test on two datasets (Minolta and K2T) containing pairs of optical and range images

Qualitative Results



Quantitative Results

	Train	Test
Minolta	0.6	0.48
K2T	0.48	0.32

Precision on 6-label problem (all saddles merged into one class)