Glass Object Localization by Joint Inference of Boundary and Depth

Tao Wang  Xuming He  Nick Barnes

NICTA & Australian National University

25 September 2012
Outline

Motivation: Finding Glass

- Glass surfaces and their interaction with the environment
- Cues for glass boundary detection
- Surface and object models for glass

Our Approach: Joint Inference of Boundary and Depth

- Glass and RGB-D sensors
- Boundary and region graph
- A Markov random field on boundaries and superpixels
- Joint prediction

Experiments: Dataset and Evaluation

- A new RGB-D glass dataset
- Performance comparisons
Color Similarity

- Histogramming hue and saturation values

Blurring

- Subtracting a smoothed image from the original
- Discrete cosine transform (DCT)

Overlay Consistency

- Adelson and Anandan’s linear model: \( I = \alpha I_B + e \)

Texture Distortion

- Response to a Gaussian filter bank

Highlights

- Highlight \textit{against} bright Lambertian regions
- A dichromatic color model
- Perimeter: affine \textit{against} piecewise constant
Surface and Object Models

Geodesic Active Contour Framework (McHenry et al., CVPR 2006)

- Discrepancy measures $D_{ij}$ and affinity measures $A_{ij}$
- Objective: \[
\max \sum_{i \in G, j \in O} D_{ij} + \sum_{i \in G, j \in G} A_{ij} - \sum_{i \in G, j \in O} A_{ij}\]

Additive Latent Feature Model (Fritz et al., NIPS 2009)

- Local features (SIFT) and latent Dirichlet allocation (LDA)
- Topic representation before quantizing into visual words
Overview to Our Approach

Figure: Illustration of the proposed approach. (a) Intensity image with ground truth foreground mask overlaid. (b) Edge detector output. (c) Triangulation result. (d) Boundary classifier output (magnified). (e) Superpixel classifier output (magnified). (f) Reconstructed depth with joint inference result overlaid.
Graph Construction

\[ E = \sum_{e_{ij} \in E} \psi_E(e_{ij}; I) + \beta \sum_{(ij, kl) \in G_E} \psi_E(e_{ij}, e_{kl}; I) + \gamma \sum_{d_i \in D} \phi_D(d_i; I) + \lambda \sum_{(i,j) \in G_D} \psi_D(d_i, d_j, e_{ij}; I) \]

- **Boundary Nodes**
  - \( \alpha \)
  - \( \phi_E(e_{ij}; I) \)
  - \( \psi_E(e_{ij}, e_{kl}; I) \)

- **Depth Nodes**
  - \( d_i \)
  - \( \phi_D(d_i; I) \)
  - \( \psi_D(d_i, d_j, e_{ij}; I) \)
  - \( d_j \)

**Symbols**
- \( E \)
- \( \psi \)
- \( \phi \)
- \( E \)
- \( \beta \)
- \( G \)
- \( D \)
- \( G_D \)
- \( I \)

**Equation**
\[ E = \sum_{e_{ij} \in E} \phi_E(e_{ij}; I) + \beta \sum_{(ij, kl) \in G_E} \psi_E(e_{ij}, e_{kl}; I) + \gamma \sum_{d_i \in D} \phi_D(d_i; I) + \lambda \sum_{(i,j) \in G_D} \psi_D(d_i, d_j, e_{ij}; I) \]
Boundary Potentials

Boundary Unary Potentials

- SVM and RF based on RGB cues and the missing depth cue: $e_{ij} \in (0, \pm 1)$
- Negative log-probability:
  \[
  \phi_E(e_{ij}; I) = -\log(P(e_{ij} | f_{ij}))
  \]

Boundary Pairwise Potentials

- Smoothness prior for connected boundary fragments:
  \[
  \psi_E(e_{ij}, e_{kl}) = 1 - \delta(e_{ij} = e_{kl} \neq 0) \\
  + C_1 \delta(e_{ij} = e_{kl} = 0) + C_2 \delta(e_{ij} \neq e_{kl})
  \]

- $C_1 = 0.3 \times \delta\left(\frac{\pi}{2} < \alpha \leq \pi\right)$
- $C_2 = (1 - \cos \alpha)^3 \delta\left(\frac{\pi}{2} < \alpha \leq \pi\right)$
Superpixel Potentials

Superpixel Unary Potentials

- Features extracted from triangular superpixels

Superpixel Pairwise Potentials

- Valid configurations of a boundary fragment and its neighboring superpixels:

\[ \psi_D(d_i, d_j, e_{ij}) = \begin{cases} -\delta(d_i = 0, d_j \neq 0, e_{ij} = +1) \\ \text{compatible configuration} \\ -\delta(d_i \neq 0, d_j = 0, e_{ij} = -1) \\ \text{compatible configuration} \\ +\delta(d_i \neq d_j, e_{ij} = 0) \\ \text{incompatible configuration} \end{cases} \]
Joint Prediction

Alternating Inference

- Goal: Applying boundary and superpixel constraints simultaneously
- Initialization: Loopy belief propagation (LBP) on boundary nodes
- Inference: Alternating mean-field approximation to compute the marginals of the boundary nodes and superpixel nodes

Depth Reconstruction

- Plane segmentation in 3D for each superpixel
- Glass objects with a piecewise planar model on supporting planes
- Median filtering for other missing regions
Our Results

Figure: Examples of glass detection results on our new RGB-D Glass dataset. Note that missing areas are shown in white, and depth readings are recovered by a piecewise planar model.
Performance Comparisons

<table>
<thead>
<tr>
<th></th>
<th>Intens.+ SVM</th>
<th>Intens.+ Depth</th>
<th>Detached Inference</th>
<th>Joint Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bound</td>
<td>19.52</td>
<td>44.38</td>
<td>54.08</td>
<td>62.27</td>
</tr>
<tr>
<td>Region</td>
<td>28.06</td>
<td>55.84</td>
<td>61.85</td>
<td>65.96</td>
</tr>
</tbody>
</table>

**Table:** F-measures at 50% recall for boundary and region accuracy metrics, respectively.

**Figure:** The precision-recall curves based on boundary matching (left panel) and pixel-wise region matching (right panel).
Thank you

Thanks for your time!

Questions or Comments?

For details please refer to the following publication:
Tao Wang, Xuming He, Nick Barnes. Glass object localization by joint inference of boundary and depth.