3D Model Assisted Image Segmentation

Srimal Jayawardena  Di Yang  Marcus Hutter
Australian National University

srimal(dot)jayawardena(at)anu(dot)edu(dot)au
http://users.cecs.anu.edu.au/~srimalj

DICTA 2011
The problem

Segmenting a mostly homogeneous (same color/texture) object into parts is a hard problem.
Methodology Overview

1. Photo of known object (Input)
2. 3D CAD Model (Input)
3. Pose estimation (hierarchically minimise gradient based loss)
4. Project 3D model parts (initialise contour detection)
5. Contour detection
6. Segmented parts
Gradient Loss for Pose Estimation

Let $\theta$ parameterize the pose of the 3D model w.r.t the camera.

(a) 3D Model Gradients $G_N(\theta)$

(b) Photo Gradients $G_I$

Loss at pose $\theta$,

$$L_g(\theta) := 1 - \left( corr\left( G_N(\theta), G_I \right) \right)^2 \in [0, 1]$$
3D Model Gradients

\( \Phi_x(u, v, \theta) \)  
\( \frac{\partial \Phi_x(u, v, \theta)}{\partial u} \)  
\( \frac{\partial \Phi_x(u, v, \theta)}{\partial v} \)

\( \Phi_y(u, v, \theta) \)  
\( \frac{\partial \Phi_y(u, v, \theta)}{\partial u} \)  
\( \frac{\partial \Phi_y(u, v, \theta)}{\partial v} \)

\( \Phi_z(u, v, \theta) \)  
\( \frac{\partial \Phi_z(u, v, \theta)}{\partial u} \)  
\( \frac{\partial \Phi_z(u, v, \theta)}{\partial v} \)

\( G_N(\theta)(u, v) = ||\nabla \Phi(u, v, \theta)||_k^k \)  

\[ (1) \]
Photo Gradients

(a) Real photo
(b) Synthetic

(c) Real $\frac{\partial I}{\partial u}$
(d) Synthetic $\frac{\partial I}{\partial u}$

(e) Real $\frac{\partial I}{\partial v}$
(f) Synthetic $\frac{\partial I}{\partial v}$

(g) Real $G_I$
(h) Synthetic $G_I$

$$G_I(u, v) = \|\nabla I(u, v)\|^k_k$$  \hspace{1cm} (2)
Overlays and Smoothing

(a) Real

(b) $n=0$

(c) $n=2$

(d) Synthetic

(e) $n=0$

(f) $n=2$
Loss Landscapes

(a) 2-norm

(b) 1-norm
Hierarchical Optimization

(a) Photo  (b) Background removed

(c) $n=2$  (d) $n=1$  (e) $n=0$

(f) Final fine pose $n=0$

Next: Initialise a *Level Set Evolution* contour detection from projected 3D model parts
Contour Detection

*Level Set Evolution* without re-initialization [Li et al., 2005, CVPR]

**Row 1:** Level set function, **Row 2:** Zero level curve

(a) Initialisation  (b)  (c)  (d) Final
Results

(a) Initialisation  (b) Result  (c) Benchmark GC  (d) Benchmark LS

(e) Initialisation  (f) Result  (g) Benchmark GC  (h) Benchmark LS
Results

(a) Initialisation  (b) Result  (c) Benchmark GC  (d) Benchmark LS

(e) Initialisation  (f) Result  (g) Benchmark GC  (h) Benchmark LS
Accuracy

Part segmentation results for two views of a Mazda Astina.

Accuracy = 1 − \left( \frac{\text{No. Misclassified. Pixels}}{\text{No. Ground. Truth. Pixels}} \right)

<table>
<thead>
<tr>
<th>Part</th>
<th>Side View</th>
<th>Semi Profile</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fender</td>
<td>97.7%</td>
<td>97.6%</td>
<td>97.7%</td>
</tr>
<tr>
<td>Front door</td>
<td>98.1%</td>
<td>95.3%</td>
<td>96.7%</td>
</tr>
<tr>
<td>Back door</td>
<td>96.8%</td>
<td>93.6%</td>
<td>95.2%</td>
</tr>
<tr>
<td>Mud flap</td>
<td>97.3%</td>
<td>95.1%</td>
<td>96.2%</td>
</tr>
<tr>
<td>Front window</td>
<td>97.8%</td>
<td>97.5%</td>
<td>97.7%</td>
</tr>
<tr>
<td>Back window</td>
<td>99.5%</td>
<td>93.9%</td>
<td>96.7%</td>
</tr>
</tbody>
</table>
Discussion

- Challenges - High amount of reflections and noise
- A closer initialisation curve - better results
- Future work - simultaneous pose estimation and segmentation

Thank you!
Level set evolution without re-initialization: a new variational formulation.