Multi-view stereo via volumetric graph-cuts

George Vogiatzis, Carlos Hernandez and Roberto Cipolla

Cambridge Research Laboratory
Toshiba Europe Research Laboratory
3D models from images

- State-of-the-art improving rapidly in recent years
- Various 3D cues have been exploited:
  - Silhouettes
  - Stereopsis (photo-consistency, multi-view stereo)
  - Shading (single/multiple images)
  - Texture
  - Defocus
Photo-consistency

Non photo-consistent point

≠
Photo-consistency
3D shape from *photo-consistency*

- Integrate this cue on surface
- Algorithms try to find the most *photo-consistent* 3D surface
- Key assumptions for object surface
  - Lambertian
  - Richly textured
  - Smooth
3D shape from *photo-consistency*

- Scene representation
  - Meshes, level-sets, voxel occupancy
- Shape prior
  - Local smoothness
- Photo-consistency metric
  - Normalised cross-correlation (NCC), sum of squared differences (SSD)
- Occlusion reasoning
  - How to determine visible images
- Reconstruction algorithm
  - How to obtain globally optimal solution
Our solution

- Volumetric Graph-cuts
  - Uses an occlusion-robust photo-consistency
  - Casts the problem as discrete Markov Random Field (MRF) optimisation, obtaining global solution
The occlusion problem

• To get 3D object must compute photo-consistency
• To compute photo-consistency, must know occluded cameras
• To know occluded cameras must know 3D object!
Find optimal depth with NCC
Find optimal depth with NCC

- Count number of local maxima of red curves for each depth (in small interval)
Combining depth-maps
Aggregating information
Occlusion-robust NCC

Averaged NCC

Robust NCC
3D MRF models

Labelling cost:
• Every voxel has a certain preference for being foreground or background

Compatibility cost:
• Every pair of neighbour voxels has a certain preference for being given the same or opposite labels

Optimal labelling can under some simple condition be obtained in polynomial time using “graph-cuts” (Boykov & Kolmogorov, ICCV 2003)
3D MRF for 3D modelling

- **Labelling cost:**
  - Constant bias towards being **foreground**

- **Compatibility cost:**
  - Pair of neighbour voxels prefers having opposite labels if photo-consistent region is between them

- Optimal voxel labelling can be computed using graph-cuts
- Computation takes approx. 7 mins for 512x512x512 grid on Pentium IV 2.6Ghz
3D Models
3D Models
3D models
## Middlebury evaluation (temple)

<table>
<thead>
<tr>
<th></th>
<th>Accuracy / Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Full</strong> (312 images)</td>
</tr>
<tr>
<td>Hernandez [10]</td>
<td>0.36mm / 99.7%</td>
</tr>
<tr>
<td>Goesele [9]</td>
<td>0.42mm / 98.0%</td>
</tr>
<tr>
<td>Hornung [12]</td>
<td>0.58mm / 98.7%</td>
</tr>
<tr>
<td>Pons [20]</td>
<td>–</td>
</tr>
<tr>
<td>Furukawa [8]</td>
<td>0.65mm / 98.7%</td>
</tr>
<tr>
<td>Vogiatzis [29]</td>
<td>1.07mm / 90.7%</td>
</tr>
<tr>
<td><strong>Present method</strong></td>
<td><strong>0.50mm / 98.4%</strong></td>
</tr>
</tbody>
</table>
Advantages

• **Accurate**
  – sub-millimetre accuracy on sequence with ground truth

• **Simple**
  – Can work with about 15-30 images

• **Fast**
  – Approximately 45’ of computation for these models
  – We believe we can bring this down to few minutes
Publications

  – Described occlusion-robust photo-consistency

  – Formulated the multi-view stereo problem as MRF inference

• New and improved version appears in PAMI (December 2007)