Robust and Accurate Multi-View Reconstruction by Prioritized Matching

Markus Ylimäki, Juho Kannala, Jukka Holappa and Janne Heikkilä
University of Oulu

Sami S. Brandt
University of Copenhagen
Outline

• Introduction
• The problem
• Related work
• Prioritized correspondence growing
• Proposed algorithm
• Comparison with the state of the art
• Conclusion
Introduction

- We propose a method for reconstruction using prioritized correspondence growing.
- The method is based on the best-first matching principle.
- The method takes a set of images and a sparse set of seed matches as input.
- The output of the method is a quasi-dense three-dimensional point cloud.
The problem

- **Input:**
- **Output:**

[Demo video wmv]  [Demo video mp4]
Related work

• Multi-view stereo is a classical problem in computer vision
• Multiple solutions using
  – Volumetric grid (e.g. Sinha ICCV07)
  – Depth maps (e.g. Merrell ICCV07) or
  – Surface expansion
    • Two-view matching (e.g. Lhuillier TPAMI02, Kannala CVPR07)
    • Multi-view matching (e.g. Furukawa TPAMI09, Koskenkorva ICPR10)
• No methods using prioritized matching with arbitrary number of images
Prioritized correspondence growing in general two-view stereo

- A set of seed matches are ordered into a priority queue based on their similarity scores
- The sorted seeds are expanded by iterating the following steps:
  a) The seed with the best score is taken from the queue
  b) New matches are searched nearby the seed
  c) Such candidates which quality satisfy a threshold are added to the queue as new seeds and to the final list of matches

Markus Ylimäki
November 14th 2012

Center for Machine Vision Research
Proposed algorithm

- Global representation of a seed $s$
Proposed algorithm (cont.)

- Expand in the reference views $a$ and $b$
Proposed algorithm – some details

- Total similarity score of a seed is a combination of pairwise ZNCC measures

- Intensity variance is used to prevent the propagation from spreading in too uniform areas

- Expansion is used only once
  - No filtering
Comparison with the state of the art

- PMVS program (Furukawa TPAMI09)
- Parameters were set so that the comparison is as fair as possible
- Experiments with four datasets:
  - Fountain-P11 and Herz-Jesu-P8 (Strecha CVPR08)
    - Evaluation of accuracy
    - Computational efficiency
  - The sparse ring Middlebury datasets of Dino and Temple (no ground truths available)
    - Visual evaluation
    - Computational efficiency
Evaluation of accuracy

• For datasets with known ground truths
• Fountain-P11
  – 11 images of size 786 x 512 pixels
Evaluation of accuracy

Three sample images from the Fountain-P11 dataset
Evaluation of accuracy

Furukawa’s result

Our result

Markus Ylimäki
November 14th 2012

Center for Machine Vision Research

UNIVERSITY of OULU
OULUN YLIOPISTO
Evaluation of accuracy

![Graph showing comparison between Occupancy (%) and Error for two methods: Ours and Furukawa's.](image)
Evaluation of accuracy (cont.)

- Herz-Jesu-P8
  - 8 images of size 786 x 512 pixels
Evaluation of accuracy (cont.)

Three sample images from the Herz-Jesu-P8 dataset
Evaluation of accuracy (cont.)

Furukawa’s result

Our result
Evaluation of accuracy (cont.)

Occupancy (%)

0 20 40 60 80 100

0 0.1 0.2 0.3 0.4

Error

Ours
Furukawa's
Visual evaluation

• For datasets without ground truths
• Temple sparse ring
  – 16 images of size 640 x 480 pixels
Visual evaluation

Three sample images from the Temple sparse ring dataset
Visual evaluation

Furukawa’s result

Our result

Markus Ylimäki
November 14th 2012

Center for Machine Vision Research
Visual evaluation (cont.)

- Dino sparse ring
  - 16 images of size 640 x 480 pixels
Visual evaluation (cont.)

Three sample images from the Dino sparse ring dataset
Visual evaluation (cont.)

Furukawa’s result

Our result

Markus Ylimäki
November 14th 2012

Center for Machine Vision Research
Computational efficiency

Number of reconstructed points

<table>
<thead>
<tr>
<th>Location</th>
<th>Number of Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fountain-P11</td>
<td>400000</td>
</tr>
<tr>
<td>Herz-Jesu-P8</td>
<td>200000</td>
</tr>
<tr>
<td>Temple</td>
<td>400000</td>
</tr>
<tr>
<td>Dino</td>
<td>200000</td>
</tr>
</tbody>
</table>

Execution time in seconds

<table>
<thead>
<tr>
<th>Location</th>
<th>Execution Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furukawa's</td>
<td>5000</td>
</tr>
<tr>
<td>Our</td>
<td>5000</td>
</tr>
<tr>
<td>Furukawa's</td>
<td>4000</td>
</tr>
<tr>
<td>Our</td>
<td>4000</td>
</tr>
<tr>
<td>Furukawa's</td>
<td>3000</td>
</tr>
<tr>
<td>Our</td>
<td>3000</td>
</tr>
<tr>
<td>Furukawa's</td>
<td>2000</td>
</tr>
<tr>
<td>Our</td>
<td>2000</td>
</tr>
<tr>
<td>Furukawa's</td>
<td>1000</td>
</tr>
<tr>
<td>Our</td>
<td>1000</td>
</tr>
</tbody>
</table>

Markus Ylimäki
November 14th 2012

Center for Machine Vision Research

Furukawa's
Our
Conclusion

• We propose a multi-view stereo reconstruction method

• The proposed approach:
  – Expands global seeds locally using the best-first matching principle
  – Uses the expansion only once
  – Produces reconstructions which quality is comparable to the state of the art but significantly faster
Thank you for your attention!