The ISPRS Benchmark on Urban Object Classification and 3D Building Reconstruction

Franz Rottensteiner a, Gunho Sohn b, Jaewook Jung b, Markus Gerke c, Caroline Baillard d, Sebastien Benitez d, Uwe Breitkopf a

a Institute of Photogrammetry & GeoInformation, University Hannover, Germany
b GeoICT Lab, York University, Toronto, Canada
c University of Twente, Faculty ITC; EOS Dept., Enschede, The Netherlands
d SIRADEL, Rennes, France
Outline

• Introduction
• Test Data
• Object Detection
  – Task and Evaluation Methodology
  – Evaluation of Building Detection
  – Evaluation of Tree Detection
  – Discussion
• 3D Building Reconstruction
  – Task and Evaluation Methodology
  – Results
  – Discussion
• Outlook
Benchmarks on Urban Object Extraction

- Benchmarks are widely used in Computer Vision to make different solutions comparable
- Photogrammetry & Remote Sensing:
  - Test data by OEEPE/EuroSDR: outdated as far as aerial sensor data are concerned
  - New data set of ISPRS WG III/4 “Complex scene analysis and 3D reconstruction”:
    http://www.itc.nl/ISPRS_WGIII4/tests_datasets.html

- Test with two tasks
  1) Urban object detection
  2) 3D building reconstruction

- Results submitted by the participants are evaluated by WG III/4 based on reference data
- This presentation: report on the results
Outline

• Introduction

• Test Data

• Object Detection
  – Task and Evaluation Methodology
  – Evaluation of Building Detection
  – Evaluation of Tree Detection
  – Discussion

• 3D Building Reconstruction
  – Task and Evaluation Methodology
  – Results
  – Discussion

• Outlook
Data Set 1: Vaihingen / Enz (Germany)

- **Multiple colour infrared (CIR) aerial images:**
  - Z/I DMC (c = 120 mm)
  - GSD: 8 cm
  - Fourfold overlap
  - Accuracy of orientation: ± 1 pixel

- **ALS data:**
  - Leica ALS 50
  - 4-7 pts/m², multiple pulses + intensities
  - DSM grid (25 cm)

- **Three test areas** (~150 m x 200 m)

- **Reference:** photogrammetric plotting
Data Set 2: Downtown Toronto (Canada)

- **Multiple colour (RGB) aerial images:**
  - Microsoft Vexcel UltraCam-D (c = 101.4 mm)
  - GSD: 15 cm
  - Stereo Overlap
  - Accuracy of orientation: ± 1 pixel

- **ALS data:**
  - Optech ALTM-ORION M
  - 6 pts/m², multiple pulses + intensities
  - DSM grid (25 cm)

- **Two test areas** (530 m x 600 m)

- **Reference:** photogrammetric plotting
Outline

• Introduction
• Test Data
• Object Detection
  – Task and Evaluation Methodology
  – Evaluation of Building Detection
  – Evaluation of Tree Detection
  – Discussion
• 3D Building Reconstruction
  – Task and Evaluation Methodology
  – Results
  – Discussion
• Outlook
Task of the Participants

• Detection of urban objects:
  – Buildings
  – Trees
  – (...)

• Deliverables per object class:
  – Object outlines as polygons in object space
  – (Geocoded images representing the detected objects)
Evaluation Methodology

• Evaluation of thematic accuracy:
  
  \[
  \text{Completeness} = \frac{TP}{TP + FN} \\
  \text{Correctness} = \frac{TP}{TP + FP} \\
  \text{Quality} = \frac{TP}{TP + FN + FP}
  \]

  ➢ TP: Number of True Positives  
  ➢ FP: Number of False Positives  
  ➢ FN: Number of False Negatives  

  per-area  
  per-object

• Evaluation of geometrical accuracy:
  
  – RMS error of distances from (correct) extracted outlines to reference outlines

• Requirements for practical relevance (Mayer et al, 2006):
  
  – Completeness ≥ 70%  
  – Correctness ≥ 85%
# Results Submitted by Participants

<table>
<thead>
<tr>
<th>Submitted by</th>
<th>Abbr.</th>
<th>Affiliation</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Moussa</td>
<td>CAL</td>
<td>University of Calgary, Canada</td>
<td>B+T</td>
</tr>
<tr>
<td>D. Bulatov</td>
<td>FIE</td>
<td>Fraunhofer Inst. Ettlingen, Germany</td>
<td></td>
</tr>
<tr>
<td>J. Niemeyer</td>
<td>HAN</td>
<td>University of Hannover, Germany</td>
<td>B+T</td>
</tr>
<tr>
<td>D. Grigillo</td>
<td>LJU</td>
<td>University of Ljubljana, Slovenia</td>
<td>B+T</td>
</tr>
<tr>
<td>C. Liu</td>
<td>TON</td>
<td>Tongji University, Japan</td>
<td>B</td>
</tr>
<tr>
<td>W. Yao</td>
<td>TUM</td>
<td>TU Munich, Germany</td>
<td>B+T</td>
</tr>
<tr>
<td>P. Dorninger</td>
<td>VSK</td>
<td>TU Vienna, Austria</td>
<td>B</td>
</tr>
<tr>
<td>Q. Zhan</td>
<td>WHU</td>
<td>Wuhan University, China</td>
<td>B+T</td>
</tr>
</tbody>
</table>

- Only one participant delivered results for Toronto (areas 4 + 5)
- One participant only delivered results for area 3
- Some participants only delivered results for building detection
## Data Used by the Participants

<table>
<thead>
<tr>
<th>Submitted by</th>
<th>Abbr.</th>
<th>Affiliation</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Moussa</td>
<td>CAL</td>
<td>University of Calgary, Canada</td>
<td>1+T</td>
</tr>
<tr>
<td>D. Bulatov</td>
<td>FIE</td>
<td>Fraunhofer Inst. Ettlingen, Germany</td>
<td>1</td>
</tr>
<tr>
<td>J. Niemeyer</td>
<td>HAN</td>
<td>University of Hannover, Germany</td>
<td>1+T</td>
</tr>
<tr>
<td>D. Grigillo</td>
<td>LJU</td>
<td>University of Ljubljana, Slovenia</td>
<td>1+T</td>
</tr>
<tr>
<td>C. Liu</td>
<td>TON</td>
<td>Tongji University, Japan</td>
<td>1</td>
</tr>
<tr>
<td>W. Yao</td>
<td>TUM</td>
<td>TU Munich, Germany</td>
<td>1+T</td>
</tr>
<tr>
<td>P. Dorninger</td>
<td>VSK</td>
<td>TU Vienna, Austria</td>
<td>1</td>
</tr>
<tr>
<td>Q. Zhan</td>
<td>WHU</td>
<td>Wuhan University, China</td>
<td>1+T</td>
</tr>
</tbody>
</table>

- Original images only (1)
- DSM (ALS) + orthophoto (4)
- DSM (ALS) (1)
- ALS points (2)
## Processing Strategies

<table>
<thead>
<tr>
<th>Submitted by</th>
<th>Abbr.</th>
<th>Affiliation</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Moussa</td>
<td>CAL</td>
<td>University of Calgary, Canada</td>
<td>B+T B+T B+T</td>
</tr>
<tr>
<td>D. Bulatov</td>
<td>FIE</td>
<td>Fraunhofer Inst. Ettlingen, Germany</td>
<td>B</td>
</tr>
<tr>
<td>J. Niemeyer</td>
<td>HAN</td>
<td>University of Hannover, Germany</td>
<td>B+T B+T B+T</td>
</tr>
<tr>
<td>D. Grigillo</td>
<td>LJU</td>
<td>University of Ljubljana, Slovenia</td>
<td>B+T B+T B+T</td>
</tr>
<tr>
<td>C. Liu</td>
<td>TON</td>
<td>Tongji University, Japan</td>
<td>B B B</td>
</tr>
<tr>
<td>W. Yao</td>
<td>TUM</td>
<td>TU Munich, Germany</td>
<td>B+T B+T B+T B</td>
</tr>
<tr>
<td>P. Dorninger</td>
<td>VSK</td>
<td>TU Vienna, Austria</td>
<td>B B B</td>
</tr>
<tr>
<td>Q. Zhan</td>
<td>WHU</td>
<td>Wuhan University, China</td>
<td>B+T B+T B+T</td>
</tr>
</tbody>
</table>

- Supervised, without segmentation (3)
- Unsupervised, with segmentation (5)
Example for the Evaluation of Building Detection – Area 1
Evaluation of Building Detection - Vaihingen

• Area-based evaluation (average Areas 1-3)

• All methods fulfil the requirements for practical relevance according to (Mayer et al., 2006)
Evaluation of Building Detection - Vaihingen

- Object-based evaluation (all objects, average Areas 1-3)

- HAN, TON and WHU do not fulfil the requirements for practical relevance according to (Mayer et al., 2006)

- Problems are largely related to small building structures
Evaluation of Building Detection - Vaihingen

• Object-based evaluation for buildings \( \geq 50 \) m\(^2\) (average Areas 1-3)

• Large structures can be detected reliably by most methods

• All methods fulfil the requirements for practical relevance according to (Mayer et al., 2006)
Evaluation of Building Detection - Vaihingen

- \( \text{RMS}_{XY} \): error of planimetric distances (average Areas 1-3)

- Order of magnitude: 2 - 4 times the resolution of ALS data

- Full potential of images is not exploited
Evaluation of Building Detection - Toronto

- Evaluation for TUM (average of Areas 4 and 5)

- More difficult scenario than Vaihingen
  - No Infrared band
  - Occlusions, extremely high-rise buildings

- Nevertheless, requirements according to (Mayer et al., 2006) fulfilled

Institut für Photogrammetrie und GeoInformation
Common Problems in Building Detection

• Horizontal roof planes in combination with complex terrain, objects on roofs
Common Problems in Building Detection

• Horizontal roof planes in combination with complex terrain, objects on roofs

• Dense tree canopies → Problems for methods directly classifying ALS points
Common Problems in Building Detection

• Horizontal roof planes in combination with complex terrain, objects on roofs
• Dense tree canopies → Problems for methods directly classifying ALS points
• Roofs covered by grass
Common Problems in Building Detection

- Horizontal roof planes in combination with complex terrain, objects on roofs
- Dense tree canopies → Problems for methods directly classifying ALS points
- Roofs covered by grass
- Large trees next to buildings for segmentation-based methods
Common Problems in Building Detection

- Horizontal roof planes in combination with complex terrain, objects on roofs
- Dense tree canopies → Problems for methods directly classifying ALS points
- Roofs covered by grass
- Large trees next to buildings for segmentation-based methods
- Small building structures
Common Problems in Building Detection

- Horizontal roof planes in combination with complex terrain, objects on roofs
- Dense tree canopies → Problems for methods directly classifying ALS points
- Roofs covered by grass
- Large trees next to buildings for segmentation-based methods
- Small building structures
- Occlusion / perspective distortion
Common Problems in Building Detection

- Horizontal roof planes in combination with complex terrain, objects on roofs
- Dense tree canopies → Problems for methods directly classifying ALS points
- Roofs covered by grass
- Large trees next to buildings for segmentation-based methods
- Small building structures
- Occlusion / perspective distortion
- Shadow
- Very low buildings in CBD
Example for the Evaluation of Tree Detection– Area 1

- CAL
- HAN
- LJU
- TUM
- WHU

TP (True Positive), FP (False Positive), FN (False Negative)
Evaluation of Tree Detection - Vaihingen

- Object-based evaluation (average Areas 1-3)

- Low performance for all methods compared to building detection

- No method achieves the requirements for practical relevance according to (Mayer et al., 2006)
Evaluation of Tree Detection - Vaihingen

• Object-based evaluation for trees \( \geq 50 \) m\(^2\) (average Areas 1-3)

• 50 m\(^2\) correspond to a crown diameter of about 8 m

• All methods except LJU fulfil the requirements for practical relevance according to (Mayer et al, 2006)

• Problems of tree detection: similar to problems in building detection
Object Detection: Discussion

• **Buildings:**
  - Main buildings per plot can be detected reliably by most methods
  - Most methods can be practically relevant
  - Small buildings remain a problem
  - Occlusions in CBD scene → Multiple overlap required?
  - Slight advantage for methods based on segmentation
  - Full geometrical accuracy potential of images not yet exploited

• **Trees:**
  - More problematic than buildings
  - Practical relevance questionable for small trees

• **Most favourable conditions:** Area 2 (high-rise residential)
Outline

• Introduction

• Test Data

• Object Detection
  – Task and Evaluation Methodology
  – Evaluation of Building Detection
  – Evaluation of Tree Detection
  – Discussion

• 3D Building Reconstruction
  – Task and Evaluation Methodology
  – Results
  – Discussion

• Outlook
Task of the Participants

• 3D Reconstruction of Buildings (LoD2 of CityGML standard):
  – Detailed roof structures
  – No roof overhangs, balconies

• Deliverables:
  – Roof plane outlines as 3D polygons in object space
Evaluation Methodology

• Quality of roof plane segmentation:
  – Completeness / Correctness / Quality of roof planes
    ➢ Check whether a plane has a substantial overlap with planes in the other data set

• Geometrical Accuracy:
  – $\text{RMS}_{XY}$: RMS errors of the planimetric distances of roof plane boundary points from reference
  – $\text{RMS}_Z$: RMS errors of height differences between synthetic DSMs
  – $\text{RMS}_Z$ is also affected by segmentation errors
### Results Submitted by Participants

<table>
<thead>
<tr>
<th>Submitted by</th>
<th>Abbr.</th>
<th>Affiliation</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>W. Zhang</td>
<td>BNU</td>
<td>Beijing Normal University, China</td>
<td>1 X</td>
</tr>
<tr>
<td>J.-Y. Rau</td>
<td>CKU</td>
<td>Cheng Kung University, Taiwan</td>
<td>4 X X X X X</td>
</tr>
<tr>
<td>D. Bulatov</td>
<td>FIE</td>
<td>Fraunhofer Inst., Ettlingen, Germany</td>
<td>4 X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE1</td>
<td>ITC, Enschede, The Netherlands</td>
<td>4 X X X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE2</td>
<td>ITC, Enschede, The Netherlands</td>
<td>4 X X X</td>
</tr>
<tr>
<td>B. Xiong</td>
<td>ITCX</td>
<td>ITC, Enschede, The Netherlands</td>
<td>4 X X X</td>
</tr>
<tr>
<td>P. Dorninger</td>
<td>VSK</td>
<td>TU Vienna, Austria</td>
<td>4 X X X</td>
</tr>
<tr>
<td>G. Sohn</td>
<td>YOR</td>
<td>York University, Canada</td>
<td>X X X X X X</td>
</tr>
</tbody>
</table>

- Two participants delivered results for Toronto (areas 4 + 5)
- Two participants only delivered results for area 3
### Data used by the participants

<table>
<thead>
<tr>
<th>Submitted by</th>
<th>Abbr.</th>
<th>Affiliation</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>W. Zhang</td>
<td>BNU</td>
<td>Beijing Normal University, China</td>
<td>1 2</td>
</tr>
<tr>
<td>J.-Y. Rau</td>
<td>CKU</td>
<td>Cheng Kung University, Taiwan</td>
<td>X X</td>
</tr>
<tr>
<td>D. Bulatov</td>
<td>FIE</td>
<td>Fraunhofer Inst., Ettlingen, Germany</td>
<td>X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE1</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X X X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE2</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X X X</td>
</tr>
<tr>
<td>B. Xiong</td>
<td>ITCX</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X X X</td>
</tr>
<tr>
<td>P. Dorninger</td>
<td>VSK</td>
<td>TU Vienna, Austria</td>
<td>X X X</td>
</tr>
<tr>
<td>G. Sohn</td>
<td>YOR</td>
<td>York University, Canada</td>
<td>X X X</td>
</tr>
</tbody>
</table>

**Areas:**
- 1
- 2
- 3
- 4
- 5

- **Images only** (2)
- **Images + ALS points** (1)
- **ALS points** (5)
<table>
<thead>
<tr>
<th>Submitted by</th>
<th>Abbr.</th>
<th>Affiliation</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>W. Zhang</td>
<td>BNU</td>
<td>Beijing Normal University, China</td>
<td></td>
</tr>
<tr>
<td>J.-Y. Rau</td>
<td>CKU</td>
<td>Cheng Kung University, Taiwan</td>
<td>X</td>
</tr>
<tr>
<td>D. Bulatov</td>
<td>FIE</td>
<td>Fraunhofer Inst., Ettlingen, Germany</td>
<td>X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE1</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE2</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X</td>
</tr>
<tr>
<td>B. Xiong</td>
<td>ITCX</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X</td>
</tr>
<tr>
<td>P. Dorninger</td>
<td>VSK</td>
<td>TU Vienna, Austria</td>
<td>X</td>
</tr>
<tr>
<td>G. Sohn</td>
<td>YOR</td>
<td>York University, Canada</td>
<td>X</td>
</tr>
</tbody>
</table>

- Semi-automatic (2)
- Fully automatic (6)
### Building Models

<table>
<thead>
<tr>
<th>Submitted by</th>
<th>Abbr.</th>
<th>Affiliation</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>W. Zhang</td>
<td>BNU</td>
<td>Beijing Normal University, China</td>
<td>X</td>
</tr>
<tr>
<td>J.-Y. Rau</td>
<td>CKU</td>
<td>Cheng Kung University, Taiwan</td>
<td>X X X X X</td>
</tr>
<tr>
<td>D. Bulatov</td>
<td>FIE</td>
<td>Fraunhofer Inst., Ettlingen, Germany</td>
<td>X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE1</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X X X</td>
</tr>
<tr>
<td>S. Oude Elberink</td>
<td>ITCE2</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X X X</td>
</tr>
<tr>
<td>B. Xiong</td>
<td>ITCX</td>
<td>ITC, Enschede, The Netherlands</td>
<td>X X X</td>
</tr>
<tr>
<td>P. Dorninger</td>
<td>VSK</td>
<td>TU Vienna, Austria</td>
<td>X X X</td>
</tr>
<tr>
<td>G. Sohn</td>
<td>YOR</td>
<td>York University, Canada</td>
<td>X X X X X X</td>
</tr>
</tbody>
</table>

- **Generic (polyhedral)** (4)
- **Primitives** (1)
- **Adaptive** (3)
Results: Examples (Vaihingen, Area 1)

3D Visualisation (YOR)
Evaluation of Building Reconstruction - Vaihingen

- Evaluation for all roof planes (average Areas 1-3)

- Correctness > 80% for all methods

- Number of undetected roof planes > 25% except for CKU (semi-automatic), FIE, YOR
Evaluation of Building Reconstruction - Vaihingen

- Evaluation for roof planes >10 m² (average Areas 1-3)

- Correctness > 89% for all methods

- Number of undetected planes is still relatively large for most methods
Evaluation of Building Reconstruction - Vaihingen

- Geometrical errors (average Areas 1-3)

- Order of magnitude (planimetry): 3 – 4 times the ALS point spacing
- Height errors are relatively large (influenced by segmentation errors)
- Full accuracy potential not exploited
Evaluation of Building Reconstruction - Toronto

• Evaluation for roof planes (average Areas 4 & 5)

  ![Graphs showing evaluation results for all planes and planes > 10 m² in CKU and YOR.]

• Correctness > 80% for all methods

• Slightly lower completeness / correctness values than in Vaihingen

• RMS errors larger than Vaihingen
  - YOR: 1.05 m (X,Y) / 15 m (Z)
Common Problems in Building Reconstruction

• Small roof planes
Common Problems in Building Reconstruction

• Small roof planes

• Small appendices to larger roof planes
Common Problems in Building Reconstruction

• Small roof planes
• Small appendices to larger roof planes
• Undersegmentation

(ITC1)
Common Problems in Building Reconstruction

• Small roof planes
• Small appendices to larger roof planes
• Undersegmentation
• Wrong segmentation
Common Problems in Building Reconstruction

• Small roof planes

• Small appendices to larger roof planes

• Undersegmentation

• Wrong segmentation

• Missing regularisation or over-regularisation
Common Problems in Building Reconstruction

- Small roof planes
- Small appendices to larger roof planes
- Undersegmentation
- Wrong segmentation
- Missing regularisation or over-regularisation
- Incorrect combination of planes
Building Reconstruction: Discussion

- Building reconstruction works well for
  - Simply-shaped buildings
  - Buildings whose dormers are small compared to dominant roof planes

- Complex objects do not just lead to more generalized models

- Accuracy potential of the sensors not yet fully exploited

- Most favourable conditions: Area 2 (high-rise residential)

- Results are generally sufficient for a ‘nice’ visualisation

- Fully automatic generation of topologically and geometrically correct models in complex environments is still a challenge
Outline

• Introduction
• Test Data
• Object Detection
  – Task and Evaluation Methodology
  – Evaluation of Building Detection
  – Evaluation of Tree Detection
  – Discussion
• 3D Building Reconstruction
  – Task and Evaluation Methodology
  – Results
  – Discussion
• Outlook
Outlook

• The data may still be obtained via the WWW:  
  http://www.itc.nl/ISPRS_WGIII4/tests_datasets.html

• Results are still accepted and evaluated

• Results are available in the WWW

• Changed conditions of use: Vaihingen data may be used for any scientific purpose

• Basis for a standard data set for urban object extraction

• Special issue of a journal will be announced soon

• Interested in individual methods?
  ➔ Sessions  ISPRS Benchmark – 1 at 14:00 – 15:30, Room MCEC 219  
  ISPRS Benchmark – 2 at 16:00 – 17:30, Room MCEC 219