UNIVERSITEIT TWENTE.

To
ISPRS Secretary General
Prof. Dr. Christian Heipke
Leibniz University of Hannover
Institute of Photogrammetry and GeoInformation
Nienburger Str. 1
D-30167 Hannover

Report for 2015, Scientific Initiative
ISPRS Benchmark on Urban Object Detection and 3D Building Reconstruction

Dear Christian,

Hereby I submit the report 2015 on behalf of ISPRS WG III/4.

Title of project:
ISPRS Benchmark on Urban Object Detection and 3D Building Reconstruction

Principal Investigator (on behalf of ISPRS WG III/4):
Markus Gerke
Postal address: ITC, EOS Department
University of Twente
Hengelosestraat 99, P.O. Box 6
7500 AA Enschede
The Netherlands
E-mail: m.gerke@utwente.nl
Tel.:   +31 53 4874 522
Position inside ISPRS: Co-Chair of ISPRS WG III/4 "3D Scene Analysis"

DE ONDERNEMENDE UNIVERSITEIT

Postbus 217
7500 AE Enschede
www.utwente.nl

UNIVERSITEIT TWENTE.

Summary

A major yet unsolved research topic in photogrammetry and remote sensing is automated extraction of urban objects from high-resolution data. In practice maps are still drawn manually to a large extent, which is very costly both in terms of time and money. With this project we aim at promoting 2D semantic segmentation that assigns object-class labels of multiple relevant urban categories to image pixels. It can be considered as a first step towards the long-term goal of completely automating mapping.

What makes this task challenging is the very heterogeneous appearance of objects like buildings, streets, trees and cars in very high-resolution data, which leads to high intra-class variance while the inter-class variance is low.

Research drivers are very high-resolution data from new sensors and advanced processing techniques that rely on increasingly mature machine learning techniques. Despite the enormous efforts spent, these tasks cannot be considered to be solved, yet. To our knowledge, no fully automated method for 2D object recognition is applied in practice today although at least two decades of research have tried solving this task. One major problem that is hampering scientific progress is a lack of standard data sets for evaluating 2D semantic labeling methods, so that the outcomes of different approaches can hardly be compared experimentally. This "semantic labeling contest" of ISPRS WG III/4 is meant to resolve this issue. The 2D semantic labeling challenge (http://www2.isprs.org/semantic-labeling.html) extends the benchmark to pixel-precise classification of urban objects.

In the first year of this project funded by the ISPRS Scientific Initiative we have prepared the data set Vaihingen and launched the 2D semantic labeling challenge. This involved preparation of all data, labeling ground truth, and the design and implementation of an evaluation strategy.
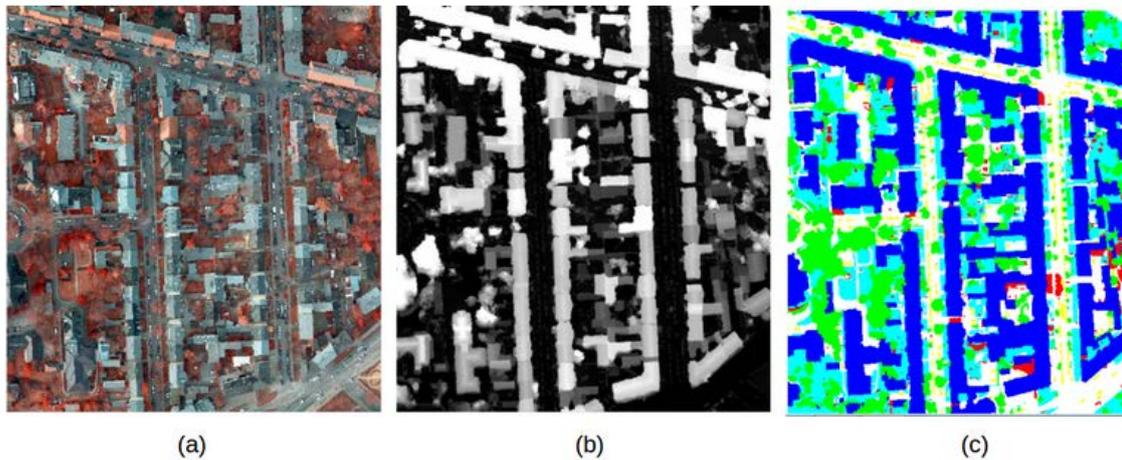
This report focusses on the work undertaken in 2015, the final year of the granting period, and gives a brief overview of results submitted to the 2D labeling benchmark so far. See http://www.isprs.org/documents/reports/SI-2014/ISPRS_SI_report-website_WGIII4_Gerke_2014.pdf for the online version of the report on the first period.

Activities from October 2014 till November 2015

The major labor work went into the creation of data for the extension of the 2D semantic labelling benchmark for Potsdam, Germany. In particular the following steps have been performed:

- Production of a photogrammetric DSM from the Potsdam images, and a true orthophoto, GSD of 5cm. Images were provided by the company BSF Swissphoto for free.
- Subdivision of the entire area into 38 tiles of fixed size (6000x6000pixels). Each tile is georeferenced, so if desired by participants, ortho mosaics (or large DSMs) of any size and tile combination can easily be compiled.
- Production of a reference for labeling in all 38 tiles, use of 5 relevant urban classes: *impervious surfaces, building, low vegetation, tree, car* (plus *background/clutter*). Note that those classes are the same as the ones used for Vaihingen
- Development of the following benchmarking strategy: we release all image data to participants, but the reference is only provided for 24 tiles. Participants deliver their classification for the remaining 14 tiles; the evaluation will only be based on the results obtained for these tiles. In contrast to the Vaihingen dataset the tiles with reference information are spatially connected. This enables participants to learn the context more efficiently if needed.
- Update of our webpage with the information on the benchmark
- Development of software to perform an efficient, fully automatic evaluation of results, including automatic update of the website with the results
- The data have been available to the scientific community since the ISPRS Geospatial Week in La Grande Motte, France, October 2015

See the figure below for screenshots of the website http://www2.isprs.org/semantic-labeling.html

UNIVERSITEIT TWENTE.

Example patches of the semantic object classification contest with (a) true orthophoto, (b) DSM, and (c) ground truth

Like for Vaihingen the results are shown in a table: major quality indicators in one row per participant, details on extra pages, see links at http://www2.isprs.org/commissions/comm3/wg4/results.html

Potsdam: 2D Labelling challenge

All quality measures except for *overall* are F1 scores in [%] using the reference with eroded boundaries. Mouse over the column headings will display more information.

Abbrev.	imp surf	building	low_veg	tree	car	Overall	Strategy	Details
SVL_1	83.5	91.7	72.2	63.2	62.2	77.8	s	D P
SVL_3	84.0	89.8	72.0	59.0	69.8	77.2	s	D P

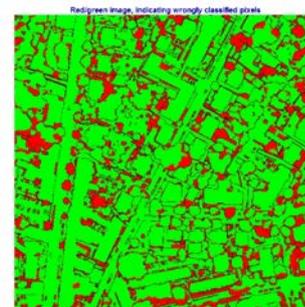Result output created on December 7, 2015

The "detailed" page link gives per submission confusion matrices, both an overall matrix and for each tile, and a visualization of wrongly classified pixels:
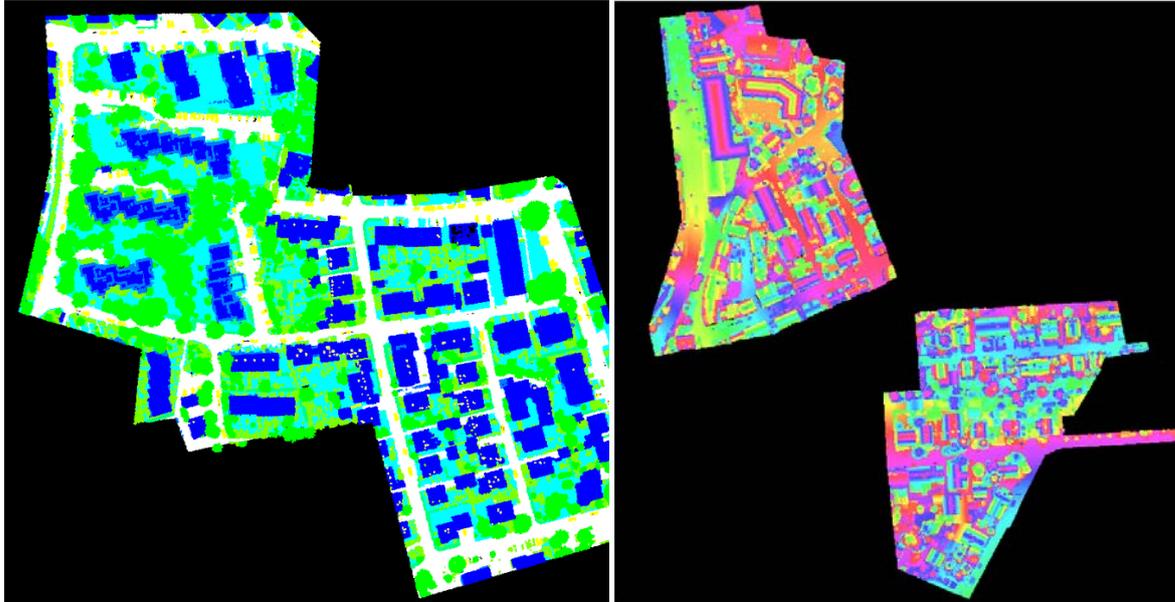
Overall statistics, reference set: no_boundary

Number of **non processed tiles** by this participant: 0

↓ predicted ‖ reference →	imp_surf	building	low_veg	tree	car	clutter
imp_surf	**0.873**	0.023	0.065	0.021	0.002	0.015
building	0.028	**0.915**	0.011	0.035	0.000	0.010
low_veg	0.088	0.015	**0.783**	0.102	0.000	0.011
tree	0.111	0.035	0.275	**0.566**	0.001	0.012
car	0.309	0.019	0.039	0.024	**0.489**	0.120
clutter	0.467	0.089	0.197	0.035	0.006	**0.205**
Precision/Correctness	0.799	0.918	0.670	0.715	0.855	0.394
Recall/Completeness	0.873	0.915	0.783	0.566	0.489	0.205
F1	**0.835**	**0.917**	**0.722**	**0.632**	**0.622**	**0.270**

Overall accuracy 0.778
===

In addition to the 2D semantic labeling benchmark for Potsdam we also released the **3D labeling benchmark for Vaihingen**. Refer to the report on this scientific initiative project from last year. In this reporting period the technical part of preparing the release data and setting up the website was finalized.

Example view on the Training area (color codes the label)

View onto 3D data provided for evaluation (color codes the height)

The result-table has a similar structure as the one for the 2D labeling part. Only some major classes are shown (the same as for 2D). The details page shows a confusion matrix with all 9 classes and an 2D raster of points indicating areas with wrong decisions. The screenshot shows the table on the landing page of 3D labeling results. Note that this entry is a "dummy" entry, just to show the frontend.

All quality measures except for *overall* are F1 scores in [%]. Note that in this table only a selection of classes - similar to the 2D case - are shown. The respective details-page shows the full statistics. Mouse over the column headings will display more information.

Abbrev.	imp surf	roof	low_veg	tree	car	Overall	Strategy	Details
MG1	--	95.1	66.2	99.4	100.0	72.8	s	D P

Result output created on December 7, 2015

Activities related to the benchmarks at conferences
- New task (announced at the PCV 2014): Semantic image labelling; based on the Vaihingen data set;
- New data set for the semantic image labelling task (announced at CMRT15): Potsdam;
- New task (announced at CMRT15): Semantic 3D point cloud labeling;
- The data set was also announced to the Geobia community; we are contemplating some kind of interaction with the next Geobia conference, which is going to take place in Enschede in 2016; details still need to be worked out.

Results submitted for the labeling benchmarks

UNIVERSITEIT TWENTE.

As of mid-December 2015 we have 20 submissions for the Vaihingen dataset and 2 for the Potsdam semantic labeling challenge. So far, 1 participant provided results for the 3D semantic labeling part. Most participants provide multiple resultant datasets by using their method with different options. This is very helpful since it allows a (relative) comparison of impact from these settings.
We observe a dominance of supervised techniques for semantic labeling and object extraction. Four of the participants are affiliated to computer vision departments. While one participant delivered results from an unsupervised (rule-based) approach, the others work with supervised methods. Approaches based on Convolutional Neural Networks are gaining importance. This is also indicated by the results of the labeling benchmark. Although a detailed analysis by the WG chairs is still pending, first results delivered by such methods seem to be quite promising.

Data request – statistics
From our records we cannot give a sound statistics on how many researchers are actually working with the data. However, the newsletter to which only people who requested the data are invited is currently listing about 200 entries. We might assume that those researchers are at least interested in this benchmark and its developments.

Outlook
Although the financial support by ISPRS will end this year we plan to continue evaluating results submitted for the datasets. Current WG officers plan to submit an overview paper on the main findings. The future structure of commissions and working group interest of the current WG chairs is only definite after the Prague congress. A detailed planning towards possible extension of the challenge will be done in due time.

Justification of money spent in 2015
In 2014 we spent CHF 6,221, while ISPRS provided us CHF 10,000. In 2015 we received another 6,000CHF. So, in total CHF 9,779 were available for 2015. Since the Potsdam data is of much higher complexity, the processing but also the labeling needed much more attention than estimated. The workflow followed was the same as in 2014, refer to the old report. The total costs for students assistants in 2015 is CHF 11,745 (see separate justification sent to the treasurer). The difference of approximately CHF 2,000 has been covered by the budget of the Institute of Geodesy and Photogrammetry of the ETH Zurich.