

Report for 2014, Scientific Initiative

ISPRS Benchmark on Urban Object Detection and 3D Building Reconstruction

Principal Investigator (on behalf of ISPRS WG III/4):

Markus Gerke, University of 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 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 an implementation of an evaluation strategy. In the remaining time of the project we will add at least one more data set that covers a bigger area and provides even higher resolution data.

Activities from January 2014 till mid-September 2014

The activities which have been carried out in the first 3 quarters of the project year (starting January 2014) deviate from the original planning due to several reasons, which will be explained below.

Upon start of the project we decided to begin with the ground truth production for the 2D labeling challenge using the Vaihingen dataset. The motivation for this was that we saw a decrease in participation in the existing benchmark which focusses on 2D object detection and 3D building reconstruction from laserscanning data or airborne images.

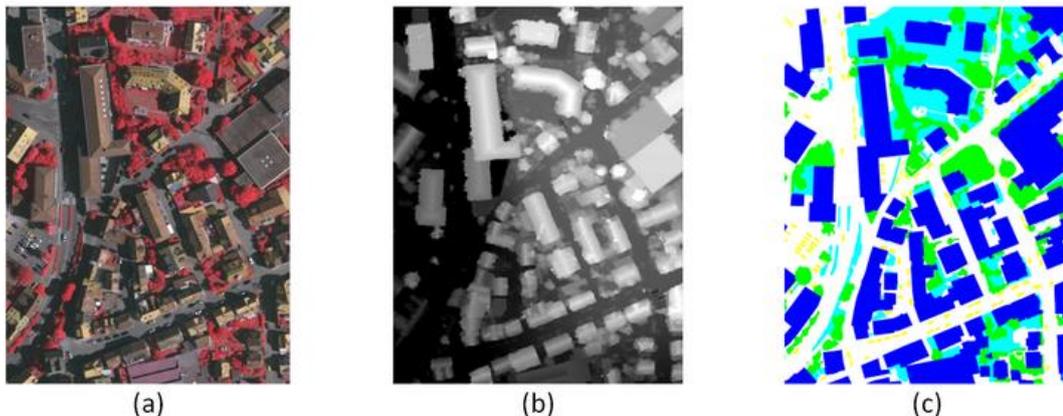
Many parties which were interested in the benchmark as such were not able to handle this specific data type. We concluded that a strategy which is based on true orthophotos and raster DSMs, which are provided by us, would attract additional researchers. We believe that the use of data formats which are easier to handle (also by common software packages for image interpretation) would increase the value of the benchmarking.

We also recognized that a complete semantic labeling of urban scenes is of interest not only for the remote sensing/photogrammetry communities, but also for computer vision research groups.

In the following list the activities we performed in Q1 to Q3 2014 related to the **2D semantic labeling benchmark (Vaihingen)** are summarized:

- Production of a photogrammetric DSM from the Vaihingen images, and a true orthophoto of good quality, GSD of 9cm.
- Subdivision of the entire area into 33 tiles of varying size (on average about 2000x2500pixels)
- Production of a reference for labeling in all 33 tiles, use of 5 relevant urban classes: *impervious surfaces, building, low vegetation, tree, car* (plus *background/clutter*).
- Development of the following benchmarking strategy: we release all image data to participants, but the reference is only provided for 16 tiles. Participants deliver their classification for the remaining 17 tiles; the evaluation will only be based on the results obtained for these tiles. This means that training in supervised classification approaches can be done using approximately half of the area.
- 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 website is online and data can be requested since the ISPRS TC III midterm symposium in Zurich, begin of September 2014

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



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

The result front page is available through <http://www2.isprs.org/vaihingen-2d-semantic-labeling-contest.html>, see screenshot below (only showing dummy data because no results from participants have been received so far)

ISPRS Semantic Labeling Contest (2D): Results

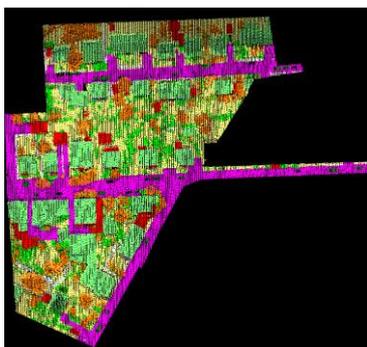
[Click here for a description of evaluation measures](#)

Vaihingen: 2D Labelling challenge

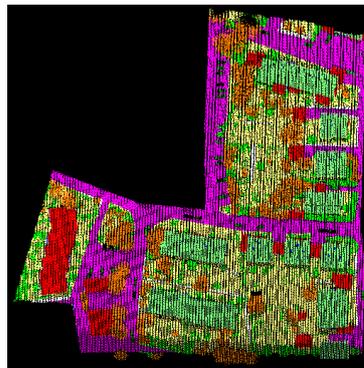
All quality measures except for *overall* are F1 scores in [%] using the reference with eroded boundaries. Mouse over the column headings will display more information.

Abbrev.	imp surf	building	low_veg	tree	car	Overall	Strategy	Details
P1	97.3	87.7	82.9	72.7	78.1	85.8	u	D P
P2	97.3	87.7	82.9	72.7	78.1	85.8	m	D P
P3	98.0	82.3	81.5	74.4	77.1	83.5	h	D
MG4	97.3	87.7	82.9	72.7	78.1	85.8	h	D P

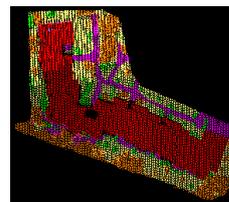
In addition to the 2D semantic labeling benchmark (Vaihingen) we prepared also the **3D labeling benchmark for Vaihingen**. That is, we obtained reference label information for a part of the ALS point cloud. Two regions with reference labels for the ALS points will again be made available for training in supervised methods, whereas the labels of the points inside the three test areas used in the object detection benchmark are to be determined by test participants (see figure below). So, this contest will be quite similar in principle to the 2D challenge, though the evaluation program needs to be adapted due to the fact that we cannot expect the participants to deliver points ordered in the same way as the input point cloud. The reference labels were received from the authors of (Niemeyer et al., 2014), who distinguished the classes *road*, *gable roof*, *flat roof*, *façade*, *grass*, *tree* and *low vegetation*.



Reference labels: Training area 1



Training area 2



Training area 3



Reference labels: Test area 1

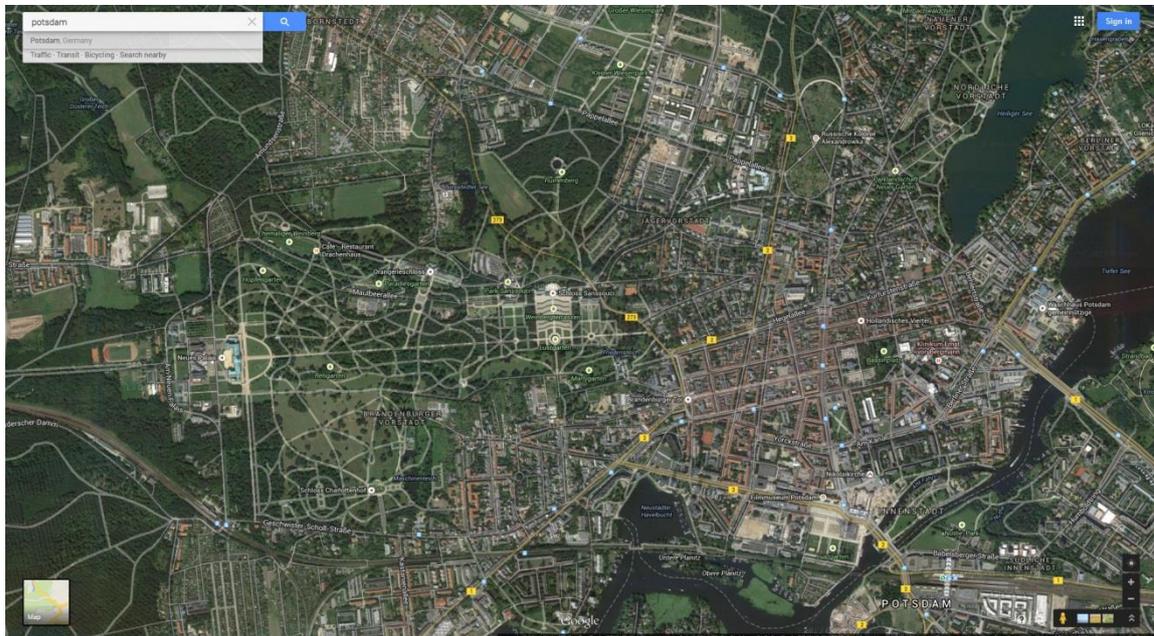
Test area 2

Test area 3

Outlook: Q4 2014 and 2015

In the remaining time of the granting period we will focus on the following actions:

- Develop/adopt the evaluation software and website for the 3D labeling benchmark of Vaihingen
- Repeat all the necessary steps to realize a 2D semantic labeling benchmark using the Potsdam data (photogrammetric DSM, true orthophoto, reference, website etc)
- Production of reference data for the 2D detection and 3D building reconstruction benchmark from Potsdam, update of the website and software accordingly



Screenshot of Google Maps of the area covered by the Potsdam data set

References

Niemeyer, J.; Rottensteiner, F.; Sörgel, U. (2014): Contextual classification of lidar data and building object detection in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 87, January 2014, pp 152-165.