An improved web-based height map for the Netherlands using AHN2

MSc Geomatics Final Graduation Plan
Delft University of Technology

Kees Jonker
1528416

January 20, 2015

Personal information

Name: Kees Jonker
Student number: 1528416
Adress: Veluwemeer 10
Postal code: 2729PC
City: Zoetermeer
Phone: 06-49094041
Email: c.jonker@student.tudelft.nl

Support and supervision

Graduation Professor (main mentor): Prof. dr. J.E. Stoter
Daily Supervisor (second mentor): dr. H. Ledoux

1 Introduction

The recent emergence of LiDAR (Light Detection And Ranging)) scanning technology has resulted in the availability of very large three dimensional point cloud data sets. For the Netherlands, the AHN2 (Algemeen Hoogtebestand Nederland) is a country-wide high-resolution dataset comprising nation's terrain by measuring height obtained by LiDAR, using airborne laser scanning (ALS). The average point density of the dataset is between 6 and 10 points per square meter (Van der Zon, 2013)). Such a high point density makes the AHN2 dataset quite large; it contains almost 640 billion laser points (Kadaster, 2014). In practice it appears that collected data on a local scale deviates from the predefined specifications Van der Zon (2013) gives different causes for this problem:

- Lower point density on different types of surface;
- Missing data next to high buildings;
- Ridges (on farmland);
- Puddles on grass land and;
- Data below trees.

Since collection of the AHN2 dataset was already finished in 2012 and made publicly available as Open Data since March 2014, research has been done on the development of (geo-)products based on the AHN2 dataset. Two groups of products can be distinguished:

- Vector-based (point-based) and;
- Raster-based.

Where the first group of products focuses directly on visualization of the point cloud data itself (Wouda, 2011; Geonext, 2014), the last group of products use spatial interpolation to calculate continuous (or non-discrete) data out of the discrete vector samples. This thesis research will focus on the latter strategy, a raster-based approach.
Since raw point clouds are not acquired on a uniform grid but can be seen as a set S of n (arbitrary) points in R^2 with an associated elevation function $h : S \to R$. To construct a grid DEM (Digital Elevation Model), h has to be extended via interpolation to a uniform grid $G \subset R^2$ at the desired resolution (Beutel, 2011).
By applying this strategy, a reconstruction will take place of the surface that has been recorded. When looking to existing raster-based products, it becomes clear that the used methodologies to apply such a reconstruction are far from good. Products contain errors, for example, holes (no-data areas) and also dynamic objects (cars, animals) are still represented in the data. Cause for these errors can be found in the deviations during the collection of the point cloud data as described by Van der Zon (2013) as well by the used methodologies to process the point cloud data. Additional, the way the raster data is represented can be qualified as unclear. This research will investigate the possibilities to decrease or even solve such problems in order to develop an improved web-based height map for the Netherlands based on the AHN2 dataset.

2 Objectives, scope and deliverables

2.1 Objectives

The objective of this thesis is to developed a methodology that automatically generates a raster map based on AHN2 point cloud data. As stated in chapter 1, compared to existing products, I think there are potential improvements possible for the production of a better AHN2 raster map. The feasibility of this statement will be researched in this thesis and therefore, some objectives and research questions are defined. The main research question is:

- Is it possible to design and prototype algorithms to automatically generate a digital raster height map for point cloud samples (retrieved by LiDAR techniques) that performs better than existing ones?

The research of the main objective can be divided into the following (sub-) research questions:

- Given raw point cloud data, is filtering needed in order to detect and remove noise and/or outliers?
- Which interpolation technique(s) is/are most appropriate to calculate/estimate a height at a given location (that is assigned to a grid cell) from (LiDAR) raw point cloud data?
- In extent to which is it possible to create a raster height map from raw input data without usage of external 2D geodata (TOP10NL, BGT)?
- The AHN2 dataset can be qualified as a large-scale data set ("big data"), what (automatic) methodology is best to process such large raw point cloud samples (retrieved by LiDAR techniques) and generate new (raster) data out of it for usage in a web-map?
- In what order can the visual representation of raster data help to improve the understanding of a map?

2.2 Scope

Focus will be on quality of the outcome of the algorithm, not on performance of the algorithm itself. Design of an efficient algorithm is something that comes secondary. Quality parameters to test outcome can be resolution, number of holes and presence of dynamic objects (cars, animals) and noise.

2.3 Deliverables

The deliverables of this thesis research will be the following:

- The thesis self;
- A (Python-)script to create automatically the input for a web-map;
- A server running a web-map containing the research results. Final goal is to come to a map for the Dutch province Zuid-Holland.