# Fast semi-automated point cloud cleaning

Rickert Mulder
Supervised by: Patrick Marais

## 1 Introduction

Point clouds are sets of vertices in three dimensional space that typically represent the surface of an object. They are used in a variety of contexts. In the cultural heritage domain, laser range scanners are used to produce accurate point cloud representations of historically significant sites. In conjunction with site photography, multiple scans from different angles can be used to create 3D models for preservation and restoration purposes.

There are a number of steps in the processing pipeline that produce such models [6]. One of the earlier steps is referred to as point cloud cleaning. When scanning a heritage site, unwanted artefacts are likely to be present in the resulting point clouds. Examples can include: objects such as power lines or smudges produced by people walking in between the scanner and target object. The point cloud cleaning process involves the removal of such artefacts from a scan.

This process involves lots of manual labour that tends to be time consuming. Some artefacts are harder to remove than others. A typical scan taken on an expedition, may take a single person anywhere between 30 minutes to 2 hours to clean. Given that one requires 500 - 1000 scans to cover a typical heritage site, this stage of the processing pipeline takes a considerable amount of time [6].

The goal of this project is to develop an open source semi-automated system that utilises point cloud classification and segmentation techniques to accelerate the cleaning process. The project will be undertaken in collaboration with the Zamani group [6].

## 2 Point cloud cleaning

Point cloud cleaning implies the classification of points, grouping points corresponding to noise before removing it. Segmentation can be performed manually by the user or can be fully automated. Between these two extremes there are semi automated approaches that lets the user manually segment the point cloud while automating subtasks to various degrees. More automated approaches save time but can compromise accuracy.

In the cultural heritage domain there is a strong emphasis on preserving detail. Every point is considered valuable information. It is thus very important that noise is correctly classified and segmented. In most industrial applications, the level of detail is reduced by creating CAD models, once the necessary features have been identified. As most point cloud editing packages are not specifically aimed at cultural heritage applications. Intelligent segmentation tools from such packages may exhibit some inaccuracies that can lead to valid points being discarded which is unacceptable when working with heritage data.

## 3 Related work

### 3.1 Existing systems

Simple lasso selection tools are quite common [3, 2, 9] and can be used to remove virtually all types of artefacts. They may however be tedious to use when noise is not easily isolated. Three dimensional selection brushes are another simple class of tools that may give a user more control over which points are removed [3]. Somewhat more intelligent cleaning can be achieved with different families of fill tools. Fill

| | Segmentation feature | | | |
|---|---|---|---|---|
| Package name | Simple | Intelligent | Auto | Open source |
| Terrascan [10] | | | ground points, vegetation, buildings | ✗ |
| Pointools Edit [3] | lasso select rectangle select, ball/cube brush select | fill select with distance threshold, fill with colour or intensity similarity, plane select | | ✗ |
| VR Mesh Studio [11] | | power lines | ground, vegetation, roofs, planes | ✗ |
| Carlson PointCloud [1] | | | clear isolated and duplicate points, extract bare earth | ✗ |
| 3D Reshaper [9] | lasso select | | clustering with distance metric, clustering colours, remove isolated points | ✗ |
| Cyclone [2] | lasso select | fill with smoothness similarity | | ✗ |
| Meshlab [12] | plane select, single point select | | isolated point removal | ✓ |

Table 1: Existing systems

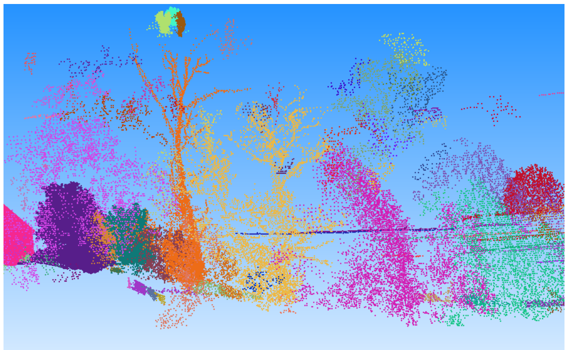

Figure 1: Distance clustering [9]

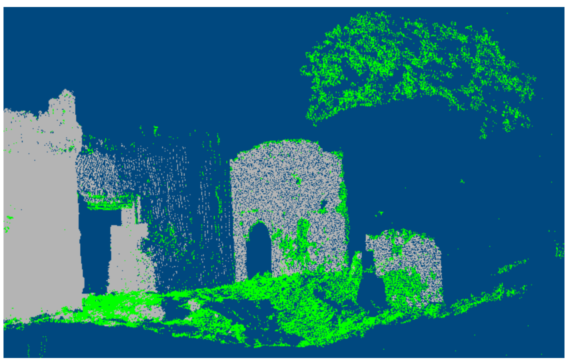

Figure 2: Automated segmentation of vegetation [11]

tools recursively add neighbouring points to a selection based on various metrics such as distance or intensity [3]. Automatically achieving the correct selection is unlikely, therefore systems may allow one to edit a selection before removing points [3].

Global segmentation or classification schemes can remove a lot of the effort associated with point cloud cleaning. In 3DReshaper [9], a point clustering approach is used to isolate distinct objects based on the distance between neighbouring points. However, because the point density of a laser scan usually decreases as points fall away from the scanner, this approach tends to produce large clusters close to the origin, and many small groups of clusters further away. Automated classification schemes take higher dimensional features into account in order to identify clusters corresponding to ground points, vegetation and buildings [10, 11, 1]. While this may be of great for land surveyors or architects, such fully automated methods tend to lack the precision required for cultural heritage purposes.

There is certainly room to reduce the amount of productivity one has to sacrifice for precision, especially in open source tools. Meshlab [12], is currently the only known open source system that provides point cloud cleaning features. It is primarily aimed at processing triangulated meshes and therefore lacks some raw point cloud cleaning features. The recently introduced point editing features, such as point picking and plane selection, performs somewhat sluggish on larger point clouds, and are not very efficient.

### 3.2 Segmentation methods

Automated point cloud segmentation is not a new problem. There is however, very little available research related to the segmentation of cultural heritage data. Some success has been achieved in automatically segmenting sites into surfaces and edges through the use of Principle Component Analysis [8]. What makes the segmentation of heritage data hard, is that scanned structures can exhibit very complex geometries that is hard to classify [8].

A variety of point feature schemes that are used extensively in robotics, could potentially be exploited in the cultural heritage domain. A number of popular point feature algorithms have been implemented and is freely available in the Point Cloud Library [4]. These include Fast Point Feature Histograms, Spin images, Global Fast Point Feature Histogram [4].

Classification can be achieved by though either heuristic schemes [8] or probabilistic models [7, 5]. Approaches based on machine learning have been shown to be effective.

## 4 Research question

La la

## 5 Aims (expand)

Why am I doing it?

Inputs, outputs How many points How large do point set become we only deal with small point sets

In collaboration with geomatics

The aim of this research is to produce a system that will allow users to clean cultural heritage data in a fast and effective manner. The focus will be on finding ways to remove the most problematic artefacts as identified by the Zamani project. In descending order of importance, the focus will thus be on removing vegetation, people, and equipment.

Existing point feature schemes will be investigated, as well as heuristic and machine learning approaches to segmentation. In order to ensure fast response times, cleaning methods will be accelerated with OpenCL.

# 6 Evaluation

Overall:

Performance metrics Total time taken to clean Given cleaning objectives Measure accuracy Perform diff against benchmark scan Compare old vs new Prior training to account for practice effects of leica point cloud Time performance

Each tool Interview : Expert user opinion Sufficient accuracy Confidence information

Remember to try clean out points in scan 1 that was already clean in scan 2.

What about pairwise registration and scanning framework?

Quantitative study (Expert users) Geomatics students Should be experienced

The system will be evaluated in a user study. The speed and accuracy of the existing system will be compared to Zamani's current system [2]. Interviews will also be conducted to assess the usability of the system.

# 7 Expected outcomes

Geomatics very happy

## 7.1 Milestones

| Task | Due date |
|---|---|
| Research Proposal | May 2012 |
| Background Chapter | October 2012 |
| Technical Report | November 2012 |
| System Design Chapter | December 2012 |
| Implementation Chapter | March 2013 |
| Conference Paper | June 2013 |
| First Thesis Draft | July 2013 |
| Final Draft | August 2013 |

# References


[1] Carlson. Carlson Point Cloud, 2012. URL `http://www.carlsonsw.com/PL\_CS\_PointCloud.html`.

[2] Leica. Cyclone, 2012. URL `http://hds.leica-geosystems.com/en/Leica-Cyclone\_6515.htm`.

[3] Pointools. Pointools Edit, 2012. URL `http://www.pointools.com/ptedit\_intro.php`.

[4] Radu Bogdan Rusu and Steve Cousins. 3D is here: Point Cloud Library (PCL). In *Library*, volume 36, pages 1–4. IEEE, 2011. ISBN 9781612843865. doi: 10.1109/ICRA.2011.5980567. URL `http://www.pointclouds.org/assets/pdf/pcl\_icra2011.pdf`.

[5] RB Rusu, Andreas Holzbach, and Nico Blodow. Fast geometric point labeling using conditional random fields. *Intelligent Robots and*, pages 7–12, 2009.

[6] Heinz Rüther, Christoph Held, Roshan Bhurtha, Ralph Schröder, and Stephen Wessels. Challenges in Heritage Documentation with Terrestrial Laser Scanning. 2011.

[7] Roman Shapovalov, Alexander Velizhev, and Olga Barinova. Non-associative Markov Network for 3D Point Cloud Classification. *Photogrammetric Computer Vision and Image Analysis*, XXXVIII(Part 3A):103–108, 2010. ISSN 16821750. URL `http://pcv2010.ign.fr/pdf/partA/shapovalov-pcv2010.pdf.pdf`.

[8] S Spina, K Debattista, and K Bugeja. Point Cloud Segmentation for Cultural Heritage Sites. *Cultural Heritage*, 0(0):1–9, 2011. doi: 10.2312/VAST/VAST11/041-048.

[9] Technodigit. 3DReshaper, 2012. URL `http://www.3dreshaper.com/en1/En\_software.htm`.

[10] Terrasolid. Terrascan, 2012. URL `http://www.terrasolid.fi/en/products/terrascan`.

[11] VirtualGrid. VR Mesh Studio, 2012. URL `http://www.vrmesh.com/default.asp`.

[12] Visual Computing Laboratory. Meshlab, 2012. URL `http://meshlab.sourceforge.net/`.