The art of collecting and disseminating point clouds

M.P. Kodde
Fugro-Inpark B.V., the Netherlands

Abstract

Point clouds can be collected in various ways, but the most commonly operated sensors include Multi Beam Echo Sounding for sampling water depths and Laser Scanners for assembling point clouds onshore. FLI-MAP and DRIVE-MAP are two examples of dynamic scanning systems, operated from a helicopter and vehicle respectively. A major development is the increasing need for combining data sets from various sources. This requires a common data format to store the point clouds as well as auxiliary imagery and metadata. The current way to do this and new upcoming standards are described in this article. To further simplify the exchange of point clouds, a Web Point Cloud Service is described and proposed. This proposed service allows data vendors and users to operate on remotely hosted point clouds.

1. Introduction

A point cloud is a set of individual 3D points that together represent a surface of an object, terrain, seabed or otherwise. The most common instruments to collect point clouds are laser scanning systems on land and acoustic scanning sensors offshore. Many of these sensors are actually used within Fugro, a world wide operating company that uses a multitude of sensors to capture and model the surface, the subsurface and the atmosphere of the earth, both on sea as well as on land. As such, point clouds have grown to great importance within Fugro over the last years. Developments that contributed to this are Fugro's FLI-MAP LiDAR system, the increasing size of terrestrial laser scanning projects and the advent of Mobile Mapping data using Fugro's DRIVE-MAP system. In addition, the need of combining datasets from various sources is growing.

These developments change the way we work with point cloud data. Existing data models, once developed for airborne applications, are no longer the most optimal system for handling point cloud data from other sources. In addition the typical client base for point cloud data sets is shifting from expert users to a more general audience. In this paper, we present some of the challenges the industry is facing with this shift to new point cloud acquisition systems, and the new demands from the end users. It will be shown that current standards need to be reviewed and that new smart methods for point cloud data distribution are needed.

First we will show some methods for point cloud data acquisition. Then, some examples of the combination of point clouds will be shown, including the combination of bathymetry and lidar data. In the subsequent paragraph, it will be shown why current standards for point cloud data models are not sufficient to store the combined data sets. Finally, a proposal is given on how to improve the way point clouds are used by novel users using web services.

2. Point cloud acquisition methods

On land, Fugro collects point clouds by using various sensors. In general, the selection of a sensor is a trade off between accuracy requirements and the area to cover. This trade off is represented in Figure 1, which shows accuracy and coverage as two dimensions.

The results from tachymetry are usually not called point clouds, since the density of the points is extremely sparse. If a denser point cloud is required over a small area, Terrestrial Laser Scan-

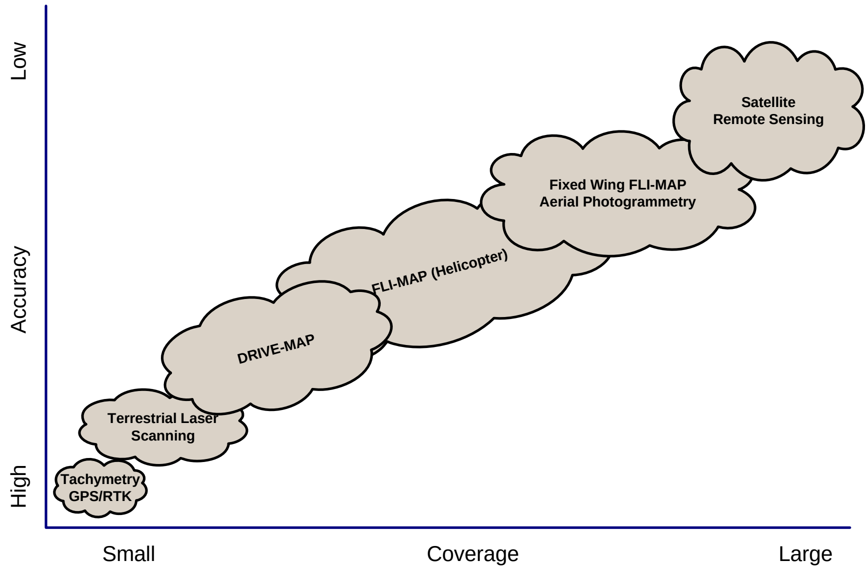

Figure 1. Land based acquisition methods.

ning is the most appropriate method. Terrestrial Laser Scanning is commonly applied in industrial and indoor environments where accuracy requirements are high. Even though a single terrestrial scan may give fewer points compared to dynamic scanning environments, multiple scans are often combined to cover larger areas. Terrestrial Laser scanning can therefore produce very large point clouds as well. This was exemplified by a survey of Fugro on a large industrial site. Up to 150 scans were needed to cover the entire area, resulting in a dataset with more than 2 billion points and a density of 100 x 100 points per m^2.

Larger areas require dynamic scanning systems. Fugro uses its DRIVE-MAP system to cover large highways and urban roads. For DTM generation, the aerial FLI-MAP system is more appropriate (Figure 2). The modern versions of FLI-MAP can be applied from both helicopters and airplanes. Though the carrying system is different, the method of operation for FLI-MAP and DRIVE-MAP are similar in that they both require good GNSS positioning and determination of the attitude using inertial navigation. Both systems acquire a combination of imagery and laser data. The combination of lidar and imagery makes it easier to identify or recognise objects and their attributed information.

Figure 2. FLI-MAP (helicopter based) and DRIVE-MAP.

Near-shore waters and inland waters are typically surveyed with multi beam echo sounding. This method is similar to laser scanning, but measures the run length of acoustic waves instead of laser pulses. The output for both systems, a 3D point cloud, is similar for both systems.

3. Combination of datasets

Often, useful information can be acquired from the combination of datasets. Combined DRIVE-MAP and FLI-MAP data show great advantages in the amount of detail that becomes available for city modelling: FLI-MAP captures the roofs, gardens and the overall DTM, whereas DRIVE-MAP gives a detailed overview of the roads, objects along the roads (such as traffic signs) and the facades of buildings. Figure 3 shows an example of a DRIVE-MAP survey of various buildings. Clearly, a lot of detail is available for the trees, cars and buildings. A combination with FLI-MAP adds information of the roofs and the back of the houses.

Figure 3. DRIVE-MAP data in an urban environment.

Another example of useful combinations of datasets are visualised in the picture below, where a terrestrial laser scan of a bridge is combined with acoustic data of the river below. This dataset, made for a US based bridge, allows engineers to effectively analyse the interface between the bridge and the river currents. An example of this combination is shown in Figure 4.

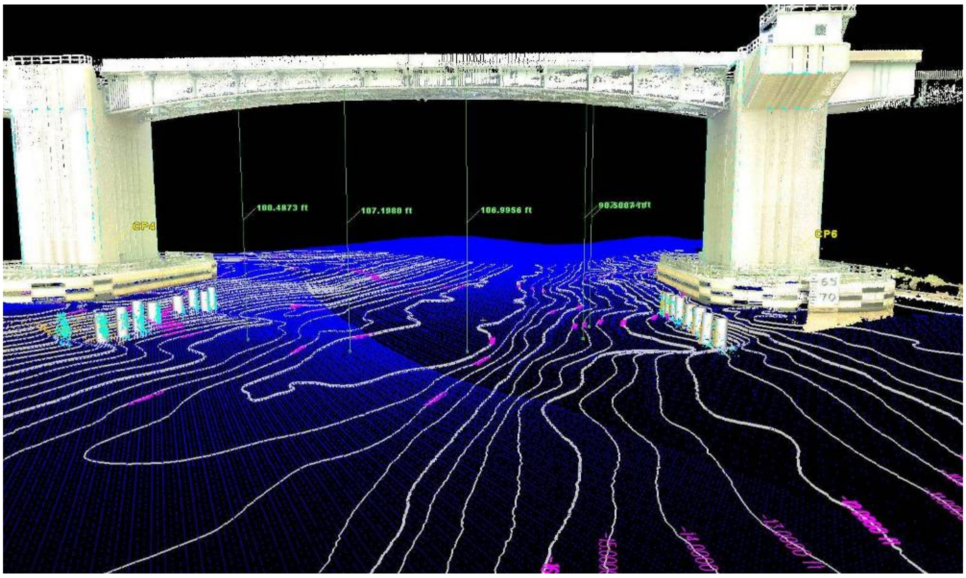

Figure 4. Combination of terrestrial and acoustic data.

4. Point cloud data model

Few data structures could be simpler than a data structure for point clouds. A point cloud is a mere collection of individual points that, together, give a representation of a surface. These points can be stored as a simple list in binary or text format. Therefore, combining multiple datasets is a matter of loading several point clouds at the same time. However, in daily practice, additional information is attributed to these points in order to enhance processing procedures and to store additional information. For the effective combination of datasets, it is important that this information, the data model so to say, remains in tact.

In dynamic laser scanning environments, such as aerial laser scanning or mobile mapping, attribute information is added to each point. For example, each point can be accompanied with the time it was acquired, RGB colour values and classification. Similarly, static scanning applications require knowledge on the scan position for each point and are often attributed with colour as well.

Conventions for adding attributes to point clouds have developed over time, often independently for each type of sensor or even manufacturer. Specifically for airborne laser scanning, a file format was developed to store airborne LiDAR laser points and its attributes. This so-called binary LAS-file, currently maintained by the ASPRS as a standard, allows the attribution of the following properties to each laser point (ASPRS, 2009):

- X, Y, Z,
- Intensity,
- Return number and total number of returns,
- Scan direction, scan angle,
- Edge of Flight line,
- Classification,
- Point source ID,
- User Data,
- R, G, B value.

All mentioned fields are required, except for the intensity value, the user data and the RGB colour value. Most fields are mainly of a technical nature. The values stored in these attributes are directly related to acquisition method of the point cloud: airborne laser scanning. The LAS-standard also gives a list of possible values for the classification field. Each point could be attributed with a certain meaning, to be selected from the list presented in Table 1.

0	Created, never classified
1	Unclassified[1]
2	Ground
3	Low Vegetation
4	Medium Vegetation
5	High Vegetation
6	Building
7	Low Point (noise)
8	Model Key-point (mass point)
9	Water

Table 1. Classification value for lidar points.

Again, these classification values are of a technical nature, derived from the filtering processes of aerial laser scanning and the typical end products.

Since the LAS file is one of the few standards for laser data, the file is commonly used for other types of scanning as well. Most notably, mobile mapping data is regularly stored in LAS files and large terrestrial laser scanning projects can be stored in LAS files as well. However, since the LAS format was not designed for this purpose, there is an inherent loss of data. For instance, Mobile Mapping systems often apply multiple scanners on one vehicle, something that cannot be accommodated in the current LAS data model. In addition, the classification rules need to be made more flexible. Terrestrial and mobile scanning applications have introduced new classes, tunnel infrastructure being the most notable one. The point classification should be seen as the semantic model for laser data. This model may be different for each application: a scan of a tree may require a hierarchical model representing leafs and branches, while a scan of a city may only need a subdivision between “ground” and “not-ground”.

To accommodate these new needs, a new version of the LAS standard, LAS 2.0, was proposed (APSRS, 2007). This new version includes room for additional classification values, such as railroad and powerline. It also provides room to select multiple scan sources and freedom in selection of coordinate datums. However, this proposal did not become an accepted standard and was followed up by the Technical Committee E57 of the ASTM standards organisation. Within the working group for data interoperability a new file format was designed, the ASTM E57 3D file format, or E57 file for short (ASTM, 2009). The file is based upon an XML structure with binary storage of large data volumes such as points and images. The file format adds among many other new features the selection of coordinate systems, point grouping and the combination of 3D point clouds and images. This latter feature, the combination of points and images is very useful, as most modern data acquisitions systems, such as FLI-MAP and DRIVE-MAP, intrinsically combine the acquisition of images and points. Finally, the format contains room for further user defined extensions, to accommodate the specific needs for some sensors and for forward compatibility to new sensors. The E57 file format is not accepted as a standard yet, but it is expected to happen soon.

5. Point cloud webservice

Since computers are getting faster, increasingly the end product of a scanning survey is just the point cloud itself. The common way to use point clouds is and was to derive a data set from the point cloud, e.g. a CAD drawing. Specifically for Terrestrial Laser Scanning, the trend is that modern software allows the end-user to do the modelling or analyses on the fly using specialised software.

Another development is that the usage base of point clouds is growing within organisations. In the past, clients usually were a small group of experts with knowledge of the acquisition technology. New DRIVE-MAP surveys show that the point clouds will increasingly be used by a wider audience, including users with a lot of domain expertise, but little GIS or point cloud processing knowledge. This requires good software to support the user, but also better ways to send, store and maintain the point clouds.

At this moment, USB discs with the point cloud data in a certain format are the most common way to distribute point clouds. However, this method requires large storage facilities at the end user’s IT-infrastructure. A new method for distributing point clouds could resolve this, preferably using web based technologies where the data storage remains at the source. Such a structure would comprise a web based set of standards for distributing, viewing and working with point clouds. This structure basically comprises three components, as shown in Figure 5:

- Point cloud data server,
- Point cloud application server,
- End-user.

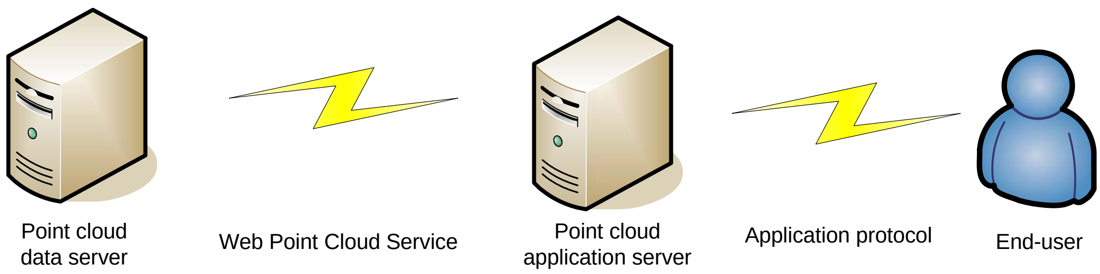

Figure 5. Point cloud distribution structure.

The point cloud data server stores the point cloud data in a standardised data format, such as the E57 file structure described in the previous paragraph. This data should be accessible using, preferably, an open standard. This standard does not exist yet, but in analogy with other OGC standards for distributing 2D GIS maps, it is called Web Point Cloud Service (WPCS) for now. This service provides an interface for:

– Selecting individual points with attributes within a certain 3D bounding box.
– Filtering the point density on the fly, depending on bounding box size and viewing distance.
– Transactional service to update point attributes such as classification, normal vector, etc.

It should be noted that the broadband connection itself is the greatest bottleneck for such a service, since the data volumes are large. This can only be overcome by smart indexing of the point clouds, data compression and the application of Level of Details. Several solutions to indexing and compression exist in the industry, however they are not interoperable.

Similar to the current OGC webservices, the WPCS data streams can be consumed by a program, currently identified as the application server. This program may be a piece of software installed locally on the user's PC, but preferably it is a web based application, allowing the user to do operation from within the browser. There could be a multitude of applications, written in any language such as HTML5/JavaScript, Flash, Silverlight, etc. The only thing they have in common is that they use the same WPCS definitions. Some of these applications, may not even have a user interface, but provide automated processing services only. An example for such a service could be the automated classification of "ground points" in the dataset.

6. Conclusion: point clouds "in the cloud"

This article has shown that various methods for point cloud acquisition exist today. Although these methods vary in there acquisition techniques, they share a very similar end product: a collection of XYZ points. However, the type of information stored alongside these points may be different for each dataset, making the exchange of data between programs and the combination of data sets difficult. It is shown that developments in the standards, including the E57 file format, are providing a solution to this.

Setting standards simplifies the exchange of point clouds, but does not solve the logistical problems of sending point clouds to the end users. In analogy with the current OGC webservices, a Web Point Cloud Service is proposed, to help the dissemination of point clouds over the internet. A smart service should be capable of distributing the data to applications in a bandwidth-efficient manner, thereby opening access to point clouds for many more users. In fact, this correlates with the current developments of "cloud computing", where storage and computing may be scattered through "the cloud", the internet. Bringing point clouds into the cloud may solve many problems and open new perspectives for new point cloud users.

References

American Society for Photogrammetry and Remote Sensing, 2007, LAS 2.0 Format Specification proposal, Madison, AL, US. Retrieved from www.apsrs.org.

American Society for Photogrammetry and Remote Sensing, 2009, LAS 1.3 Format Specification, Maryland, US. Retrieved from www.apsrs.org.

ASTM, 2009, E57 3D File Format Design proposal v2, retrieved from http://www.libe57.org/.