Features of Point Clouds and Functionalities of Processing Software

M.J.P.M. Lemmens
Delft University of Technology,
Faculty of Architecture and the Built Environment, The Netherlands

To be useful, the millions or even billions of 3D points generated by a variety of active and passive sensors need to be stored, organised, combined, geo-referenced, measured, analysed and distributed within organisations or outward. Initially, the data is unorganised; software has been designed, developed and put on the market place to organise the unorganised and to extract information from the point clouds. In principle, the curved 2D surface can represent any instance such as soil pollution, forest biomass, rainfall, terrain elevation or the seabed. In the field of geomatics, the phenomenon will usually be the terrain surface or the seabed. This paper focuses on point clouds from which DEMs or DSMs can be generated, stemming from sources including airborne Lidar; Terrestrial Laser Scanning (TLS); airborne and spaceborne Radar; close-range, airborne and spaceborne imagery; and sonar.This paper presents first the main features of point clouds and focussesnext on the diverse functionalities of point clouds processing software presently on the market, including : data storage; geo-referencing; the filtering aspect of point-cloud creation; interpolation; andvisualisation and editing.

Features of Point Clouds

A point cloud is a set of data points represented in a preferred coordinate system. The dataset consists of measurements taken at discrete points of a curved 2D surface in 3D space. This 2D surface may be smooth or it may contain discontinuities such as facades of building. The number and density of the points should be such that the 2D surface can be reconstructed, i.e. at every point on the surface a value can be calculated from the measured values of points in the vicinity. In principle, the curved 2D surface can represent any instance such as soil pollution, forest biomass, rainfall, terrain elevation or the seabed. In the field of geomatics, the phenomenon will usually be the terrain surface or the seabed in the form of a digital elevation model (DEM) or digital surface model (DSM).

The data of such a point cloud comprises a set of x,y coordinates to which height/depth values have been assigned: one value per x,y location. Added to height or depth, other attributes may be attached to the x,y component, such as reflection intensities of the laser/sonar pulse or RGB of a colour image recorded by a camera at the same time. As a result, the x,y component and its attributes forms the nucleus of the point cloud and the number of nuclei may run into billions (Lemmens, 2014a). Indeed, as a result of rapid advances in laser scanning and dense image matching it is now relatively easy to produce high-resolution terrain data at very high rates. Returns from the active Lidar systems are first stored in the form of range, angle and time. These are then integrated with 3D location data (latitude, longitude and altitude) collected by on-board GNSS and pose data (roll, pitch and heading) measured by an IMU, which is usually integrated with the GNSS, and if applicable compasses, barometers and odometers. The integration of the raw data is largely automated and results in a geo-referenced 3D point cloud which often needs editing to remove outliers and other improper points. Usually the point cloud is transferred to a 2.5D surface in the form of regular grids or triangulated irregular networks (TIN). These products often act as the basis for input in CAD, BIM and other software tools used by planners, designers, inspectors and others.

Point clouds may also stem from overlapping imagery, captured by nadir-looking or oblique cameras, using photogrammetric image-matching software, such as Pix4Dmapper, Timble'sInpho, Imagine from Hexagon Geospatial or Racurs' Photomod (Lemmens, 2011). The cameras may be mounted in manned or unmanned airborne systems (Lemmens, 2014c).

Functionalities

The functionalities of processing software differ widely. To understand types of use, it is key to gain insight into the ins and outs of the different packages. The functionality may start at the creation of the point cloud itself, as is the case for image matching software which creates a DSM from overlapping imagery. We tag the functionality of point cloud processing software into eight groups:

- Point cloud creation
- Geo-referencing
- Data storage & Interoperability
- Visualisation & Editing
- Interpolation
- Feature Extraction
- Taking Measurements
- Analysis

The sequel to this paper will focus on: data storage; geo-referencing; the filtering aspect of point-cloud creation; interpolation; and visualisation and editing.

Data Storage

A first challenge in the use and exchange of point cloud data is the considerable sizes of the files: a 32-bit operating system can run files of up to several GB, and most computers are not designed for such large datasets. The amount of data produced by today's sensors is growing faster than the processing and storage capacity of database management systems (DBMS). The storage of billions of 3D points is an issue in itself as their collection, archiving and distribution is increasingly becoming a task of governmentalorganisations. Manufacturers of hardware have developed proprietary formats for serving the individual surveyors, but these are not well suited for massive volumes distributed from one central provider to multiple users.The same is true for Oracle Spatial and PostGIS, which are the major DBMSs on the market and as a result organisations store point clouds by dividing them into tiles and strips (Kodde, 2014). Storage as ASCII text files may be appropriate but the files may become large and their access slow while their untroubled use requires good communication on record contents and the definition of separation characters. The ASPRS created the LAS format to address the shortcomings of ASCII for Lidar point clouds (Olsen et al., 2013). The binary LAS format is an open standard, requires less storage space and has become ade facto standard.LAZ is a lossless compressed version of the LAS format, resulting in a reduction of a factor 5 to 10 with respect to the original LAS file size. Also the recently released E57 uses binary storage to reduce storage space. The file size is virtually unlimited and E57supports storage of ancillary data of such imagery next to the point cloud data. As software support for E57 is limited up till now, proprietary formats are often preferred as they are optimised for system-specific outputs.

Geo-referencing

3D patches of point clouds – i.e. individual scans – must be transformed to a single coordinate system to enable geo-referencing or registration, which involves identifying common points in the overlaps. As a geometric transformation model, the 3D similarity (Helmert) transformation is usually applied (three translation parameters, three rotation parameters and one scale parameter). This process requires at least seven coordinates, i.e. two 3D ground control points (GCP) and one heightGCP. To obtain redundancy and thus the ability to detect outliers and to compute precision through error propagation, seven or more GCPs that are evenly distributed along the border of the site are commonly measured, usually with differential GNSS or total stations. The stitching of the single point clouds requires tie points in the overlaps. These are distinctive points in the scene or consist of targets placed, marked or painted in the overlaps. The latter requires a more detailed, accurate and thorough planning of the survey to ensure that the marks are placed in the overlaps and be visible from the point of scanning. Added to this, attention shouldbe paid to make sure that the GCPs are clearly identifiable. In addition to the indirect method described above, direct geo-referencing can also be conducted in some cases. For example, a TLS can be centred over a known point and levelled as if it were a total station. The position can also be determinedusing a GNSS antenna mounted onusing a GNSS antenna mounted on top of a TLS.

Filtering

An important aspect of point-cloud creation is the removal of unwanted points (filtering).

Airborne Lidar, TLS,Radar as well as automatic matching of overlapping imagery capture all objects present on the site, including vegetation, cars, bikes, dogs, fish, shipwrecks and suchlike, all of which will, in general,be irrelevant for the survey at hand and are thus 'noise' from the surveyor's point of view. The software should therefore be able to remove such points, either through manual editing or by automatic filtering. Most software contains tools to quickly and easily reduce such noise. Ground filtering is a specific method aimed atthe removal of points reflected from vegetation when creating a 3D city model, or from vegetation and buildings when creating a bare-ground DEM. In course of time many different filtering methods have been developedsee e.g. Filin and Pfeifer (2006), which is a token of the complexity of the filtering problem (Lemmens, 2011a). The resulting algorithms are based on a wide variety of mathematical modelling of the point cloud data. A broad overview of Lidar ground-filtering methods toguide users in selecting the optimal method for their specific applications is given by Meng et al. (2010).

Interpolation

Usually the continuous surface is reconstructed from the sample points by estimating the height value of individual points by using some interpolation technique. Interpolation techniques are also used to transfer an irregular distributed set of elevation data into a regular raster. Interpolation aims thus at providing elevation data in regions where no data exist or to transfer an irregular distributed set of height points into grid format. Many interpolation techniques have been developed in course of time. These interpolation techniques have similar performance, provided that the data behaves well with respect to point distribution and fluctuations in heights. Nevertheless, selection of a proper interpolation method for a given input dataset is a difficult endeavour (Burrough and McDonnell, 1998). Various geo-science related disciplines have developed a variety of interpolation methods for spatial data of specific types. Sample size, sampling design and other characteristics of the data influence the performance of an interpolation method chosen by the user but it is difficult to determine in advance the effects of the characteristics of the data set on the final result (Li and Heap, 2008). However, for terrain heights one may state, as a rule of thumb, that the closer a known point lies to the unknown point, the more similar the behaviour will be. Of course, the underlying smoothness assumption is violated in cases where buildings and other constructions are present. Here, two nearby positions may have very different heights and this property is used for ground filtering that is the removal ofunwanted objects to arrive at a bare ground representation. In case the smoothness assumption is valid the value of an unknown point is usually computed from measured points in the vicinity using a weighting scheme.It is often feasible to use the inverse of the distances as weightings. The inverse distance weighting (IDW) weights the known points in a search area around the unknown point using distance . The search area may be a circle, a square or any other shape. The result is a distanceweightedmean. Nearest neighbour(NN) uses area as the weightingcriterion. IDW and NN both compute values which are within the range of those of the known points and thus do not generate peaks, pits, ridgesor valleys if they are not present in the input. No action from the user is needed and the output is smooth while the values of the known points are preserved. Another often-used method is Kriging whereby the weighting function is not based on distance or area, but rather on the covariance of the measured points.

Ingeneral, the quality of the computed value of the unknown point of any interpolation method mainly depends on whether the points in the vicinity belong to the same type of points. For example, if the measured points are partly located on a building roof and partly on the street, the height of an unknown street point will be computed somewhere between street and roof level. So, a key issue is the robustness of a certain interpolation technique against the changes in the geo-morphological structure within the vicinity of the point of which the elevation has to be determined. It is important that the points used in the interpolation and the point to be interpolated all belong to the same landscapetype or land use type. For example, when we determine the height value of a point, which is located in a valley from two valley points and one slope point, the interpolated height value may be too large and thus erroneous. So, the goodness of an interpolation can be judged by considering how well the method is insensible to the changes in geo-morphological characteristics of the terrain. Locally adaptive gridding techniques enhance the sensitivity of interpolated regular grids to terrain structure, including ridges and streamlines. Although these methods are certainly an improvement, proper interpolation results can only be warranted by performing a pre-processing stage pinpointed on the segmentation of the point cloud in homogeneous sub-areas in terms of geo-morphological structure based on additional information stemming from other sources and inputfrom a human operator or by using region growing techniques well-known in the realm of digital image processing (Lemmens, 1999a; 1999b).

Visualisation and Editing

The software should be able to visualise and edit point clouds with many billions of points at a reasonable rate; it should not take hours before a view appears on the screen. All the points may be depicted in a single colour and size, but a colour code may also portray the strength of the return signal, the elevation or the RGB values of the same point in animage. Inspection of the point cloud requires zooming, panning, rotating, fly-through and adjusting point size. Editing operations include the mouse-clicking of individual points for the removal of outliers or a group of points, addition of missing points or calculation of distances, angles, areas and volumes. The boundary of the group may be identified by drawing a cube, a rubber surface or other volume shape. Selection of a group of points allows geometric primitives such as planes, spheres, cylinders or NURBS (non-uniform rational basis spline) to be fit using an optimisation method – usually this will be least squares. Selection of a group of points also allows themanual removal of unwanted pointssuch as those reflected on vegetationand buildings if one wants to create a bare-ground DEM. Staking out a baseline enables the extraction of a cross section.

Point Cloud Processing Software

Processing software may be general purpose and handle point clouds from a diversity of sensors, or it may be dedicated to specific outputs such as data acquired by terrestrial laser scanners, airborne Lidar, mobile mapping systems, or sonar. Manufacturers of point cloud-generating sensors have recognised that clients need to process the outputs of their sensors, and have complemented their hardware with proprietary software for managing, geo-referencing, visualising, editing and exporting the outputs to dedicated software. Some software builders have spotted potential in offering tools for creating a broad pallet of end products from a particular sensor type such as Lidar, possibly combined with pixel data, from pixel data alone or sonar.

Other packages stem from the other end of the spectrum, i.e. the application domains. For example, constructors who were used to a certain CAD software package started to appreciate TLS point clouds and asked vendors to add modules for processing them. Some manufacturers discovered new opportunities along the way and built dedicated modules on top of one or more base modules aimed at, for example, the mining industry or 3D models of crash sites. This development process is far from complete, and new tools are being added all the time. However, a generic package which can handle all types of sensor output and generate all types of end product does not exist, because each package has its own peculiarities. Therefore, before purchasing software, it is not only important to look at the functionalities of the software, but also advisable to examine its design ideas, any current or planned extensions, its ability to join modules into one workflow, and interoperability with other software and services provided.

Key for specific, industrial use is not only the manual measuring of length, height, distance, angle, area and volume, but also the extraction of geometric features such as lines and planes andthe fitting of solid models through a set of points. The solid models may include cubes, spheres, cylinders or user-specified objects such as valves and elbows.

Concluding Remarks

Vendors often offer plug-in modules, which provide a sequence of functionalities adapted to particular needs, while othersprovide 'lite' versions aimed at users who do not need all the functions. Fernandez et al. (2007) provides a survey of point cloud processing software focussed on Lidar data. Lemmens (2014b) presents an overview of point cloud processing software aimed at creating DEMs or DSMs and products derived from these from a variety of sources including airborne

Lidar; Terrestrial Laser Scanning (TLS); airborne and spaceborne Radar; close-range, airborne and spaceborne imagery; and sonar.

References

- Burrough, P. A., McDonnell, R. A., 1998, Principles of Geographical Information Systems 2nd Edition, Oxford University Press, New York.
- Fernandez, J.C., Singhania, A., Caceres, J., Slatton, K.C., Starek, M, Kumar,R., 2007, An Overview of Lidar Point Cloud Processing Software, GEM Center Report No. Rep_2007-12-001, University of Florida.
- Filin, S., Pfeifer, N., 2006, Segmentation of airborne laser scanning data using a slope adaptive neighbourhood, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 60, no. 2, pp. 71-80.
- Kodde, M., 2014, Future-proof data storage – File and database storage methods for huge point clouds, GIM International, 28: 7, pp. 16-21.
- Lemmens, M.J.P.M, 1999a, Uncertainty in automatically sampled Digital Elevation Models, Chapter 47, in: Lowell, K., Jaton, A., (Edts), Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resourses, Ann Harbor Press, Chelsea.
- Lemmens, M.J.P.M. 1999b: Quality description problems of blindly sampled DEMs. In Shi, W., Goodchild, M.F. and Fisher, P.F., editors, Proceedings of the International Symposium on Spatial Data Quality '99, Hong Kong Polytechnic University, 210–18.
- Michigan,1 ISBN 1-57504-119-7, pp. 399-407.
- Lemmens, M., 2011a, Geo-information – Technologies, Applications and theEnvironment, Springer,ISBN 978-94-007-1666-7.
- Lemmens, M., 2011b, Digital Photogrammetric Workstations Status and Features, GIM International, 25:12, pp. 28-31
- Lemmens, M., 2010a, Airborne Lidar processing software, GIM International, vol. 24, no. 2, pp. 14-15.
- Lemmens, M., 2014a, Point Clouds (1) – The functionalities of processing software, GIM International, 28:6, pp. 16-21.
- Lemmens, M., 2014b, Point Clouds (2) – Survey of processing software, GIM International, 28:7, pp. 27-33.
- Lemmens, M., 2014c, UAS for mapping – A product survey on systems and features, GIM International, UAS Special, pp. 12-15
- Li, J., Heap, A.D., 2008, A Review of Spatial Interpolation Methods for Environmental Scientists, Geoscience Australia Record 2008/23.
- Meng, X., Currit, N., Zhao, K., 2010,Ground filtering algorithms for airborneLidar data: a review of critical issues,Remote Sensing, vol. 2, no. 3, pp.833-860.
- Olsen, M.J., Roe, G.V, Glennie, C., Persi, F., Reedy, M., Hurwitz, D., Williams, K., Tuss, H., Squellati, A., Knodler, M., 2013, Guidelines for the Use of Mobile LIDAR in Transportation Applications, National Academy of Sciences, ISBN 978-0-309-25914-9.