Virtanen, Juho Pekka; Kukko, Antero; Kaartinen, Harri; Jaakkola, Anttoni; Turppa, Tuomas; Hyyppä, Hannu; Hyyppä, Juha

Nationwide point cloud-The future topographic core data

Published in:
ISPRS International Journal of Geo-Information

DOI:
10.3390/ijgi6080243

Published: 01/08/2017

Please cite the original version:
Nationwide Point Cloud—The Future Topographic Core Data

Juho-Pekka Virtanen 1,2,* , Antero Kukko 2 , Harri Kaartinen 2 , Anttoni Jaakkola 2 , Tuomas Turppa 2 , Hannu Hyyppä 1 and Juha Hyyppä 2

1 School of Engineering, Aalto University, P.O. Box 14100, FI-00076 Aalto, Finland; hannu.hyyppa@aalto.fi
2 Finnish Geospatial Research Institute FGI, Geodeetinrinne 2, FI-02430 Masala, Finland;
antero.kukko@nls.fi (A.K.); harri.kaartinen@nls.fi (H.K.); anttoni.jaakkola@nls.fi (A.J.);
tuomas.turppa@nls.fi (T.T.); juha.coelasr@gmail.com (J.H.)
* Correspondence: juho-peka.virtanen@aalto.fi; Tel.: +358-50-405-7791

Academic Editor: Wolfgang Kainz
Received: 7 June 2017; Accepted: 3 August 2017; Published: 8 August 2017

Abstract: Topographic databases maintained by national mapping agencies are currently the most common nationwide data sets in geo-information. The application of laser scanning as source data for surveying is increasing. Along with this development, several analysis methods that utilize dense point clouds have been introduced. We present the concept of producing a dense nationwide point cloud, produced from multiple sensors and containing multispectral information, as the national core data for geo-information. Geo-information products, such as digital terrain and elevation models and 3D building models, are produced automatically from these data. We outline the data acquisition, processing, and application of the point cloud. As a national data set, a dense multispectral point cloud could produce significant cost savings via improved automation in mapping and a reduction of overlapping surveying efforts.

Keywords: point cloud; multispectral; laser scanning; mobile laser scanning; data integration

1. Introduction

Until today, the most extensive and accurate nationwide data sets have been maintained by national mapping agencies. These datasets include aerial photos, raster maps, and vector data. Commonly the vector data is stored in topographic databases (see, e.g., [1]). It includes the details of the terrain and built-up objects, divided into several object classes, in vector format. The positional accuracy of the objects typically ranges from map scales of 1:5000–1:10,000. The topographic database is applied as source material in map production. Part of the information is updated continuously (e.g., roads) or yearly (buildings), with more extensive updating carried out at intervals of 5–10 years [1,2]. Creation and updating of topographic databases involves a large amount of manual work.

The most common input data for production and updating of a topographic database is aerial imagery, but airborne laser scanning (ALS) is increasingly used. Common laser scanning (LS) methods include terrestrial laser scanning (with a stationary instrument) (TLS), mobile laser scanning (MLS) [3], airborne laser scanning (ALS), and scanning from an unmanned aerial vehicle (UAV-LS) [4]. The position and orientation of mobile, airborne, and UAV systems are solved using a combination of Global Navigation Satellite System (GNSS) and the inertial measurement unit (IMU). Recently, multispectral airborne laser data has also been used to provide point clouds with radiometrically calibrated intensities allowing single-sensor mapping solutions [5–8]. In addition to LS, point clouds can be obtained with stereophotogrammetry, depth camera based techniques [9], or synthetic aperture radar (SAR) [10]. Other imaging sensors, such as thermal cameras, can add information to point clouds [11].
Dense point clouds and fully automated processing can be applied to forestry [12], road infrastructure maintenance and monitoring [13,14], city modeling [15], construction [16], fluvial studies [17,18], and autonomous driving [19]. Several countries are applying ALS for statewide elevation modeling; the benefits include 5–10 times better elevation model accuracy (when compared to photogrammetry), highly increased automation in processing, and significantly decreased costs. Methods for automated object reconstruction have been reported for city environments [20], roads [21], and forests [22]. ALS-based forest inventory is operational in at least Scandinavia, the Baltic countries, Spain, Switzerland, the USA, Canada, Australia, and New Zealand. In addition to reconstruction and analysis methods, direct visualization of point clouds [23], annotation of point clouds with semantic data [24], and online storage and retrieval of massive point clouds [25] have been developed. Point clouds have even been suggested to replace existing 3D city models [26].

With the development of measuring technology, applications, and utilization methods, point clouds become source data for an increasing number of applications and processes. Their wider application and acquisition has been suggested by several authors, especially using airborne sensors [27,28]. In this article, we present the concept of producing a dense nationwide point cloud, originating from multiple sensors and containing multispectral information, operating as the core data for geo-information. The surveying activities focus on producing a dense point cloud, which is then applied in further processes. Geo-information products, such as digital terrain and elevation models and 3D building models, are produced automatically as needed.

2. Data Acquisition

As measuring techniques have varying properties in terms of viewpoint (terrestrial, airborne), efficiency, and accuracy, it makes sense to combine them when measuring complex environments [29]. Since roadside data has high importance to each country, the future nationwide point cloud should consist of car-based MLS data integrated with a national coverage of ALS. In the following, we review some relevant developments in measuring systems.

2.1. Mobile Laser Scanning

MLS has proven to be very efficient in measuring road and city environments [30]. After the introduction of multi-platform MLS, the use cases have expanded to natural environments, industrial installations, and urban environments that cannot be easily accessed by car [3]. With the development of algorithms that allow simultaneous localization and mapping (SLAM), MLS has also advanced to indoor environments, where GNSS are not available [31]. UAV-LS permits a very fine scale mapping of both urban and natural environments from a less occluded viewpoint than that of TLS or MLS [4]. In addition to laser scanners, UAVs can carry cameras to obtain data for dense point cloud calculation [32].

The advantage of working with point clouds is that point density can vary depending on the importance of the target, allowing a different level of detail to be extracted from the point cloud data [25]. In urban areas, a higher level of detail can be obtained with UAV-LS or backpack laser scanning [3,33]. For examples of MLS and UAV-LS data the reader is referred to Sections 4.2.1 and 4.2.2.

Autonomous Vehicles for Crowdsourced Mapping

Autonomous-driving technologies have attracted considerable academic and industrial interests in recent years. Following the success of the autonomous car technology competition, the DARPA Grand Challenge [19], many companies have announced their future performance providing automatic or automated driving based on advanced sensor technology. Therefore, most cars in the future will likely carry a mapping system similar to the ones currently installed in high-end MLS cars collecting roadside data. In practice, all future autonomous cars will be equipped with mapping sensors producing point clouds (e.g., lidars, cameras, radars, sonars). See Figure 1. As a result, huge amounts of point cloud data will be acquired from urban and road environments on a continuous basis. Current procedures in mapping are based on defined measuring campaigns, with map updating
left unsolved due to high costs. Continuous flows of car-based data could revolutionize these procedures by providing very interesting alternative solutions for keeping map databases up-to-date. In addition to vehicles, pedestrians having mapping sensors in their smart phones can contribute to mapping [34]. Combined, this technology disruption means that all centralized mapping will be turned into decentralized, distributed mapping. Maps can be frequently updated.

**Figure 1.** An example of the point cloud collected from an urban environment by an autonomous car using a Velodyne VLP-16 sensor (Velodyne LIDAR, San Jose, CA, USA). A wall of a nearby building, a passing vehicle and city trees can be observed from the sample point cloud. By 2030, such data will be collected by several sensors in each car, and such cars are provided with more than 10 M units per year.

2.2. Airborne Multispectral Laser Scanning

LS technology is also developing with the increase of spectral information recorded. The first example of this was the recording of laser backscatter intensity and the use of intensity values in the visualization of point clouds and some classification tasks. The recording of the return waveform has been a significant enabler in forestry applications [35]. The role of laser intensity has been relatively small in automated classification until multispectral LS (e.g., [36,37]), the geometric information of LS having been of higher importance.

The emerging multispectral LS (e.g., Optech Titan for ALS) increases the amount of spectral information obtainable with LS. Earlier intensity studies (for example, [38,39]) showing the concepts for radiometric calibration of ALS intensity have been precursors for multispectral ALS. Briese et al. realized one of the first multispectral point clouds by using three separate ALS systems [40]. Wang et al. used two separate ALS systems to acquire dual-wavelength data and found that the use of dual-wavelength data can substantially improve classification accuracy compared to one-wavelength data [41]. More extensive discussions and reference lists on recent multispectral airborne laser scanning point clouds are available from [5–8,42].

Results with the first multispectral ALS systems have been promising. For example, the overall accuracy of the land cover classification results with six classes (building, tree, asphalt, gravel, rocky, low vegetation) can be achieved at the 96% level compared with validation points [8]. An example of multispectral data from the built environment is given in Figure 2.
Spectral information can also be produced with the means of sensor integration by combining passive imaging sensors to LS [43], but the quality of active multispectral point clouds is significantly better since the errors due to measuring geometry leading to a bidirectional reflectance distribution of passive sensors are overcome with active backscatter data. Thus, the classification accuracy of active multispectral LS data is always higher than with passive techniques using the same wavelengths.

2.3. Airborne Single Photon Systems

There is currently a coming technology breakthrough in the collection of airborne laser data in the form of using single-photon (and Geiger-mode) technology. Single-photon systems require only one detected photon compared to hundreds or even thousands of photons needed in conventional laser pulse time of flight (TOF) or waveform lidars [44]. As a result, the pulse density in a single-photon system can be 10–100 times higher than in conventional LS. Due to the use of a single photon, the systems are easy to make eye-safe and the maximum range is higher than in conventional LS. However, the number of noise points is also increased.

In Degnan (2016) [44], high-range resolution was achieved through sub-nanosecond laser pulse widths, detectors, and timing receivers. Additionally, efficient noise filters, suitable for near real-time imaging, have effectively eliminated the solar background during daytime operations. In Degnan, elevation errors were at the sub-decimeter level [44]. In Tang et al., canopy heights derived using the single photon lidar had a strong agreement with field-measured heights, and automated spatial filtering algorithms can support large-scale, canopy-structure mapping from airborne single photon laser data [45]. In Stoker et al., Harris Corporation’s Geiger-mode IntelliEarth sensor and Sigma Space Corporation’s Single Photon HRQLS sensor were evaluated and compared to large area elevation modeling, and, although not directly applicable for the specifications, were found to possess much potential for increasing the efficiency of data acquisition [46].

3. Data Integration and Processing

3.1. Co-Registration

As not all measuring methods produce a georeferenced point cloud (e.g., indoor slam without GNSS), other methods are needed for co-registration to prevent bias between data sets. Rönholm identified thirteen types of orientation methods developed by the participants for solving general
registration problems [47]. Major strategy types included the extraction of corresponding 3D features from both data sources, the extraction of 3D features from ALS data and the corresponding 2D features from images, or the creation of a synthetic image from LS data and the extraction of corresponding 2D features from both synthetic laser-derived images and aerial images. The types of tie features included points, lines, surfaces, unfiltered laser point clouds, and a combination of lines and surfaces. One of the most promising solutions include the matching of 3D surfaces to each other. Measuring methods that can cover large areas in high detail and produce georeferenced data are a good starting point for producing multi-source data of large areas. The non-georeferenced data sets can be co-registered with this large, georeferenced ‘block’, thus bringing the entire multi-source point cloud to a single, known coordinate system [48]. In addition, it is possible to utilize ALS data for improving the registration of MLS systems, operating in varying GNSS visibility [49].

3.2. Downsampling

In LS, the density of the resulting point cloud is not homogenous, but depends on the operating geometry of the system [50]. Typically, the TLS instruments operate in a fixed angular step, producing a point density that follows the inverse square law, dropping dramatically when the distance to the scanner grows. In MLS, the same effect is visible in the density being reduced when going further away from the system’s path [50]. This can be countered with suitable downsampling methods that aim to produce an even point density in the data. This serves two purposes: firstly, to reduce the amount of redundant data that is typically located right at the foot of the scanner; and secondly, to produce data sets where the variation of density does not interfere with applications, such as modeling [50,51].

In addition, sampling is required for producing a single cloud from a set of overlapping data. In practice, the data has to be analyzed to identify overlapping regions, which then have to be quantified in terms of density and downsampled [51].

3.3. Data Integration

3.3.1. Temporal Information

Point clouds depict the environment as it was at the time of acquisition. Multi-temporal point clouds have been applied for change detection in forestry [52] and fluvial studies [17]. Temporal information can be integrated with the point data [53], allowing temporal filtering of the data (e.g., to obtain points from the latest scans only) or to allow extraction of point cloud pairs for change detection. For example, in some forestry applications, the data acquired in the leaf-on state is significantly more useful than one in the leaf-off state [54].

3.3.2. Accuracy Information

The sensors that are used have an impact on the quality of the point cloud. LS instruments have a limited ranging and angular measurement accuracy, and other error sources also affect the final accuracy [55]. In MLS systems, the accuracy of the platform localization depends on the GNSS visibility and quality of IMU used. GNSS occlusions cause a momentary deterioration in data [3,48]. For some applications, high dimensional accuracy is required. Therefore, the accuracy information should also be included in the point cloud, ideally stated as the global positional accuracy. This would allow the setting of application-specific accuracy demands, the estimation of data reliability, and more intelligent updating of data by omitting less accurate points if better ones are available.

3.3.3. Spectral Information

Point RGB color is commonly used in visualizing large point cloud data sets. However, spectral information present in point clouds, obtained by either sensor integration [37], or by laser backscatter intensity [56,57] and waveform [35] can be applied in analysis and segmentation [57]. In particular, multispectral sensors have attracted interest for improving classification results [41,42].
3.3.4. Semantic Point Clouds

Classification and segmentation of the point cloud is a prerequisite for successfully extracting object parameters or reconstructing a model of an object. The term “semantic point cloud” refers to a cloud where every point is assigned a label describing the object the point is representing (e.g., “building”, “tree”) [58]. Different classification methods can be applied to identify different objects from the point cloud, gradually producing more and more semantic information. Object-specific semantic information can also be included in the point cloud by, for example, adding a national building ID all points belonging to corresponding building, after it has been segmented. This allows binding the point cloud into other data sets containing object specific information, e.g., building characteristics.

By producing a semantic point cloud and segmenting accordingly, object type specific analysis and reconstruction algorithms can be applied. In addition, the semantic information allows filtering points in visualization and analysis applications, for example by omitting vegetation in urban environments.

4. Application of Dense Point Cloud Data

Several applications for dense point clouds are reported in research literature, across several fields. In several cases, point clouds have been used in analysis of the environment: In construction, they facilitate validation of as-built models [59], and construction progress monitoring [60] through integration with planning data. In forestry, the point clouds have been used for estimating foliage structures [56] and gaps [61], or estimating forest inventory parameters [59,62].

In robotics, point clouds can be applied as a 3D map of the environment. They have been applied for path planning for autonomous vehicles [63], or for estimating accessibility in urban environments [64], using the point cloud to identify passable areas.

In a similar manner, point clouds allow localization of imaging sensors [65] if they contain matching key points obtained through a structure-from-motion algorithm. This is useful for creating markerless augmented reality applications in indoor environments [66], where the localization of the device has to be solved from its onboard camera view.

4.1. Visualization of Dense Point Clouds

In addition to analyses and segmentation methods, visualization of point clouds (rather than more processed models) has been developed, with the advanced systems being capable of visualizing data sets of billions of points [23]. Commercial software dedicated to visualizing large point clouds has emerged (e.g., Bentley Pointools [67], Euclideon Geoverse MDM [68]), and open source projects for visualization also exist (e.g., Potree [69]). In addition, tools that allow the user to study the point cloud with CAD software have been developed. De Haan argues that dense point cloud visualization may contain less errors than the results of automated modeling in some cases [70]. Visualization methods that utilize immersive display devices, such as CAVEs, have also been developed [71].

For maintaining and distributing large point clouds, tools that allow the storage of point clouds in a database and spatial queries have to be applied. Oosterom et al. have carried out a benchmark of several hosting systems, including Oracle Spatial and PostgreSQL, using a dense aerial data set of the Netherlands as test material [25].

In addition to dedicated visualization tools, existing game engines may be leveraged. Figure 3 shows a point cloud in the Unity 5 engine. An octree data structure is used to achieve fast point retrieval. Visualization is enhanced by calculating estimated normal information for points. The point cloud shown is produced with photogrammetry, using a combination of UAV and terrestrial images.
4.2. Spatial Information Products

Several geo-information products are currently produced from point cloud data sets. These include building models [72,73], DTM & DSM [74], and individual tree characterization [30]. Several computational methods have been developed for extracting object parameters from dense point clouds and automatically reconstructing the models of objects. These methods are typically specific to object types (e.g., trees [75] or buildings [76]) and are therefore spread across a number of research disciplines. These computational methods can be applied automatically, permitting frequent updating of data sets produced with these methods. In practice, the analysis can be performed synchronously with measuring: every time new data is obtained, the resulting geoinformation data sets are regenerated.

4.2.1. Built and Road Environments

In the built environment, the automated generation of simple building models (e.g., [72,77]) has become mainstream in production of 3D city models [78]. In further research, several algorithms for more detailed building model generation have been introduced, e.g., [20,79]. For a review on urban reconstruction, see [76].

Most cities and main roads will be documented in the future with high-quality mobile point clouds (Figure 4). They allow detailed analysis of the road environment, including road surface analysis. The combination of such data with ALS data sets allows automated 3D modeling and keeping maps updated. Algorithms have been developed for extracting road markings [57], curbs [80], road edges [81], and pole-like objects [82] from MLS data. For dense MLS data sets, methods for automatic identification and segmentation of various urban furniture have been developed [83,84]. Raw point cloud data, however, is valuable for all kinds of engineering applications: direct measurements are possible from raw data. Multispectral ALS data is also feasible for automatic road detection and has a significant improvement compared to the use of optical aerial imagery. In a test using Optech Titan...
multispectral ALS data, 80.5% of the points representing roads were classified correctly. When aerial images were used, the corresponding percentage decreased to 71.6% [7].

**Figure 4.** Example road environment data collected with the Riegl VUX-1HA (RIEGL Laser Measurement Systems GmbH, Horn, Austria) with NovAtel SPAN (UIMU-LCI and Flexpak6) IMU (NovAtel Inc., Calgary, AB, Canada). The point cloud has been colorized with laser backscatter intensity.

### 4.2.2. Forests

Approaches for obtaining forest and forestry data from ALS point cloud data are divided into two groups: area-based approaches (ABA) and individual/single-tree detection approaches (ITD) [30]. ABA prediction of forest variables relies on the statistical dependency between the variables measured in the field and ALS point height metrics, which results in plot- or stand-level information. In ITD, individual trees can be detected and tree-level physical variables (such as height, crown size and, tree species) directly measured, and other variables (such stem volume, biomass and DBH) can be predicted. For inventory purposes, the stand-level forest inventory results are aggregated by summing up the ITD data. For a comparison of various ITD techniques, the reader is referred to [75,85–87]. An accuracy of about 1 m in tree height determination can be achieved. The majority of the trees in the dominant story can be detected.

Multispectral ALS allows discrimination of tree species. In a test by Yu et al., point cloud features described tree species with 76.0% accuracy, where the use of intensity features of all Optech Titan channels resulted in 85.6% accuracy [5]. Isolated and dominant trees can be detected with a detection rate of 91.9% and classified with high overall accuracy of 90.5%. The corresponding detection rate and accuracy are 81.5% and 89.8% for a group of trees, 26.4% and 79.1% for trees next to a bigger tree, and 7.2% and 53.9% for trees under a bigger tree, respectively.

Since diameter cannot be directly measured from ALS, TLS and MLS have been studied to provide diameter information at the tree and plot levels. In Liang et al., the DBH estimation from TLS data resulted in an RMSE of 1.29 cm [88]. In Liang et al., an MLS system was tested, and the RMSE of the DBH estimates was 2.4 cm [89]. The tree stem curve determines the tapering of the stem as a function of the height. The RMSE of the stem curve measurements was 4.7 cm when utilizing single-scan TLS data [62]. In [90], the RMSE of the stem curve estimation of the pine tree was 1.3 cm and 1.8 cm with the multi- and single-scan TLS data, respectively; and the RMSE of the curve measurement of the spruce tree was 0.6 cm for both the multi- and single-scan data.

Utilization of UAV-LS data for forest inventory has not been extensively studied yet. With a mini-UAV laser scanning system utilizing the Velodyne VLP-16, the estimation of diameter at breast height from point cloud metrics showed an accuracy of 2.6 cm, which is comparable to accuracies obtained with terrestrial surveys using MLS, TLS, or photogrammetric point clouds. An example is shown in Figure 5.
4.3. Change Detection from Multi-Temporal Point Clouds

National laser scanning is currently in progress in many countries. In the Netherlands, the whole country has already been surveyed twice. On a smaller scale, multi-temporal Lidar data sets have been applied to change detection [53]. We assume that in the coming years, national laser scanning can be performed every five years, possibly with a multispectral system producing dense point clouds. This allows spectral and geometrical change detection of unseen details at the national level. Figure 6 shows data collected in 1998 and 2000 from a Finnish forest. A powerline has been constructed between these years. Due to construction, two trees have been removed and a couple of branches have been taken away. The growth of tree crowns is also visible. Having dense point clouds with suitable additional data allows documenting of the environment and unseen automated change detection.

Figure 5. UAV-based point cloud from a forest. (The depicted plot is located in WGS84 coordinates 61°11’ N, 25°07’ E).

Figure 6. Example of changes that can be mapped at the national level. Laser scanning data from the year 1998 (a), 2000 with a powerline (b), and corresponding change (c) (two trees and some branches removed).

5. Point Clouds as National Core Data Sets

Out of the available LS methods, ALS is best suited for collecting nationwide point clouds in a repeated manner. Understandably, national laser scanning programs are emerging. Single-photon ALS
systems provide the potential for producing dense point clouds at low cost. Even annual country-level data collections with such a system are feasible after the technology has matured.

For obtaining a comprehensive data set from road and urban areas, MLS can be combined with ALS (Figure 7). UAV-LS and backpack laser scanning can raise the level of detail where needed. Finally, the emerging autonomous vehicles that already employ LS are a potential future data source for decentralized mapping.

For integrating mapping data from different systems, co-registration is required. By labeling the points with temporal and sensor information, filtering by time or accuracy requirements is possible. Once produced, dense point clouds with multi-temporal information facilitate change detection. Downsampling algorithms have to be applied to remove overlapping points and produce a more homogenous point density. With classification and segmentation methods, the point cloud is segmented to individual objects. After this, information concerning these objects can be attached to the point cloud. This data can be drawn from the topographic database, which in turn can be updated with new, automatically produced vector objects. With automated analysis methods, data can be extracted for various applications, such as forestry or surveying of road environments. For example, automated reconstruction methods produce LoD2 building models if needed.

The produced point cloud can either be visualized directly or processed further. With the developed visualization methods, data sets of billions of points can be stored and visualized interactively. Game engines are also applicable for application development. This way, dense-colored point clouds are directly usable for planning and illustration, reducing the need to produce geometric models for visualization.

6. Discussion

6.1. Automation

Applying largely automated processing with point clouds reduces labor costs. For example, the current Topographic Database of Finland is maintained by 150–200 full-time workers at NLS. A dense multispectral point cloud seems to offer automated data processing possibilities in mapping products. Furthermore, the costs savings also come from the reduction in overlapping measuring campaigns. For example, city environments are currently scanned by both city- and state-level actors. If high-quality 3D model products could be developed from the same data sets, parallel mapping by
NLS and cities would end, providing significant savings. In that concept, also the big cities which are now mapping their territories with about 20 points/m² would benefit from the national core data set. The collection of dense point clouds allowing higher automation is significantly easier using single-photon (and Geiger-mode) technologies. ALS data acquisition has also been found more affordable per surface area unit in larger campaigns [28]. It should be reminded that laser scanning data costs are remarkably affected by the scanning area. If national dense point clouds were available as open data, this would also dramatically stimulate development of value-adding services by companies. While some tasks, such as automated building modeling on a low level of detail, have been possible to automate for a considerable amount of time (e.g., [72]), unsolved issues remain in the context of automatically producing geo-information assets from LS data, especially if the detail level is to be raised [91]. In some cases, the objects of interest may also be hidden from airborne and terrestrial surveying methods [92].

It is worth noting that not all topographic objects are derivable from geometric or radiometric data: for example, land and property ownership and rights are geometric, potentially 3D entities that can’t be reconstructed from geometric measurements [93]. Even if the generation and updating of other geo-information assets were automated, the changes to these would still have to be made manually.

6.2. Data Amount and Point Density

Dense point clouds easily become large in terms of data amount. The ALS data set used by Oosterom et al. contained 640 billion points [25]. With the point density of 6–10 points/m², this was sufficient to cover the entire Netherlands, requiring slightly above 11 TBs of storage space. With single-photon systems, the point density is typically at least 10 times higher. Data amounts are not only increased by higher point densities, but also by the amount of data carried with the points. In addition to conventional classification information [94], points may also carry more detailed semantic labels [58], object ids [26], time stamps [53], or sensor information, leading to a further increase in data amount. Having a time series of multiple overlapping point clouds also multiplies the data amount. However, for point densities remaining in tens of points per square meter, the data amounts remain feasible with current technology.

In MLS data sets, the density is of a different magnitude (over 1000 points/m²). If the point density is increased to MLS-grade, storage space requirements in excess of 100s of TBs can be expected. Several authors have discussed challenges of utilizing DB for storage of large point clouds [25,53]. For coping with large data amounts, pre-generated LoDs and geometric segmentation are suggested [25]. Having a point cloud with more point specific information would allow more sophisticated optimization methods, for example, by retrieving terrain points in low LoD from a large area, and building points from nearby regions only. Further, it would even be possible to separate the cloud to different storage servers as per object type, maintaining the MLS-based detailed clouds as smaller regional “blocks”.

For dense point clouds, downsampling is beneficial for reducing data amounts, and for producing a homogenous point density [25,50]. At the same time, lowering the point density limits potential application. Even removing overlapping data may hinder some processes, if the overlaps are used in, for example, co-registration [49]. Generally speaking, the higher the density, the easier the reconstruction of the object automatically with high detail. For classification tasks, a density below 5 points/m² [94] has been found to reduce performance. In single tree detection, increase above 10 points/m² did not significantly improve results. In building modeling, densities of 4–20 points/m² are used in [95]. For automatically identifying building components from MLS data sets [20,82], significantly higher point densities have been used. In [20] the data set consisted 340 M points from an area of 180 by 280 m, resulting in roughly approximated point density of over 6000 points/m². These results are, however, also dependent on the characteristics of the environment [86].

Based on existing research, a point density of 20 points/m² has sufficed in ALS cases, whereas higher point densities are required for more detailed reconstruction from MLS. Comparing to [25], it seems that the current technology is applicable for maintaining nationwide ALS clouds suited for
automatically producing geo-information assets for a topographic DB. For utilization of MLS data, a higher point density and, therefore, a revised storage solution would be required.

6.3. Multi-Sensor Integration

Combining data from different measuring devices and acquisition times is not completely straightforward if a high-quality result is to be obtained. For example, MLS systems typically experience a reduction in positioning accuracy in poor GNSS visibility, such as dense urban environments and forest [3]. Co-registration by data becomes increasingly difficult if the viewpoints have less overlap [96]. Not only can data from several platforms be co-registered [97], but ALS can be used to remedy some positioning accuracy issues of overlapping MLS [49]. In a research setting it is possible to ensure that the environment remains sufficiently unchanged to facilitate co-registration. For larger campaigns, change detection methods are needed prior to co-registration.

Ideally, the integration should be carried out in terms of all data associated with points, not only geometry. While radiometric calibration of ALS can be performed to some extent even with natural targets [98], the radiometric calibration of TLS (and MLS) is more challenging, both due to instrumental effects and highly altering environments and ranges encountered [99]. Further, if imaging sensors are used to produce spectral information, their performance limitations affect the final result [100]. While multispectral systems are already available for ALS, they are still absent from TLS and MLS. As some applications utilize multispectral data [42,101], they cannot utilize the monospectral parts of a multi-sensor data set, the same applying for full waveform Lidar data.

Finally, even the geometric quality assurance of multi-sensor data may become problematic. If, for example, an indoor data set is registered using correspondences with a georeferenced MLS data set, the absolute accuracy of indoor data depends on both the accuracy of the indoor measuring method used and the accuracy of the georeferenced point cloud used. This creates a point cloud with varying accuracy, dependent on the accuracy of other data.

6.4. Emerging Applications

The point cloud allows accurate distance measurements for arbitrary visible objects. In addition, the visualization capability of a dense colored point cloud is significantly higher than for polygon data only. From the application development point of view, the emergence of 3D game engines [102] that utilize mesh models has been a strong driving force for development of other 3D applications as well. For point clouds, some libraries [103] and platforms that allow the storage and maintaining of large point cloud data sets [69] have emerged. If the point clouds are to contain a large amount of additional data, the choice of platforms is even more limited. For GIS data, there are also standardized APIs for requesting and transferring data online, e.g., WMS and WFS [104]. For point clouds, suitable, easily applicable development environments, open tools, and standardized interfaces have not yet fully emerged. These, along with increasing availability of data, can be expected to stimulate application development in the future. Whole new applications would be enabled by the availability of dense, frequently updated point cloud data with large coverage. We would be able to document the impact of climate change on Finnish nature, e.g., using laser-based indicators for global change. We would be able to create intelligent services for road safety, such as warnings about visual barriers and bad road conditions attached to information where our forest animals are living. There are numerous applications based on such country-wide data.

Point clouds can also be applied for localization of imaging sensors [65]. In augmented reality applications, point clouds are already being applied for markerless localization of mobile devices in indoor environments [66]. Having a nationwide point cloud would also enable its use for GNSS-free localization. In a similar manner, the autonomous vehicles commonly utilizing LS and SLAM [105] may also utilize an existing point cloud for path planning [63]. In this respect, dense MLS from road environments is actually becoming an asset for autonomous vehicles in the future.
Secondly, the emergence of national, openly available point cloud data would stimulate application business. When the national multispectral point cloud is given to the public as an open data set, like the current national laser scanning data, companies would be able to create an industry ecosystem achieving significant turnover. Data would also have an impact on individual industry sectors: the data would allow the precision forest concept in Finland, optimizing the cross stumpage earnings and optimization of wood products. The electronic wood trade would be based on accurate species-specific data using multispectral ALS. In the same way, national road point clouds could be made a national infrastructure.

7. Conclusions

Topographic databases maintained by national mapping agencies and based on aerial images are currently the most common nationwide data sets in geo-information. Currently, the maintenance of these includes significant amounts of manual work. However, the application of laser scanning as source data for surveying is increasing. Along with this development, several analysis methods that utilize dense point clouds as source data have been developed. As more applications utilize these dense point clouds, they become more important. We suggest the use of a dense multi-sensor point cloud as the national core data set in geo-information. Upcoming single-photon technology has the most potential as a sensor solution for providing dense point clouds with low unit costs for country-level data acquisition.

Multipath laser scanning should be applied to obtain data from both airborne and terrestrial perspectives. Processing that includes co-registration, integration of temporal data, and segmentation is required to produce a data set applicable to analysis, automated reconstruction, and change detection. As some applications, the dense colored point cloud can also be used directly, without modeling. It should be noticed that for many applications, the visualization capability of a dense colored point cloud is significantly higher than for polygon data only. As a national data set, a dense multispectral point cloud could produce significant cost savings via improved automation in mapping, and a reduction in overlapping surveying efforts by cities and the state. Potentially, a large amount of data in a topographic DB could be produced and updated automatically. As open data, dense, high-quality point clouds possess a significant business potential for improved forestry and road infrastructure maintenance, and they operate as a “platform” for several novel applications.

To realize the concept, further research and development is needed for coping with data amounts in MLS, for determining the needed point densities for more detailed reconstruction tasks, and for ensuring and characterizing the dimensional quality of multi-sensor point clouds. In addition, the development of analysis methods utilizing point clouds further increases their applicability.

Acknowledgments: We are grateful to the Academy of Finland for their support of the Center of Excellence in Laser Scanning Research (CoE-LaSR, 292735 and 307362) and project “COMBAT” (293389). The Finnish Funding Agency for Innovation is acknowledged for support in project “VARPU” (7031/31/2016). Solid Potato Ltd. is acknowledged for Figures 1 and 4. Leena Matikainen is acknowledged for Figure 2.

Author Contributions: Juho-Pekka Virtanen and Juha Hyyppä envisioned and wrote the article. Antero Kukko, Harri Kaartinen, and Anttoni Jaakkola contributed to passages concerning MLS systems, autonomous vehicles, and road and forest environments. Tuomas Turppa contributed to the section on point cloud visualization. All authors participated in finalizing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References


Fritz, A.; Kattenborn, T.; Koch, B. UAV-based photogrammetric point clouds—Tree stem mapping in open stands in comparison to terrestrial laser scanner point clouds. *ISPRS Arch.* 2013, 40, 141–146. [CrossRef]


49. Teo, T.A.; Huang, S.H. Surface-based registration of airborne and terrestrial mobile LiDAR point clouds. Remote Sens. 2014, 6, 12686–12707. [CrossRef]
60. Kim, C.; Son, H.; Kim, C. Automated construction progress measurement using a 4D building information model and 3D data. Autom. Constr. 2013, 31, 75–82. [CrossRef]
74. Kraus, K.; Pfeifer, N. Advanced DTM generation from LIDAR data.
80. Rodríguez-Cuenca, B.; García-Cortés, S.; Ordoñez, C.; Alonso, M.C. Morphological operations to extract urban curbs in 3D MLS point clouds. ISPRS Int. J. Geo-Inf. 2016, 5, 93. [CrossRef]

94. Richter, R.; Behrens, M.; Döllner, J. Object class segmentation of massive 3D point clouds of urban areas using point cloud topology. *Int. J. Remote Sens.* **2013**, *34*, 8408–8424. [CrossRef]


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).