TRAJECTORY-BASED VISUALIZATION OF MMS POINT CLOUDS

G. Takahashi ^{1, 2}*, H. Masuda ¹

¹ Dept. of Mechanical Engineering and Intelligent Systems, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, Japan (genki_takahashi@kk-grp.jp, h.masuda@uec.ac.jp)
² Dept. of R&D, KOKUSAI KOGYO CO., LTD., 2-24-1 Harumi-cho, Fuchu-shi, Tokyo, 183-0057, JAPAN genki_takahashi@kk-grp.jp

Commission II, WG II/4

KEY WORDS: Mobile Mapping System, point clouds, visualization, data construction, labelling, road facilities

ABSTRACT:

MMSs allow us to obtain detailed 3D information around roads. Especially, LiDAR point clouds can be used for map generation and infrastructure management. For practical uses, however, it is necessary to add labels to a part of the points since various objects can be included in the point clouds. Existing automatic classification methods are not completely error-free, and may incorrectly classify objects. Therefore, even though automatic methods are applied to the point clouds, operators have to verify the labels. While operators classify the point clouds manually, selecting 3D points tasks in 3D views are difficult. In this paper, we propose a new point-cloud image based on the trajectories of MMSs. We call our point-cloud image *trajectory-based point-cloud image*. Although the image is distorted because it is generated based on rotation angles of laser scanners, we confirmed that most objects can be recognized from point-cloud images by checking main road facilities. We evaluated how efficient the annotation can be done using our method, and the results show that operators could add annotations to point-cloud images more efficiently.

1. INTRODUCTION

Recently, with the advent of Mobile Mapping System (MMS), dense LiDAR point clouds over a wide range of area are available. An MMS is a vehicle that equips cameras, laser scanners such as LiDAR, a GNSS (Global Navigation Satellite System) receiver, and so on. MMSs allow us to obtain detailed 3D information along roads, which can be used for map generation and infrastructure management.

However, various objects can be included in the point clouds, so that it is necessary to add labels to a part of the points for practical uses. Since manual segmentation and classification are tedious and time-consuming tasks, automatic methods have been intensively studied among the ISPRS community. For example, Yang et al. introduced contextual features to recognize road facilities from MMS data (Yang, 2017). Soilán et al. proposed road markings' segmentation and classification methods using neural networks (Solian, 2017).

Despite these studies, existing methods are not completely errorfree, and may incorrectly classify objects. Therefore, even though automatic methods are applied to classify point clouds, operators have to verify the results in case of commercial uses. Typically, operators find incorrect and missing labels from hundreds of hours of MMS data while watching videos or time-series 2D images driven from MMS data, resulting in a costly and timeconsuming task. In addition, supervised machine learning algorithms require a lot of manually created training data, which is also a time-consuming task.

In this paper, we aim to reduce the operators' task of adding labels to MMS point clouds. Currently, typical manual labelling methods are as follows.

(1) Operators select a region of a 2D image (i.e. camera image) and add a label to the region as shown in Figure 1. Since an MMS captures a huge number of 2D images while driving, operators have to go through many time-series images. Note that laser scanners and cameras are fixed on the MMS vehicle, meaning their relative positions are known. With the relative



Figure 1. Labelling of a camera image



Figure 2. Labelling of 3D point clouds

^{*} Corresponding author

positions and a well-known pin-hole camera model, 3D points (i.e. LiDAR points) can be projected on a 2D image. In other words, it is possible to select 3D points corresponding to the selected region of the 2D image, and labels can be added to the selected points.

(2) Operators display 3D points on the computer screen and select a part of them to add labels as shown in Fig. 2. Since point clouds are too large to display, operators have to subdivide and sequentially process the point clouds.

This manual procedure is not efficient because it restricts the working space of operators to camera views or viewing frustums of 3D CG. Therefore, we discuss a new method for efficiently labelling point clouds obtained by MMSs. Our approach allows operators to efficiently process point clouds using a wide range of visualization of MMS data.

Figure 3 shows an overview of our method. An MMS outputs point clouds, trajectories of the vehicle, and camera images. In our method, a point-cloud image is generated using the point clouds and vehicle trajectories. The point-cloud image consists of 3D points arranged in a lattice manner on a 2D image. This image is useful for adding annotations to MMS data, because all objects are displayed on a single image. In addition, annotations on the point-cloud image can be automatically copied to the point clouds and camera images. Since the point-cloud image and the point clouds are equivalent, annotations on the point-cloud image can be copied to segmented point-clouds. Annotations can also be copied to camera images, because 3D points can be mapped on a camera image using a pin-hole camera model when the relative position between laser scanners and cameras is known.



Figure 3. Our Proposal for MMS data annotation

2. RELATED WORK

Munoz (2008) proposed to classify MMS point clouds by combining local feature descriptors and context based features, reaching 91.66% accuracy. Babahajiani et al. (2014) classified non-ground points into several classes based on local feature descriptors, achieving an accuracy from 72% to 95%. Guan (2014) tried automatic segmentation of road signs by performing

segmentation with multiple thresholds considering the density of MMS point clouds. Yang (2015) extracted objects along roads such as buildings, street lamps, trees, telegraph pillars, traffic signs and cars by generating multi-scale super-voxels based on 3D point clouds, colours, intensities, and their distribution, and applying graph-based segmentation. However, these methods are not error-free, with an accuracy of around 70–90%. Classification errors are not allowed in most practical applications. In practical cases, all of the automatically classified results must be verified by operators before being delivered to customers.

Recently, machine learning algorithms have been surprisingly improved. Yang (2013) classified MMS point clouds by calculating geometric features and applying SVM. Yu (2016) extracted and classified traffic signs using super-voxels of MMS point clouds. However, machine learning approaches are not error-free as well. In addition, supervised machine learning approaches require a lot of training data, which is not easy to collect because training data often has to be manually created. Unfortunately, at present, automatic labelling methods are not sufficient to obtain practical results not requiring operator's work.

When measuring using an MMS, trajectories of the vehicle are recorded as well as point clouds. The trajectories are calculated using GPS and IMU, which is used to calculate the position of laser scanners. Trajectories data consists of a sequence of 3D coordinates. The vehicle trajectories, camera images, and point clouds are synchronized, and they have a common GPS time. Since the relative positions between the vehicle, laser scanners, and cameras are known, data gathered from different devices can be linked to each other. For example, point clouds can be mapped to a camera image so that each point can be coloured using the pixel values corresponding to the projected position.

In this paper, we focus on visualization methods where operators can efficiently and intuitively add labels to point clouds. Our method is based on mapping point clouds to a single 2D image. We call point clouds mapped onto a 2D image a point-cloud image. Bruno (2015) introduced a sensor-space to map 3D points onto a 2D image. They quantized the rotation angle of the laser scanners at each point and determined the pixel position. However, rotation angles are not generally available for commercial MMSs. Kohira (2017) proposed a more general mapping method using GPS time of the point, the rotation frequency and the pulse repetition frequency of laser scanners. Since the rotation frequency and the pulse repetition frequency are the basic specification values of the laser scanners, this method can be applied to general MMSs. However, each row in the image is drifted since parameters for mapping are deviated during measurement as shown in Figure 4.



Figure 4. Conventional point-cloud images



(a) Road view

(b) Feature view

Figure 5. Trajectory-based point-cloud images

In this paper, we propose a method for generating point-cloud images where operators can easily recognize objects. While a lot of automatic labelling methods have been discussed, efficient labelling methods for operators have rarely been discussed so far. We believe that efficient labelling methods are very important in practical applications, such as infrastructure management and map generation. Furthermore, efficient labelling is also useful to create training data for machine learning.

3. TRAJECTORY-BASED POINT-CLOUD IMAGE

3.1 Point-cloud image

In this paper, we propose a new point-cloud image based on the trajectory of an MMS. We call our point-cloud image *trajectory-based point-cloud image*. Figure 4 shows conventional point-cloud images, which are drifting due to a slight deviation of laser scanners frequency and GPS time. Figure 5 shows trajectory-based point-cloud images, in which points are ordered along the trajectories of an MMS.

In the conventional method, point clouds acquired while the laser beam performs one rotation are allocated in a horizontal row of the image. However, we learned that the rotation frequency of laser scanners was not only slightly shifted from the specified value, but also that it deviated a MMS proceeded. We measured the amount of deviation of laser scanners which perform one rotation every 1/100th of a second. Figure 6 shows a histogram of the difference between the specified rotation cycle time and the actual time calculated using point clouds with GPS time. Although the deviation is very small, it causes pixel shifts in conventional point-cloud images.



Figure 6. Deviation of rotational frequency of laser scanners

On the other hand, in the trajectory-based point-cloud image, points are regularly arranged even if the rotation frequency deviates, because point clouds are adjusted using the MMS trajectories. In our method, point clouds can be mapped on an image in two modes. In the first mode, point clouds are mapped so that the trajectories coincide with the vertical center line of the image as shown in Figure 5(a). This mode is suitable for displaying roadways and objects on the road without separating them. In the other mode, point clouds are mapped so that the trajectories coincide with the vertical lines at the left and right edges of the image as shown in Figure 5(b). This mode is suitable for displaying objects located in high positions along a road, such as buildings, poles, trees, and cables.

The trajectory-based point-cloud image is created in two steps. First the reference points are calculated using point clouds and the polyline of trajectories. Then each 3D point is mapped on the

2D image according to the relative position with the reference point. In the followings, this process is described in detail.

3.2 Reference points on the trajectories

Laser beams are emitted from laser scanners while an MMS runs forward. Therefore, the scan-lines become helical as shown in Figure 7. Trajectories of an MMS are represented as a sequence of 3D coordinates, which thus indicates a path of the IMU installed on the MMS.

When the scan-line and the MMS trajectory line are displayed from the top view as shown in Figure 7(b), the scan-line and the trajectory line intersect twice per laser beam revolution, at the upper and bottom sides. We refer to each intersection as reference point, and to the GPS time at that point as reference



Figure 7. An MMS trajectory, a scan-line, and reference points



Figure 8. Intersection between a scan-line and a trajectory line.

reference points or lower ones is required to identify the trajectory, we use lower reference points. This is because upper time. With regard to the reference point, since one of upper reference points cannot be obtained normally when no object exists above the MMS.

We suppose that the scan-line consists of points $\{\mathbf{p}_i\}$, and the MMS trajectories consist of points $\{q_i\}$. They are represented as polylines generated by connecting the points sequentially. We also denote the GPS time of \mathbf{p}_i as t_i , and the coordinates of \mathbf{p}_i as $(p. x_i, p. y_i, p. z_i)$. For calculating reference points from the scanline and the MMS trajectories, their polylines are projected on the x-y plane. The projected polylines are denoted as $\{\overline{\mathbf{p}}_i\}$ and $\{\overline{\mathbf{q}}_i\}$.

The reference points are calculated as crossing points between two polylines $\{\overline{\mathbf{p}}_i\}$ and $\{\overline{\mathbf{q}}_i\}$ as described in Figure 8. We suppose that two line segments $\overline{\mathbf{p}}_s \overline{\mathbf{p}}_{s+1}$ and $\overline{\mathbf{q}}_t \overline{\mathbf{q}}_{t+1}$ intersect at $\overline{\mathbf{r}}$. Then the reference point \mathbf{r} and the reference time T are calculated as:

$$\mathbf{r} = \frac{|\overline{\mathbf{p}}_{s+1} - \overline{\mathbf{r}}|\mathbf{p}_s + |\overline{\mathbf{p}}_s - \overline{\mathbf{r}}|\mathbf{p}_{s+1}}{|\overline{\mathbf{p}}_s - \overline{\mathbf{r}}| + |\overline{\mathbf{p}}_{s+1} - \overline{\mathbf{r}}|}$$
(1)

$$T = \frac{|\overline{\mathbf{p}}_{s+1} - \overline{\mathbf{r}}|t_s + |\overline{\mathbf{p}}_s - \overline{\mathbf{r}}|t_{s+1}}{|\overline{\mathbf{p}}_s - \overline{\mathbf{r}}| + |\overline{\mathbf{p}}_{s+1} - \overline{\mathbf{r}}|}$$
(2)

The reference point **r** where $p.z_s < q.z_t$ is a lower one and where $p. z_s > q. z_t$ is an upper one. In addition, the directions of the scan-lines at the lower and upper references are opposite to each other. We then select the lower reference points to make point-cloud images. When the reference time of the lower reference points are obtained, they are sorted in ascending order as $\{T_i\}$.

3.3 Trajectory-based point-cloud image

ι

Suppose that point \mathbf{p} with GPS time t is mapped to the coordinates (u, v) of the image. In our method, points can be mapped in the road view and the feature view as shown in Figure 5.

For the road view point-cloud image, points are mapped based on the fact that the reference points are on the vertical center line of the image. First, we select the reference time T_k that satisfies:

$$\frac{T_{k-1} + T_k}{2} \le t < \frac{T_k + T_{k+1}}{2} \tag{3}$$

Then the pixel coordinates (u, v) of the image are calculated as:

$$u = INT\left(\frac{2t - (T_{k-1} + T_k)}{T_{k+1} - T_{k-1}} \cdot W\right)$$

$$v = k$$
(4)

where INT(x) is a function that returns the integer part of x, and W is the width of the point-cloud image. The value of W can be determined using the number of laser beam shot per laser beam cvcle.

For the point-cloud image in the road view, the reference points are mapped onto the right and left edges of the image. Thus the following reference time T_k is selected.

$$T_k \le t < T_{k+1} \tag{5}$$

Then the pixel coordinates (u, v) are calculated as:

$$u = INT\left(\frac{t - T_k}{T_{k+1} - T_k} \cdot W\right)$$

$$v = k$$
 (6)

4. FEATURES OF POINT-CLOUD IMAGE

The point-cloud image is distorted because it is generated based on rotation angles of the laser scanners. In our research objective, it is required that operators can correctly recognize objects in point-cloud images and add labels to them. In this section, we evaluate 2D views of various objects on the point-cloud images qualitatively.

Example images are shown in Figure 9 to 18. In each figure, a camera image is shown on the left, and a point-cloud image is on the right. Camera images are taken from a front camera on the MMS. Some camera images are rotated or flipped to be compared with the point-cloud images.

Figure 9 - 11 show pole-like objects such as street lights, traffic signs, and trees. Although tall objects tend to be bent, they can easily be identified from the point-cloud images. Figure 12 and 13 show a guard pipe and a guardrail. They can also be clearly identified. Figure 14 and 15 show road markings. Although the point-cloud images are generated only using the intensity of point clouds, road markings can be recognized. In these figures, the boundary of the roadway and the walkway are recognizable. However, it is difficult to recognize manhole covers on the road, because points of the manhole covers are missing. In Figure 16 and 17, signals are shown. Although points of signal lights are missing, signals can be identified by the shape of signal frames.

From these examples, we confirmed that most objects can be recognized from the point-cloud images, but some objects such as objects emitting light or not reflecting light may be difficult to be identified.



Figure 9. Street light



Figure 10. Traffic sign



Figure 11. Roadside tree



Figure 12. Guard pipe





Figure 13. Guardrail







Figure 15. Road marking (crosswork sign) and manhole covers



Figure 16. Traffic signal



Figure 17. Signal for pedestrians

5. EVALUATION

5.1 Conditions of experiments

We evaluate how efficiently the annotation is added using our method. An operator adds annotations to (1) 2D camera images, (2) 3D point clouds, and (3) trajectory-based point-cloud images. Then we compared the operation time for each task.

Point clouds were captured using an MMS in Inagi-shi, Tokyo. The MMS is an MX-8 developed by Trimble. This MMS has two VQ-250 LiDAR scanners on the right and left of the vehicle. Camera images were taken with a camera mounted at the front of the MMS.

5.2 Quantitative evaluation

Table 1 shows the time for adding annotations by the operator. In this evaluation, the operator searched street lights and road signs from camera images, point clouds, and point-cloud images, and placed annotations at detected positions.

Camera image	3D point cloud	Point-cloud image
11m 27s	11m 4s	6m 51s
Table 1. Annotation time by the operator		

As shown in this table, the operator could add annotations to point-cloud images more efficiently. In the case of camera images, the operator had to identify the same object in multiple images and add an annotation to each image, because the same object appeared in multiple images, which was time-consuming. In the case of point-clouds, the operator sometimes added annotations to incorrect positions, which is because the operator did not notice that he selected the background points instead of the one they wanted. It was difficult for the operator to recognize the depth on the 2D screen. On the other hand, the point-cloud image was useful for the operator, because the data is 2D (without depth), and each object appears only once on the point-cloud image.

5.3 Qualitative evaluation

For qualitative evaluation, we interviewed the operator and obtained the following comments.

• In camera images, the same object appears in many different images. It was necessary to confirm whether each object was annotated or not. This confirmation was very stressful. In the point-cloud image, confirmation was very easy, because all objects were displayed in a single image.

- In the point cloud, objects were often occluded, so that the 3D scene had to be rotated to detect the target object. It was difficult to identify the positional relationship among objects. In some cases, the positional relationship was only identified by moving objects.
- In the point-cloud image, objects of the same type tended to be displayed in the neighborhood. This was helpful to find the target object.

From the above comments, it is also shown that our method is useful for annotating MMS data.

6. CONCLUSIONS

In this paper, we discussed visualization method for efficient labelling of MMS point clouds. The conclusions are as follows:

- We propose a new point-cloud image based on the trajectories of an MMS; trajectory-based point-cloud image. In our method, point clouds can be mapped on an image in two modes. Therefore, we can choose the mode depending on the target facilities of a road side.

- The point-cloud image is distorted because it is generated based on the rotation angles of laser scanners. We investigated 2D views of various objects on the point-cloud images and confirmed that most objects can be recognized from the pointcloud images.

- We evaluated how efficiently the annotation input was made using our method. In the result, an operator could add annotations to point-cloud images more efficiently.

REFERENCES

Babahajiani, P., Fan, L., Gabbouj, M., 2014. Object recognition in 3D point cloud of urban street scene, *Computer Vision-ACCV* 2014 Workshops, pp. 177–190.

Bisheng Yang, Zhen Dong, 2013, A shape-based segmentation method for mobile laser scanning point clouds, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 81, pp.19-30.

Bisheng Yang, Zhen Dong, Gang Zhao, Wenxia Dai, 2015, Hierarchical extraction of urban objects from mobile laser scanning data, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 99, pp.45-57.

Bisheng Yang, Zhen Dong, Yuan Liu, Fuxun Liang, Yongjun Wang, 2017, Computing multiple aggregation levels and contextual features for road facilities recognition using mobile laser scanning data, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 126, pp.180-194.

Bruno Vallet, M. Bredif, A. Serna, B. Marcotegui, and N. Paparoditis, 2015. Terramobilita/iqmulus urban point cloud analysis benchmark, *Computers & Graphics* 49, pp. 126–133.

Daniel Munoz, Nicolas Vandapel, and Martial Hebert, 2008, Directional associative markov network for 3-D point cloud classification, *Proceedings of 3DPVT'08 - the Fourth International Symposium on 3D Data Processing, Visualization and Transmission.* Haiyan Guan, Jonathan Li, Yongtao Yu, Cheng Wang, Bisheng Yang, 2014, Using mobile laser scanning data for automated extraction of road markings, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 87, pp.93-107.

K. Kohira, H. Masuda, Point-cloud compression for vehiclebased mobile mapping systems using portable network graphics, 2017, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume IV-2/W4, pp.99-106.

Mario Soilán, Belén Riveiro, Joaquín Martínez-Sánchez, Pedro Arias, Segmentation and classification of road markings using MLS data, 2017, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 123, pp.94-103.

Yongtao Yu, Jonathan Li, Chenglu Wen, Haiyan Guan, Cheng Wang, 2016, Bag-of-visual-phrases and hierarchical deep models for traffic sign detection and recognition in mobile laser scanning data, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 113, pp.106-123.