LINE SEGMENT-BASED APPROACH FOR ACCURACY ASSESSMENT OF MLS POINT CLOUDS IN URBAN AREAS

Martyna Poreba[1,2] and François Goulette[1]
[1] CAOR – Centre de Robotique, Mathématiques et Systèmes, MINES ParisTech
60 Bd Saint Michel, 75272 Paris Cedex 06, France;
E-mail: (martyna.poreba, francois.goulette)@mines-paristech.fr

[2] Dept. of Geoinformation, Photogrammetry and Environmental Remote Sensing,
AGH University of Science and Technology
Al. A. Mickiewicza 30, 30059 Cracow, Poland;

KEYWORDS: Accuracy, Mobile, Point Cloud, Edge, Reference Data

ABSTRACT: This paper presents an accuracy assessment of Mobile Laser Scanning (MLS) point clouds and the initial results of practical experiments carried out for a custom-built mobile system**.** To minimise the difficulties of identification and precise measurement of control points, an approach based on line segments is proposed. The main aim of this technique is to compare a series of 3D line segments extracted from MLS data with those from a reference data set. The lines represent edges found as intersections between the principal planes previously modelled within the MLS point cloud. In order to calculate the modified Hausdorff distance, a correspondence-finding algorithm automatically locates similar line segments inside two different data sets: Test and Model. The results achieved confirm that the proposed method appears to be effective for accuracy assessment, especially when evaluating low density point clouds. The experiment site for data acquisition was located in an urban area rich in tall buildings and narrow streets. An estimation of the precision of the mobile system was made using repeated data collection passes, indicating a measurement error value of about 0.3 m. The absolute accuracy was evaluated by comparison with a reference point cloud collected by a static laser scanning survey. The assessment indicates an average difference of less than 0.7 m along two-way street with relatively unobstructed view of the sky.

1. INTRODUCTION

The need for precise spatial data in various applications is continually increasing, complemented by the growing number of mobile laser scanning systems (MLS) available on the market. This technology is a multi-sensor system that integrates various navigational, laser scanning, and other data acquisition sensors mounted on a rigid, moving platform for acquiring data. The point clouds collected by such systems are susceptible to local and global deformations. To ensure a high level of confidence in mobile laser scanning data for an end-user, a validation of the measurements made by the system itself is necessary. It is possible to define several criteria which provide information about the quality of the MLS data such as resolution, accuracy, precision, repeatability or completeness.

1.1 Related Work

A number of methods have been employed in order to verify the accuracy of measured 3D point clouds. The task is challenging, especially in urban canyons, as the accuracy of the geo-referenced point cloud is highly dependent on the GNSS (Global Navigation Satellite Systems) visibility during data acquisition. Buildings, trees and other structures often cause disturbances in satellite visibility and hence a temporary outage in GNSS signal reception. A common approach to accuracy assessment is to compare the measurements with reference values possessing a precision of at least an order of magnitude better than the data being assessed. For this purpose a pair of data sets is used, one collected from MLS and the other a reference set from conventional surveying techniques, a previously geo-referenced terrestrial laser scanner (TLS) point cloud, an orthophoto, or an existing city model. Nevertheless, the majority of these methods compare check points with the corresponding points in the point cloud, chosen such as to be readily distinguishable in the scene, e.g. the ends of white line markings on the road (Gandofi *et al.*, 2008 and Barber *et al.*, 2008), tips of architectural details (Poreba and Goulette, 2012a), poles, building and curb corners (Kaartinen *et al.*, 2012), or signs in the form of newly designed 3D targets points (Feng *et al.*, 2008). This comparison approach suffers the uncertainty of whether the point selected in the point cloud does actually correspond to the check point in reference set. The ability to select appropriate points decreases for lower point cloud resolutions. However, as indicated in (Barber *et al.*, 2008), elevation accuracy is typically easier to determine than the horizontal one. In this vein (Ray and Graham, 2008) discuss a method to verify horizontal accuracy of LIDAR data by using extracted feature lines such as pavement markings which are then adjusted to an orthophoto. The elevation errors may be computed with the aid of a filtered digital elevation model (Kaartinen *et al.*, 2012). Meanwhile (Haala *et al.*, 2008) report the use of planar surfaces selected semi-automatically from an available 3D city model in order to assess the accuracy of MLS data. In the same context (Poreba and Goulette, 2012a) propose to perform a geometric matching via the ICP (Iterative Closest Point) algorithm of the MLS point cloud against the corresponding TLS high-resolution data. (Mano *et al.*, 2012) suggest performing the accuracy assessment by applying the Least Squares 3D (LS3D) surface matching method to overlapping shapes derived from two point clouds.

In this paper we propose an accuracy assessment approach based on line segments. Our method minimises the difficulties of identification of a specific control point. It also allows to measure the quality of MLS data in terms of absolute accuracies as well as to examine the repeatability of the mobile system being used. We propose to use the Hausdorff metric to measure the similarity of two sets of edges.

2. METHODOLOGY

The main aim of the method is to compare series of 3D line segments extracted from MLS data (Test set) with those from a reference data set (Model set) collected using a Total Station, extracted from TLS surveys or alternatively obtained from an existing Topographic Data Base. Line segment correspondence is found automatically by examining the similarity of features. The task becomes more challenging under the assumption that the number of line segments and their length in both data sets is varying. Moreover, some of the feature lines from one set do not correspond to any geometric entity from the other one (Figure 1).

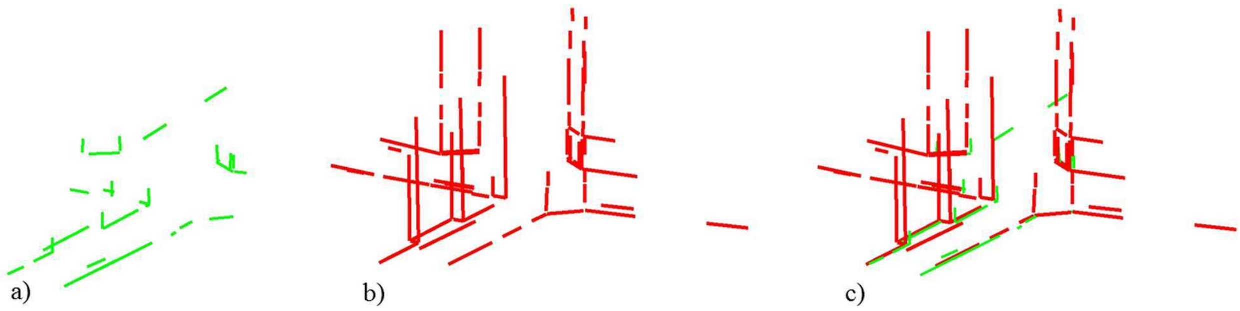

Figure 1 Sample sets of line segments to compare: a) line segments from Test database; b) line segments from Model database; c) Two data sets overlapping each other

The most common method to find the correspondence and reciprocal offset between elements in two data sets is to compute the minimal Euclidean distances between them. Unfortunately, this approach is not robust enough for our applications as it neither takes into account the orientation of lines nor distinguishes the partially overlapping line segments. Thus, a modified LHD (Line segment Hausdorff Distance) technique, proposed by (Gao and Leung, 2002) and complemented by (Chen et al., 2003) originally employed for matching two sets of 2D line segments generated from logos, was adopted to compare 3D line segments. First, the distance between each pair of edges in the two data sets is used to build a similarity matrix. Next, the distances contained within this matrix are used to produce a binary correspondence matrix by searching for values smaller than a prescribed threshold. Once the segments have been assigned, the LHD may be computed in order to measure the absolute accuracy of the acquired point cloud or the reproducibility of the MLS data. The resulting values represent not only the accuracy but also the error encountered during the procedure of edge extraction.

2.1 Hausdorff Distance

The Hausdorff distance is a shape comparison metric, one of the dissimilarity measures frequently used by Geometric Pattern Matching algorithms. It is a max-min distance that has been employed in computer vision for object recognition, mostly in the 2D case. The main goal of the Hausdorff metric is to determinate how much two given objects resemble each other. We denote a set by an uppercase letter and an individual element by a lowercase letter. We compare the set $T=\{t_1, t_2 \dots t_p\}$ representing a Test database with $M =\{m_1, m_2 \dots m_q\}$ representing a Model database. The Oriented Hausdorff Distance (OHD) from T to M is defined as:

$$\mathrm{OHD}(T, M) = \max_{t \in T} \min_{m \in M} d(t, m) \tag{1}$$

where d(t,m) is a particular metric, most often the Euclidean distance. Finally the symmetric Hausdorff (HD) distance is calculated as the maximum of two directed distances, (Eq.2), which in fact expresses the degree of mismatch between the two compared objects.

$$\mathrm{HD}(T, M) = \max(\mathrm{OHD}(T, M), \mathrm{OHD}(M, T)) \tag{2}$$

2.2 Similarity matrix

When the Hausdorff metric is considered, a suitable definition of the d(t, m) referred as the distance between any element of sets T and M must be introduced. Therefore, in keeping with (Gao and Leung, 2002), we propose to build the vector $\vec{d}(t_i^l, m_j^l)$ which represents the distance between two line segments consisting of angle distance $d\theta(t_i^l, m_j^l)$, parallel distance $dII(t_i^l, m_j^l)$ and perpendicular distance $d^{\perp}(t_i^l, m_j^l)$. Such a definition uses the following additional attributes: line orientation and line-point association. Figure 2 shows all the distance types defined to measure the dissimilarity of 3D line segment pairs.

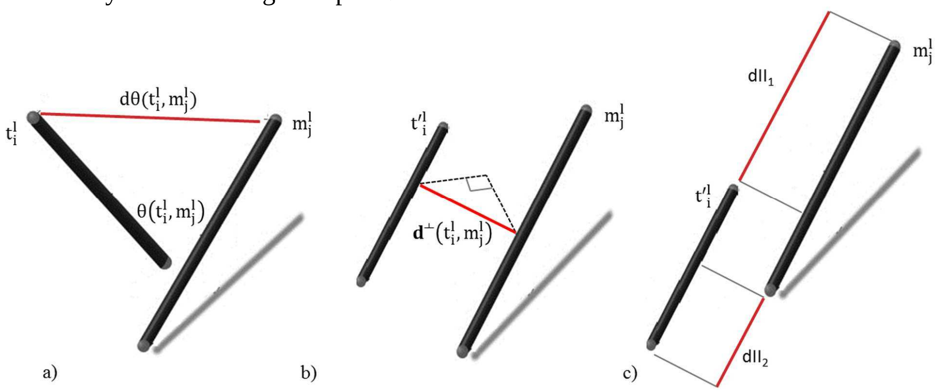

Figure 2 Line displacement measures: a) Angle distance; b) Perpendicular distance; c) Parallel distance

The angle distance is defined according to the equation proposed by (Chen *et al.*, 2003) as:

$$d\theta\left(t_i^l, m_j^l\right) = \min(\| L_{t_i} \|, \| L_{m_j} \|) \cdot \sin(\theta(t_i^l, m_j^l)) \tag{3}$$

where $\sin\theta(t_i^l, m_j^l)$ is the sinus of the angle between line directions; it is multiplied by the length of the shorter line segment. Both parallel and perpendicular distances are built to quantify the translation needed to align two line segments. Thus, the parallel distance is the minimum displacement required to align either the left end points or the right end points of the line segments (Eq. 4). The parallel shifts dII_1 and dII_2 are set to zero if one line is within the range of the other. The perpendicular distance $d^{\perp}$ is simply the minimum distance between two parallel lines.

$$dII\left(t'^l_i, m_j^l\right) = \min(dII_1, dII_2) \tag{4}$$

In fact, none of the pairs of lines are exactly parallel. Every line in the Test data set has its own rotation where the midpoint is used as the center of rotation. Thus, it is necessary to get the correct orientation before computing parallel and perpendicular distances. Finally, the distance between any pair of line segments is:

$$d\left(t_i^l, m_j^l\right) = \sqrt{W \cdot \left(d\theta(t_i^l, m_j^l)\right)^2 + \left(dII(t_i^l, m_j^l)\right)^2 + \left(d^{\perp}(t_i^l, m_j^l)\right)^2} \tag{5}$$

where W is a non-dimensional weight for angle distance, chosen experimentally.
On this basis, the similarity matrix of dimension *p* by *q* for each possible pair of line segments among the two sets is calculated (Figure 3a).

2.3 Correspondence matrix

A correspondence between two sets of lines is determined through the analysis of the similarity matrix. However, such computation does not necessarily produce a one-to-one correspondence because of the incompleteness of linear feature extraction, the observed over-segmentation problem or the inconsistency of the direction and length of line segments. The matching of candidate lines may in fact result as a one-to-null, one-to-one, one-to-multiple, multiple-to-one and even multiple-to-multiple type of correspondence, simultaneously.
To locate the corresponding line segments between sets, a threshold value must be found. For this purpose we propose to link each Test line with the nearest corresponding line from the Model set. Obviously, for every line segment we can find its most similar feature in the other set, even if their resulting distance is quite significant. Starting from this initial correspondence, a threshold value is used to find and reject outlier pairs. For this purpose, the distance values linked to the current pairs are sorted in ascending order and their successive differences are computed.

This operation is repeated for newly created list of values. Subsequently, a function for finding the local maxima of a data set is applied to the results of the previous step as illustrated in Figure 3b (magenta plot). Assuming an “a priori sensibility” for the search function, the first detected peak value denotes the threshold value (δ). Using this parameter a correspondence matrix (Cor) is created as:

$$\mathrm{Cor}_{ij} = \left\{ (i, j) \in \aleph^2 : \forall (t_i^l \in T, m_j^l \in M), d(t_i^l, m_j^l) \leq \delta \right\} \tag{6}$$

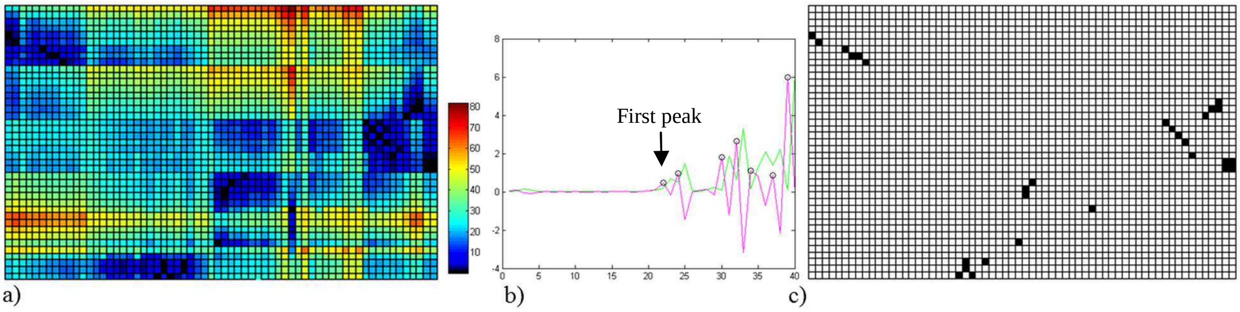

Figure 3 Modified LHD computations steps: a) Similarity matrix; b) Thresholding – searching for the local maxima; c) Binary correspondence matrix

The performance and robustness of this approach was tested using a ground-truth. The results are discussed in the next section.

2.4 Hausdorff Distance between Sets of 3D Segments

The classical definition of a Hausdorff distance is not adequate for comparing objects represented as sets of line segments. Thus, the definition introduced by (Gao and Leung, 2002) was adapted to be useful for measuring the accuracy of 3D straight line segments originating from different sources. It is based on a modified Line segments Hausdorff Distance (LHD), calculated, in our case, solely between corresponding line features in two sets. The details of the correspondence matrix (Cor) search step were described in the previous section. The Oriented modified LHD (OLHD) is defined as:

$$\mathrm{OLHD}\,(T, M) = \frac{1}{\sum_{m_j \in M} L_{m_j}} \sum_{(i,j) \in \mathrm{Cor}} L_{m_j} d(t_i^l, m_j^l) \tag{7}$$

where L_{mj} is the length of line segments of set M, ordered by (Cor). By taking the weighted average of the line segment distances, this definition decreases the impact of outliers, making it more robust. Since longer lines are more reliable, the dissimilarity created by a longer line should more significantly affect the final result, the measure of accuracy. The definition of symmetric modified LHD from OLHD remains the same as in Equation 2.

3. EXPERIMENTAL RESULTS

Our study focuses on two aspects: determining absolute accuracy and assessing repeatability (precision) of MLS point cloud measurements. To minimise problems with recognition and identification of targets in the MLS data, all cases employed edges corresponding to intersections between principal planes (representing walls and roads) modelled in both the MLS and the reference point clouds (Poreba and Goulette, 2012b). In order to check the repeatability of the MLS system, overlapping point clouds acquired from passes at different times and with varying GNSS satellite visibilities were examined. Numerous permutations such as Course1-Course2, Course2-Course3 etc. were compared. Meanwhile, the absolute accuracy assessment determined how accurately the point cloud was positioned in a ground-fixed coordinate system.

3.1 Test Site Setup

For the purpose of this study the Stéréopolis II mobile system operated by the French National Geographic Institute (IGN) was chosen as the main source of data. The system configuration used in our survey is described in (Paparoditis *et al.*, 2012). Data acquisitions were made on February 28, 2012. The test site chosen was an area approximately 130m x 120m in size, located in Paris near Saint-Sulpice Square. This area can be characterized by compact, high developments, tight urban canyons created by tall buildings and narrow, mostly one-way streets where sky visibility was limited. Consequently, the test area was considered as a very difficult acquisition site as it is affected by poor

GNSS receiving conditions. In order to provide an accurate evaluation of the precision of the scanning system, data sets from repeated MLS passes were compared. Figure 4a shows the acquisition trajectories carried out at the test site. A total of 3 independent passes of the Stéréopolis II system were examined. For the purpose of absolute accuracy assessment, a 3D point cloud was collected from a stationary Leica ScanStation C10 scanner from three different positions. The test area scanned this way is denoted in Figure 4a by the red dashed circle. The final reference point cloud density was set to 4mm. The data was geo-referenced to the National Survey Grid (Lambert 93) via two GNNS stations working in static mode during the TLS acquisitions. The absolute accuracy of the reference point cloud depends on the GNNS positioning. In this case the accuracy is estimated to be of about 3cm, which provides a reliable reference (Figure 4c).

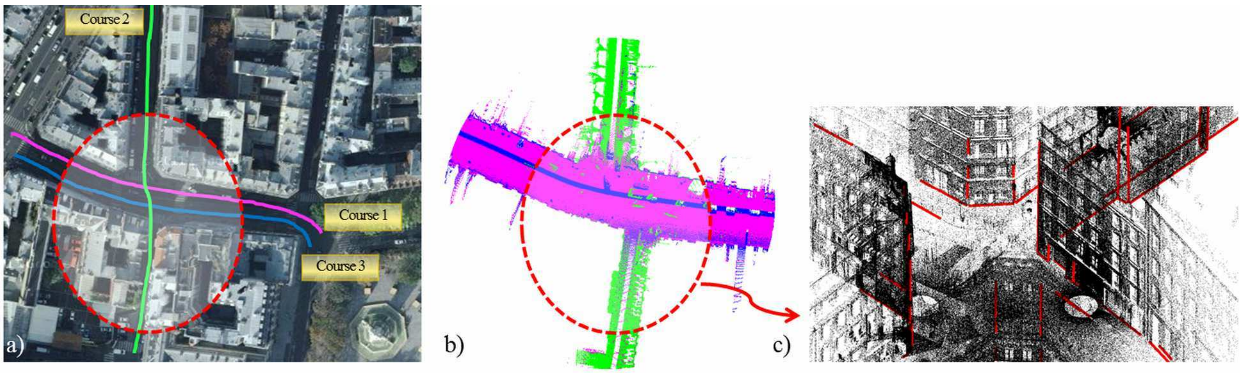

Figure 4 a) Test field and trajectory of the courses; b) MLS point clouds (colour corresponds to the courses); c) TLS point cloud with line segments used as a reference data

3.1 Repeatability Assessment

Among the collected MLS point cloud data sets, the overlapping ones were chosen and examined (Figure 4b). At all times the comparison was made in two ways: Test set to Model set and Model set to Test set. For each combination of MLS point clouds acquired for one course, a correspondence matrix between detected principal edges was created, using the calculated threshold value. Additionally, the Oriented modified LHD (OLHD) was computed. The final symmetric modified Hausdorff distance (LHD) representing the quality is marked with an asterisk (*) – Table 1. The performance of the proposed automatic correspondence-finding algorithm was validated against the ground-truth. The pairing between the two databases was determined correctly with a 99% confidence level. This detection rate was defined as the result of dividing the sum of correctly rejected (True Negative - TN) and correctly identified (True Positive - TP) pairs by the total number of combinations. The accuracy of the applied edge-detection method was not analysed in this study. The weigh W for the distance calculations d(t, m) was set as 10.

Table 1 Results of Repeatability Assessment

Data sets	Number of Segments		Pairing existing	False Negative		False Positive		Detection rate [%]		OLHD [m]	
	Test set p	Model set q						(TP+TN)/(p·q)			
Course1-Course2	36	28	11	0	1	1	2	99.90	99.70	**0.518***	0.472
Course1-Course3	36	41	29	2	2	5	5	99.53	99.53	**0.302***	0.297
Course1-Course3	33	32	23	2	2	2	3	99.62	99.53	0.343	**0.370***
Course2-Course3	28	41	12	1	1	2	2	99.74	99.74	0.481	**0.560***

Table 1 provides a summary of this precision assessment. The study shows that for urban areas a precision of about 0.3 m is achievable, despite the very challenging conditions for satellite navigation. In fact, the worst precision assessment was found when comparing with Course2. Thus, in order to verify the accuracy of the values obtained, the calculations were repeated for a new set of MLS edges extracted from the combination of Course1-Course3. The results are consistent which leads to the conclusion that the Course2 measurements were most likely affected by the poor quality of the GNNS signal. Low point cloud quality from this course is also shown by the absolute accuracy estimation (Table 2).

3.2 Absolute Accuracy Assessment

Table 2 summarises the statistical results obtained by comparing the line segments extracted from collected MLS data against the same line segments from TLS surveys. As expected, the absolute accuracy is worse for measurements performed along narrow streets and under tight turns, where the data is likely burdened by the poor quality of the GNSS solution. However in other cases, e.g. for passes with relatively open sky visibility and smoother navigation conditions, the absolute accuracy is of about 0.7 m. The evaluation procedure was the same as for the repeatability assessment method.

Table 2 Results of Absolute Accuracy Assessment

Data sets	Number of Segments		Pairing existing	False Negative		False Positive		Detection rate [%]		OLHD [m]	
	Test set p	Model set q						(TP+TN)/(p·q)			
Course1-TLS	36		21	1	1	9	8	99.57	99.61	**0.671***	0.656
Course2-TLS	28	64	22	2	2	10	18	99.33	98.88	1.104	**1.257***
Course3-TLS	41		24	2	2	7	9	99.66	99.54	0.563	**0.657***

* final modified LHD between two line segment sets

4. CONCLUSIONS

In this paper, a method for accuracy assessment of MLS point clouds has been presented. The proposed correspondence-finding algorithm for selecting related pairs among two sets of line segments produces convincing results. Nevertheless, an incorrect coupling may still be obtained in some cases, most often caused by a bad threshold value, which depends on the mutual displacement between line segment sets being compared. Consequently, lines in relatively close proximity in one set are paired with the corresponding feature lines in the other set as their distance does not exceed the selected threshold value. A one-to-one correspondence may be determined in further experiments in order to reduce the number of false positive results. The use of edges for the purpose of MLS point cloud quality assessment appears to be effective and well-suited, especially for low density point clouds. Built-up urban areas are a rich source of line features, which can be easily extracted. The performed studies have proved that the modified LHD is able to quantify MLS data and provide reliable results. It should be highlighted that the calculated average distance between two line segments sets represents not only the accuracy but also the error incurred during the edge extraction procedure. In future work we plan to isolate this component from the final computed accuracy value.

REFERENCES

Barber, D., Mills, J., 2008. Geometric validation of a ground-based mobile laser scanning system. ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 63, pp.128-141.

Chen, J., Leung, M.K., Gao, Y., 2003. Noisy logo recognition using line segment Hausdorff distance. The Journal of the Pattern Recognition Society, 36, pp. 943-955.

Feng, J., Zhong, R., Yang, Y., Zhao, W., 2008. Quality Evaluation of Spatial Point-Cloud Data Collected by Vehicle-Borne Laser Scanner. International Workshop on Education and Training & International Workshop on Geosciences and Remote Sensing.

Gandofi, S., Barbarelle, M., Ronci, E., Bruchi, A., 2008. Close photogrammetry and laser scanning using a mobile mapping system for high detailed survey of a high density urban area. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, part B5, pp.909-914.

Gao, Y., Leung, M.K., 2002. Line segment Hausdorff distance on face matching. The Journal of the Pattern Recognition Society, 35, pp.361-371.

Haala, N., Peter, M., Kremer, J., Hunter G., 2008. Mobile Lidar Mapping for 3D point cloud collection in urban areas – a performance test, In proceedings XXI ISPRS Congress.

Kaartinen, H., Hyyppä, J., Kukko, A., Jaakkola, A., Hyyppä, H., 2012. Benchmarking the Performance of Mobile Laser Scanning Systems Using a Permanent Test Field. Sensors 12, pp. 12814-12835.

Mano, K., Ishii, K., Hirao, M., Tachibana, K., Yoshimura, M., Akca, D., Gruen, A., 2012. Empirical accuracy assessment of MMS laser points clouds. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIX-B5, pp.495-498.

Paparoditis, N., Papelard, J-P., Cannelle, B., Devaux, A., Soheilian, B., David, N., Houzay, E., 2012. Stereopolis II : A multi-purpose and multi-sensor 3D mobile mapping system for street visualization and 3D metrology. Revue Française de Photogrammétrie et de Télédéction, pp.69-79.

Poreba, M., Goulette, F., 2012a. Assessing the accuracy of land-based mobile laser scanning data. Geomatics and Environmental Engineering, Vol.6, no. 3, pp.73-81.

Poreba, M., Goulette, F., 2012b. RANSAC algorithm and elements of graph theory for automatic plane detection in 3D point clouds. Archives of Photogrammetry, Cartography and Remote Sensing (in press).

Ray, J.A., Graham, L., 2008. New horizontal accuracy assessment tools and techniques for LIDAR data, In Proceedings of the ASPRS Annual Conference, Portland Oregon.