

AN AUTOMATIC ALGORITHM FOR MINIMIZING ANOMALIES AND DISCREPANCIES IN POINT CLOUDS ACQUIRED BY LASER SCANNING TECHNIQUE

Fabiane Bordin^a, Luiz Gonzaga Jr^{a b}, Fabricio Galhardo Muller^{a *}, Mauricio Roberto Veronez^{a c}, Marco Scaioni^d

^a Advanced Visualization Laboratory (VIZLab), University of Vale do Rio dos Sinos (UNISINOS)
São Leopoldo, Brazil - fabianebor@unisinos.br - fabriciogmuller@gmail.com

^b Graduate Program in Applied Computing (PIPCA), University of Vale do Rio dos Sinos (UNISINOS)
São Leopoldo, Brazil - lgonzaga@unisinos.br

^c Graduate Program in Geology (PPGEO), University of Vale do Rio dos Sinos (UNISINOS)
São Leopoldo, Brazil - veronez@unisinos.br

^d Department of Architecture, Built Environment and Construction Engineering, Politecnico di Milano
Milano, Italy - marco.scaioni@polimi.it

Commission V, WG V/5

KEY WORDS: Laser Scanning, LiDAR, Remote Sensing, Point Cloud, Return Intensity, Edge Effect, Quadtree, Clustering

ABSTRACT:

Laser scanning technique from airborne and land platforms has been largely used for collecting 3D data in large volumes in the field of geosciences. Furthermore, the laser pulse intensity has been widely exploited to analyze and classify rocks and biomass, and for carbon storage estimation. In general, a laser beam is emitted, collides with targets and only a percentage of emitted beam returns according to intrinsic properties of each target. Also, due interferences and partial collisions, the laser return intensity can be incorrect, introducing serious errors in classification and/or estimation processes. To address this problem and avoid misclassification and estimation errors, we have proposed a new algorithm to correct return intensity for laser scanning sensors. Different case studies have been used to evaluate and validated proposed approach.

1. INTRODUCTION

Laser scanning technique from airborne and ground platforms has been largely used for the reconstruction of high-resolution 3-D topography in the field of geosciences and other applications (Shan and Toth, 2008) (Laefer et al., 2014) (Bordin et al., 2013) (Ehlert and Heisig, 2013) (Sahin et al., 2012) (Buckley et al., 2008). The capture of spatial data through a laser scanner has become popular due the difficulties to access some targets or part of them, and mainly by the capability for reconstructing high accurate 3D models. Besides, the equipment can scan thousands of points per second (Burton et al., 2011) that reduces the field time. Furthermore, using digital data, it's possible to make the analysis in the office or laboratory, which reduces the fieldwork cost, time and in the most cases, human risks.

The digital acquisition by laser scanning records not only the target's position and color, but also the return intensity of the laser pulse. In the latest years, the reflection strength of the laser pulse intensity which returns to the sensor, or simply return intensity, has been also exploited to analyze and classify rock properties and for biomass and carbon storage estimation. However, anomalies and discrepancies can occur during the remote sensing, mainly the edge effect. The edge effect may occur when the laser collides partially with the target and part of the signal is lost or, also, collides with other undesired objects in the background. This effect makes the laser's return intensity be incorrect. When the purpose of the application is to analyze the reflectance of the laser's return intensity to exploit properties of materials that compose the target, the edge effect should be compensated to avoid misclassification errors.

To address this important problem, we have proposed a new algorithm to recover the correct returned laser intensity when an edge

effect is detected in collected data by laser scanning sensors. The methodology approaches mainly three different parameters: the intensity for each returned pulse, the laser scanners position, and the beams divergence. The algorithm core is based on point-cloud clustering, when each group corresponds to a different material. When the edge effect appears, those points which were affected are grouped into a class (existing or new). Under presence of the edge effect, one of the groups is tagged as edge effect group and the intensity value in this group needs to be corrected. Then, to determine the correct intensity of the returned pulse, and then to minimize the border effect, the algorithm estimates the collision percentage of the laser pulse with the target and the amount of the returned beam back to the laser station.

Three different case studies have been planned to demonstrate proposed algorithm effectiveness, based on three different targets: (1) artificial, but controlled target; (2) rock target; and (3) a forestry target. The results show that the proposed algorithm for edge-effect compensation is a feasible solution to recover the correct lasers return intensity according to the reflectance of the target structure, providing significant improvements and promising results for the development of applications based on laser scanning technique for different applications, including geoscience, civil engineering, environmental and geological studies.

2. ANOMALIES E DISCREPANCIES ON POINT CLOUDS

On laser scanning, beams are emitted from the laser towards the target, hit the target and return to the laser station. Through the LiDAR principles (Wehr and Lohr, 1999) the point position is calculated and the return intensity of that beam pulse is measured. Laser scanners have RGB cameras, too. They take pictures of the targets and, on the post-processing stage, pixels on that images

*Thanks to CAPES and FINEP agencies for funding

are associated to 3D-points. Finally, each point on point cloud has the following attributes (X, Y, Z, R, G, B, i) , where (X, Y, Z) represents 3D point coordinates, its respective colour (R, G, B) obtained from images and finally the return intensity.

We mean by edge (or border), the target silhouette where the laser's beam stop colliding with the target and start to collide with the background or a percentage of the laser beam is lost. Figure 1.(a) shows the rock's edges delimited by a red contour. Figure 1.(b) shows controlled target, with edges delimited by green contours.

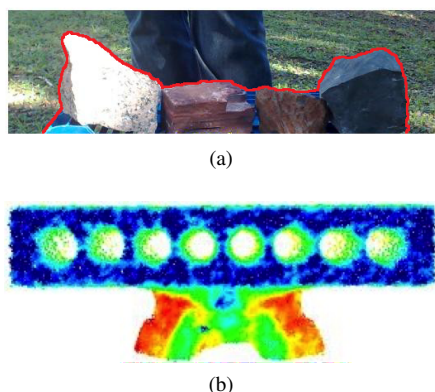


Figure 1: (a) The red area is the edge of the target where edge effect may appear. (b) The green points inside the circles and around the blue rectangle represent the edge effect seen in the 3D point cloud.

The intensity data (i) is directly proportional to the target's reflectance and inversely proportional to the target's distance, and suffer influences of atmospheric conditions and other agents (Ahokas et al., 2006) (Höfle and Pfeifer, 2007). The greater is the distance from the laser scanner to the target, the wider and less intense the laser pulse will be. The atmospheric conditions like the presence of rain also can modify the intensity data (Hodgetts, 2013). In general, a wider and less intense laser pulse due to the distance or weather results in an intensity data that does not correspond to the target's material. There are works which propose to correct the intensity data on the basis of the distance sensor–target (Höfle and Pfeifer, 2007) (Kaasalainen et al., 2011) and, also, applied in geological researches (Burton et al., 2011) as a remote sensing technique for rock properties analysis (Hodgetts, 2013).

Pingbo et al (Tang et al., 2009) also have proposed a method for estimating edge loss in laser scanned data by considering the impacts of scanning distance, density of data and incidence angle on the edge loss. The proposed method explores the return intensity, laser scanner's position and signals divergence, on the estimation algorithm, and a couple of controlled are performed to show the algorithm effectiveness. We have considered similar parameters to perform estimation and on corrected points recovering.

3. THE PROPOSED ALGORITHM

By using clustering algorithms, the point cloud can be classified into different groups based on the intensity data. Each group corresponds to one different material. Because the existence of the edge effect, one of the groups is going to be the edge effect group. The intensity value in the last group needs to be recovered. In the Figure 2, each color from the point cloud belongs to a different group of intensity.

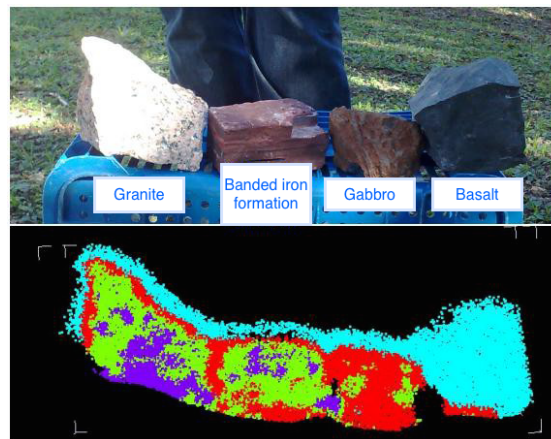


Figure 2: In the upper part four blocks scanned for the second experiments are shown. In the lower part, the effect of clustering intensity data is depicted.

On algorithm development, we have assumed the relation between the return intensity and the percentage of the collision is linear. So it can be defined by the equation 1 as the intensity recovery formula, to be applied to the edge effect clustered group.

$$I_r = \frac{I_m}{c_e} \quad (1)$$

where I_r is the recovered intensity value, I_m is the measured intensity by the laser and the c_e is the estimated collision value of the laser pulse with its target, where $\{c_e \in \mathbb{R}; 0 < c_e \leq 1\}$.

Once the c_e value is defined, it's easy to calculate the correct intensity value. Let's consider P as the set of clustered points which belong to the edge effect group, Q the set of all points that belong to the entire point cloud, p is a point from the set P being analysed and q is a point from the set Q . For each point p that belongs to P , an axis-aligned bounding box (AABB) is created centered in p . All points q from the set Q are tested to verify if q is inside the AABB box. The AABB size is defined through the spacing between points, which generates the *spacing* variable. This can be computed based on the scanning configuration because it's necessary to have a minimum number of points to calculate the collision approximation. If the q point is inside the AABB, then q is inserted in a quadtree of n level created in the same AABB's position and with same AABB's size. Figure 3 shows the AABB centered on the point p with neighborhood points and also shows the quadtree positioned in the point p (the point being analysed) with the points around.

Using the number of points that are in each quadrant of the quadtree, estimated collision value (c_e) between the laser pulse and the laser's target can be evaluated. Each quadrant has a collision percentage depending on the quadtree level (Equation 2). If the quadtree level (n) is 1, then the quadtree has 4 quadrants. In this case, each quadrant corresponds to 25% of collisions. Then the quadrant that has higher points number is considered that the quadrant collides 100% with the target. So, it collides a 100% of the 25% of the quadrant. The quadrant has 100 points and it's the quadrant with the higher number of points in a level 1 quadtree, it means that a 100 points correspond to 25% of collisions. If the second quadrant has 50 points, it means that it collides 50% of the 25% total of the quadrant, resulting in a 12.5% of total collisions. In the end, all of the collision percentage quadrants are summed up to estimate the laser pulse collision in that point (Algorithm 1).

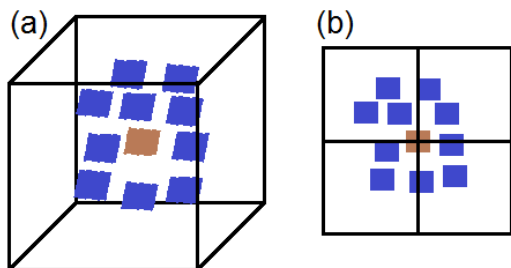


Figure 3: Subfigure (a) shows the AABB bounding box with the points inside. (b) Shows the quadtree centered in the same place of the AABB with points split in different cells.

$$perc = \frac{1}{4^n} \quad (2)$$

where $perc$ is the total collision percentage for each quadtree quadrant and n is the quadtree level.

Algorithm 1 Laser intensity recovery algorithm

Require: $P = \{p_1, p_2, \dots, p_n\}$, $Q = \{q_1, q_2, \dots, q_m\}$, n , $spacing$

- 1: **for all** $p \in P$ **do**
- 2: $aabb \leftarrow \text{CreateAABB}(p.x, p.y, p.z, spacing)$
- 3: $quadtree \leftarrow \text{CreateQuadtree}(p.x, p.y, p.z, spacing)$
- 4: **for all** $q \in Q$ **do**
- 5: **if** $q \subset aabb$ **then**
- 6: $quadtree \leftarrow q$
- 7: **end if**
- 8: **end for**
- 9: $quads \leftarrow \text{Sort}(quadtree, n)$
- 10: $highest \leftarrow quads[0]$
- 11: $percPerQuad \leftarrow \frac{100}{4^n}$
- 12: $sum = 0$
- 13: **for all** $quad \in quads$ **do**
- 14: $sum += \frac{(quad * percPerQuad)}{highest}$
- 15: **end for**
- 16: $sum = \frac{sum}{100}$
- 17: $p.intensity = \frac{p.intensity}{sum}$
- 18: **end for**

In the worst case scenario, the algorithm has numerical burden of $O(n^2)$, where the most part of the points from the point cloud are clustered and considered belonging to the edge effect group. Also, to determine the collision percentage is required to go through all the point cloud testing if each point is inside the AABB Bounding Box.

4. RESULTS AND DISCUSSION

The proposed algorithm has been evaluated to correct the return intensity signal on the occurrence of the edge effect for three different case studies : (1) synthetic point cloud, (2) rock's classification and (3) forestry target.

4.1 Case Study I: synthetic point cloud

For the first study, it was created a synthetic point cloud where all the points, with one line exception, have the intensity value of 100, over a normalized range between 0-255. The "edge effect" line is the sequence of points in a straight line which has intensity

value of 50. This point cloud was clustered and the points were separated into two groups. Figure 4 shows the synthetic point cloud before and after the recovery algorithm. In Figure 4(a), there are two groups, the blue points correspond to the correct intensity value while the gray ones correspond to the edge effect group. The red line highlights the edge effect points. After applying the intensity recovery algorithm in the synthetic point cloud and clustering that point cloud again, almost all the points have been assigned to the same group. Figure 4(b) shows almost all the points in the same group and the red line highlight the points which were part of the edge effect and now is part of the same group. Since the synthetic point cloud has few points, its data is detailed in the table 1.

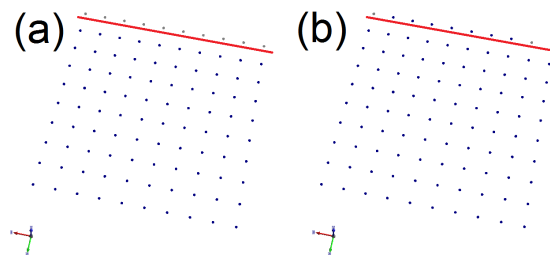


Figure 4: Synthetic point cloud clustered before the intensity recovery (a). Synthetic point cloud clustered after the intensity recovery (b).

Table 1: Intensity data recovered through the created algorithm

Original	Collision	Recovered	Difference (100)
50	30.8642%	162	62
50	34.7222%	144	44
50	34.7222%	144	44
50	46.2963%	108	8
50	55.5555%	90	10
50	50.9259%	98	2
50	43.6508%	114	14
50	39.6825%	126	26
50	30.8642%	162	62
50	30.5555%	163	63

It's noticeable that the points in corners of the point cloud feature worst results than the points in the center of the point cloud. This fact occurs because the synthetic point cloud ends abruptly in the corners. While it does not happen in the point clouds of real objects captured by laser scanning. Indeed, in such case, the points in the corners would not have the same intensity value of the points in the center. Figure 5 shows why the recovered intensities differences occur. But the collision estimation is valid, since the collision percentage for corner points are smaller than the collision estimation of the points in the center.

4.2 Case Study II: rock's classification

The second evaluation, it was used a real point cloud scanned with TLS - Optech ILRIS-3D¹. Different types of rock blocks were put in a basement to be scanned with the TLS ILRIS-3D. Figure 2 shows that after clustering the point cloud, one intensity group is clearly the edge effect.

It's also possible to see that one rock is part of the edge effect since this specific rock reflects low laser pulse intensity. The Figure 6a shows original target. After applying the algorithm to

¹<http://www.teledyneoptech.com/index.php/product/optech-ilris/>

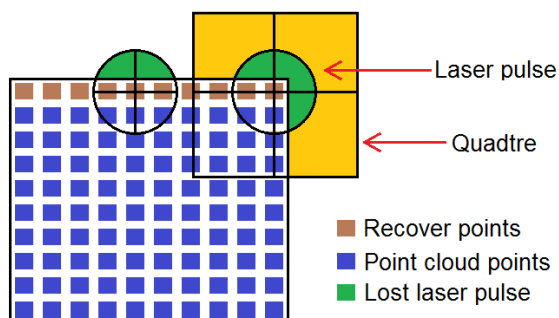


Figure 5: The green area is the lost part of the laser pulse that was estimated correctly by the quadtree, but the synthetic intensity value generated was equal to all points.

correct the intensity data it's still possible to see a residual edge effect, but smaller than the one in the original point cloud (Figures 6c and 6d). Moreover, it's possible to distinguish the rock that was masked by the edge effect. The difference of the edges in the Figures 6c and 6d occurs because of the different parameters of the algorithm, but both parameters recovered the intensities of the original point cloud.

These parameters make the AABB and the quadtree bigger and smaller, which makes the collision estimation to reach a value closer to the reality. There's a relation between the size and number of points in the AABB and the quadtree that can provide better results.

4.3 Case Study III: forestry target

Also, the algorithm was applied to a point cloud of a tree (Figure 7) to see its behaviour with many more points and with points that were distributed in a more randomly way. The process was the same, cluster the points, detect the edge effect and apply the recovery algorithm. Afterwards, we re-cluster the point cloud. And the clustered groups were rearranged as Table 2.(a) shows. It can be seen that the edge effect was not totally removed, but the algorithm reduced approximately 51,765%, Figure 7.(b).

Table 2: Number of points in each group before and after the edge effect correction

Group	Points (before)	Points (after)
Edge	520.053	250.844
Stem	81.812	221.901
Leaf	481.131	610.251

5. CONCLUSION AND FUTURE WORKS

In this work, we presented an automatic algorithm to recover the intensity of points that suffered from the edge effect on targets scanned by terrestrial laser scanners. The algorithm has recovered the intensity data based on the collision estimation and point cloud clustering, without previous setup or additional accessory.

This algorithm was applied to a synthetic point cloud and in a real point cloud. In both cases, it succeeded to effectively minimize the edge effect. As future work, other clustering algorithms, not based on centroids, can be evaluated for different targets. Computational optimizations are also needed to improve the algorithm's efficiency. The algorithm sourcecode will be distributed under GPL license, through VIZLab/Unisinos site. Please, visit us at <http://vizlab.unisinos.br> or contact us by e-mail.

REFERENCES

- Ahokas, E., Kaasalainen, S., Hyyppä, J. and Suomalainen, J., 2006. Calibration of the optech altm 3100 laser scanner intensity data using brightness targets. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36(Part 1), pp. 1Á6.
- Bordin, F., Teixeira, E. C., Rolim, S. B. A., Tognolic, F. M. W., Veronez, M. R., da Silveira Junior, L. G. and de Souza, C. F. N., 2013. Methodology for acquisition of intensity data in forest targets using terrestrial laser scanner. *IERI Procedia* 5, pp. 238–244.
- Buckley, S. J., Howell, J. A., Enge, H. D. and Kurz, T. H., 2008. Terrestrial laser scanning in geology: data acquisition, processing and accuracy considerations. *Journal of the Geological Society* 165(3), pp. 625–638.
- Burton, D., Dunlap, D. B., Wood, L. J. and Flaig, P. P., 2011. Lidar intensity as a remote sensor of rock properties. *Journal of Sedimentary Research* 81, pp. 339–347.
- Ehlert, D. and Heisig, M., 2013. Sources of angle-dependent errors in terrestrial laser scanner-based crop stand measurement. *Computers and Electronics in Agriculture* 93, pp. 10–16.
- Hodgetts, D., 2013. Laser scanning and digital outcrop geology in the petroleum industry: a review. *Marine and Petroleum Geology* 46, pp. 335–354.
- Höfle, B. and Pfeifer, N., 2007. Correction of laser scanning intensity data: Data and model-driven approaches. *Photogrammetry e Remote Sensing* 62(6), pp. 415–433.
- Kaasalainen, S., Jaakkola, A., Kaasalainen, M., Krooks, A. and Kukko, A., 2011. Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: Search for correction methods. *Remote Sensing* 3(10), pp. 2207–2221.
- Laefer, D. F., Truong-Hong, L., Carr, H. and Singh, M., 2014. Crack detection limits in unit based masonry with terrestrial laser scanning. *NDT & E International* 62, pp. 66–76.
- Sahin, C., Alkis, A., Ergun, B., Kulur, S., Batuk, F. and Kilic, A., 2012. Producing 3d city model with the combined photogrammetric and laser scanner data in the example of taksim cumhuriyet square. *Optics and Lasers in Engineering*.
- Shan, J. and Toth, C. K. (eds), 2008. *Topographic Laser Ranging and Scanning: Principles and Processing*. CRC Press.
- Tang, P., Akinci, B. and Huber, D., 2009. Quantification of edge loss of laser scanned data at spatial discontinuities. *Automation in Construction* 18(8), pp. 1070–1083.
- Wehr, A. and Lohr, U., 1999. Airborne laser scanning – an introduction and overview. *ISPRS Journal of Photogrammetry & Remote Sensing* 54, pp. 68–82.

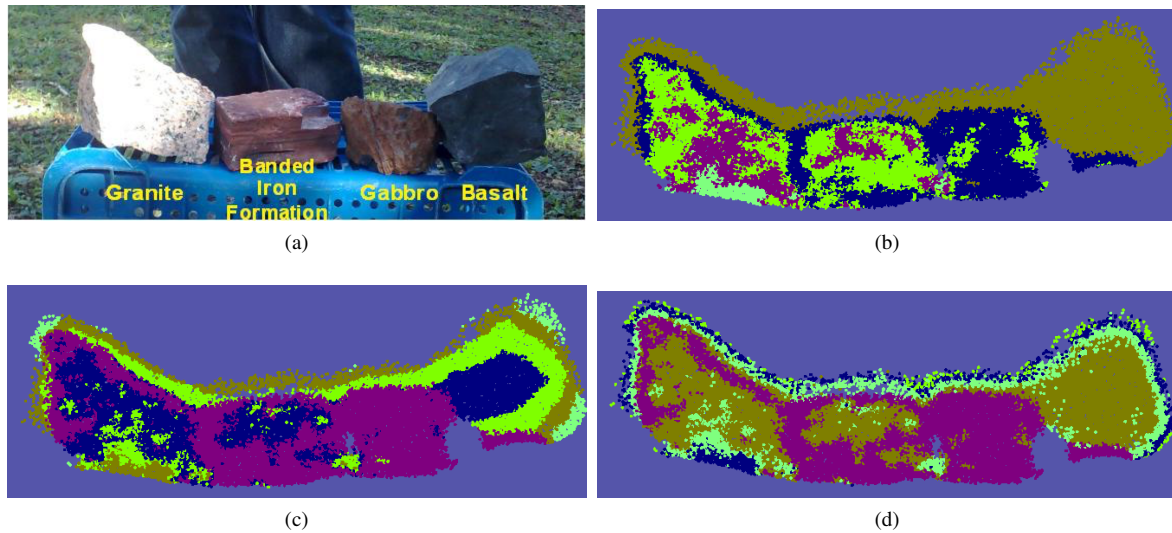


Figure 6: (a) Rock samples, includes granite, banded iron formation, basalt and gabbro; (b) rock samples before running the algorithm to correct the intensity data; (c) Clustered rock samples after running the algorithm to correct the intensity data; (d) Clustered rock samples after running the algorithm to correct the intensity data with different set of parameters.

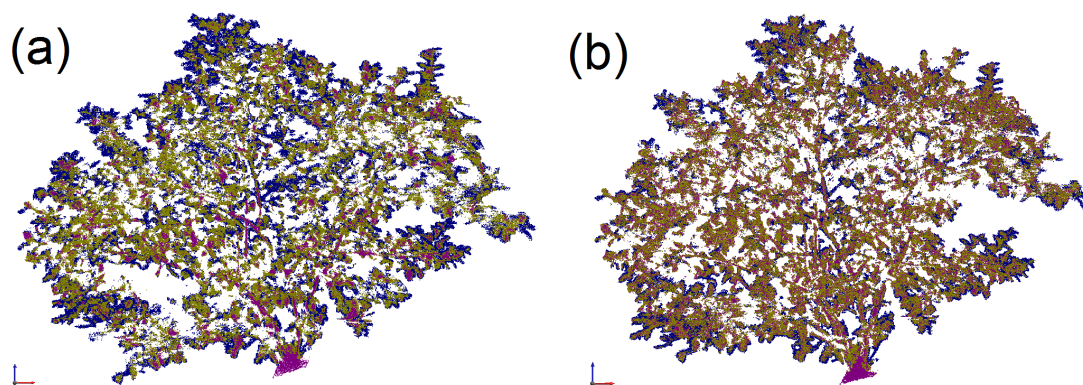


Figure 7: Clustered tree before (a) and after (b) the return intensity recovery algorithm to be applied