HEURISTICALLY TARGETED MINIMUM DESCRIPTION LENGTH TEST FOR STONE DETECTION FROM PUBLIC POINT CLOUD DATA

Paavo Nevalainen[1] *and Juuso Suomi*[2] *and Jukka Heikkonen*[1]

[1]Dept. of Information Tech., [2]Dept. of Geography and Geology,
University of Turku, FI-20014 Turku, FINLAND, ptneva@utu.fi

ABSTRACT

Coarse cross-terrain point clouds are gathered by aerial laser scan (ALS) and dense point clouds by unmanned vehicle (UAV) operation. These two data sources have complementary nature and should be combined for various applications. This paper uses minimum description (MDL) length approach to detect individual stones and their physical dimensions from UAV data. The MDL procedure is spatially targeted by a two-step heuristics: local stoniness likelihood derived from ALS data and the curvature detection on UAV data. Comparison of the performance of MDL principle and a geometric approach, namely mean square error (MSE) minimization is presented. The MDL approach can be applied to cloud point densities $\rho \geq 3\,m^{-2}$.

1. INTRODUCTION

The problem of estimating terrain surface stoniness is related to locating automatically potential gravel deposits nearby infrastructure sites in Northern Finland. One can produce a 20$\times$ 20 m raster file about stoniness likelihood [1]. Also, geomorphological models are getting more complex using many micro-topological features, and it seems that the stone size distribution and stone coverage ratio could serve as two new, ubiquitous features.

Publicly available nation-wide aerial laser scan (ALS) 3D point clouds have the ground point density $\rho = 0.7\,m^{-2}$, which does not allow the detection of individual stones. At some sites higher ground point densities $\rho = 42\,m^{-2}$ are produced by unmanned aerial vehicles (UAV). The minimum description length principle (MDL) is proposed to detect individual stones and their radius from these dense clouds. The MDL process needs a good initial guess of each stone location, and a spatial angle filtering (SAF) algorithm is suggested to produce a list of most likely stone positions at each site with a high stoniness likelihood.

The encoding cost approximation of the point cloud points is based on so called crude MDL [2], where the model and data are encoded separately. A further assumption is made at the proximity of an object about normal distribution of point cloud points in orthogonal direction from the object, and about uniform distribution along the surface of the object. This assumption is common in point cloud research, see e.g. [3].

The Section 2 introduces the crude version of the MDL principle applied to detecting the ground as a planar object and stones as hemispherical objects in the 3D point cloud. A generic "potato" shape is addressed in Section 2.2. Choosing the likely stony areas for further analysis and pinpointing MDL check to likely stone locations is briefly outlined in Section 3. A comparison between geometric stone parameter fitting and MDL, and some early results from a test site with a dense point cloud and known geometry of some stones is presented in Section 4. Conclusions are provided in Section 5.

2. MDL OF VARIOUS POINT SETS

In this treatise a point set $\mathcal{X}$ gets associated to an object X, if the set has an advantageous distribution at the normal of the surface of the object X. One has to compare the cost of encoding a point set $\mathcal{Q} \subset [0, R]^3 \subset \mathbb{R}^3$ associating points either to a symbolic outliers object O or to the object X. The first case is assumed to have uniform distribution (even it is known to be too pessimistic). A pixel accuracy ϵ and a nominal length R define the encoding cost $\phi_3 = 3\,lb(G)$ of one 3D point, where $G = R/\epsilon$ is the characteristic pixel amount and $lb(.) = \log_2(.)$ is an abbreviation of a binary logarithm.

Now, the encoding length of whole cloud assuming all points belonging to an outlier object O, is: $\Phi(\mathcal{G}|O) = |\mathcal{G}|\phi_3$, where $|G|$ stands for the size of the local point cloud $\mathcal{G}$. Assuming an object X included to the model, the encoding length becomes:

$$\Phi(\mathcal{G}|O, X) = C|\mathcal{G}| + \phi_3|\mathcal{O}| + \Phi(X) + \Phi(\mathcal{X}|X), \quad (1)$$

where $C = 2$ is the encoding length to distinguish the point classes. In this case $\mathcal{Q}$ has to be divided to two distinct sets, outliers $\mathcal{O}$ and the set $\mathcal{X}$: $\mathcal{Q} = \mathcal{O} \cup \mathcal{X}$. The object parameter encoding length $\Phi(X)$ and the object-specific point information $\Phi(\mathcal{X}|X)$ are also needed. A summary of objects and their parameter encoding lengths is given in Table 1. The only 2D objects is the line $\mathcal{L}_2$. The rest are 3D objects: $\mathcal{L}_3, \mathcal{C}, \mathcal{P}, \mathcal{S}$ and $\mathcal{E}$ as a line, a circle, a plane, a hemisphere and an ellipsoid, respectively. A generic smooth object $\mathcal{M}$ (potato) will be introduced in Section 2.2. The shapes $\mathcal{S}, \mathcal{E}, \mathcal{M}$ are alternative representations of a stone.

The question is, which encoding is shorter: $\Phi(\mathcal{G}|O)$ or $\Phi(\mathcal{G}|X, O)$? Initially any encoding based on the presence

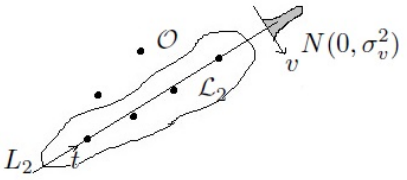

Figure 1. An example of 2D line L_2 and the point set $\mathcal{L}_2$ it encodes. Point projection distances v have normal and tangential components t have uniform distribution.

of an object X is more costly, but there is a break-even point at $|\mathcal{X}| = n_2$. There is also a geometric definition limit $|\mathcal{X}| = n_1 < n_2$, at which the object X becomes geometrically defined and where the geometric error of a fit equals zero.

Each point $p \in \mathcal{X}$ will be projected on the object surface by $proj(p, X)$ over an orthogonal distance $d(p, X) = ||p - proj(p, X)||_2$. Both components will be encoded separately by tangential and vertical encoding with point-specific costs $\phi_v(.)$ and $\phi_v(.)$, respectively:

$$\Phi(\mathcal{X}|X) = \sum_{p\in\mathcal{X}} \{\phi_t(proj(p, X)) + \phi_v(d(p, X))\}. \quad (2)$$

The cost $\phi_v(v)$ of vertical encoding of a single real value $v \sim N(0, \sigma_v^2)$ is the classical Huffman encoding result [3]:

$$\begin{aligned} \phi_v(v) = &\quad \tfrac{1}{2\ln(2)}(\tfrac{v}{\sigma_v})^2 + lb(\tfrac{\sigma_v}{\epsilon}) + \tfrac{1}{2}lb(2\pi) \quad (3)\\ \phi_v \approx &\quad \mathbb{E}_{p\in\mathcal{X}}[\phi_v(d(p, X))] = \tfrac{2}{\ln(2)\pi^2} + 2, \end{aligned}$$

where the expected encoding cost ϕ_v of vertical distances v follows by choosing $\sigma_v = 1.5\,\epsilon$ which leads to $\mathbb{E}[v/\sigma_v] = \sqrt{2/\pi}$. The choice is for informal consideration only, and it can be justified by the observed ground height distribution in the dense point cloud.

The tangential part is uniform distributed, and the cost $\phi_t = \mathbb{E}_{p\in\mathcal{X}}[proj(p, X)]$ for expressing one point projection can be derived case by case (see Table 1). Derivation of ϕ_t of a line L_2 and a D3 plane P have been given in [3]. Other cases are similar, and the derivation is excluded from this presentation due the space considerations. The break-even point cloud size n_2 in Table 1 can now be calculated by:

$$\begin{aligned} \Phi(\mathcal{X}|X) = &\quad \Phi(\mathcal{X}|O)\\ \Phi(X) + n_2\,(\phi_t + \phi_v) = &\quad n_2\,\phi_3 \quad (4) \end{aligned}$$

A summary of minimum detectable stone radii with sparse and dense data is given in Table 2. One can see that the MDL principle cannot be used for nation-wide sparse data! A rather typical stone with radius $r = 0.6\,m$ would need a point density $\rho \geq 11/m^2$ to be detected as a hemispherical object ($\rho \geq 23/m^2$ for an ellipsoid).

An efficient implementation of stone detection consists of two tasks: to properly initialize an object X (its location and radius) and to improve the initial parameter choice in MDL sense by addressing points to two possible classes: outliers $\mathcal{O}$ and object-specific points $\mathcal{X}$. We

Table 1. An informal summary of some geometric objects and their model costs. Uniform horizontal point cloud distribution has been assumed.

$\mathcal{X}$	$\Phi(X)$	$\phi_t(p)$	n_1	n_2
$\mathcal{O}$	0	-	1	1
$\mathcal{L}_2$	$2\phi_3/3$	$\phi_3/3$	2	6
$\mathcal{L}_3$	$5\phi_3/3$	$\phi_3/3$	2	4
$\mathcal{C}$	$2\phi_3 + 1$	$\phi_3/3$	3	5
$\mathcal{P}$	$\phi_3 + 1$	$2\phi_3/3$	3	6
$\mathcal{S}$	$4\phi_3/3$	$2\phi_3/3 + 1$	4	12
$\mathcal{E}$	$3\phi_3$	$1...2\phi_3/3 + 1$	9	26

Table 2. Minimum size of detectable stones.

ground point density ρ $(1/m^2)$	0.7	42
hemisphere min. radius	2.3	0.3
ellipsoid min. axis	3.5	0.5

propose a heuristical likelihood for a stone derived from coarse point cloud for the first task. The second task is best done by a random sample consensus (RANSAC) algorithm [4]. Note that actually there are three possible models: $\Phi(\mathcal{G}|\mathcal{O}), \Phi(\mathcal{G}|P, O)$ and $\Phi(\mathcal{G}|S, P, O)$, for outliers only, a plane (as the local ground surface) and outliers, and for a stone, a plane and outliers, respectively. The class encoding length C equals $0, 1$ and 2 for each three cases. A basic RANSAC needs to be modified for the last case, details are not included in this presentation.

2.1. The striped point cloud

The vertical projection of the point cloud is very seldom uniformly distributed. The laser scan process produces stripes, therefore the tangential encoding cost will be affected. The projection $proj(p, X)$ changes to a projection to a nearest sweep line L_3. Lines are at regular intervals and one can associate points to correct line by assuming continous tangential co-ordinate on subsequential lines or circles. Thus following changes are needed in Table 1:

$$\phi_t = \Phi_3/3 \;\text{ for plane } \mathcal{P} \text{ and sphere } \mathcal{S} \quad (5)$$

Also, one has to encode the indexing of the sweep planes by $\phi_i = lb(R'/\Delta)$, where $R' > R$ is an approximate width of the potential stone location and Δ the stripe separation.

2.2. A generic continuous shape $\mathcal{M}$

MDL can be understood as an alternative regularization methodology for usual geometric fitting by minimizing the mean square error (MSE) of orthogonal distances $d(p, M), p \in \mathcal{M}$. From various possible geometric regularization terms, the minimization of Gaussian curvature κ_G leads to a very close relation to MDL minimization results. Equation 6 defines the κ_G regularization:

$$l(M|\mathcal{M}) = \sum_{p\in\mathcal{M}} [d(p, M)]^2 + \lambda \overbrace{\sum_{p\in\mathcal{M}} \kappa_{Gp} A_p}^{\approx \int_M \kappa_G dA}, \quad (6)$$

where the last sum represents an approximate integration of Gaussian curvature κ_G over the surface of M via discrete differential geometry (DDG) approach using a triangularization T of the generic surface M produced by the SAF algorithm of [1]. Approximation of the curvature κ_G can be based on spatial angles of SAF using so called spherical excess (see [1]) or so called angle defect (see[5]). The curvature κ_G is not available (or is not meaningful) for other objects in Table 1. This equals setting the regularization weight $\lambda \equiv 0$ for other target functions $l(X|\mathcal{X})$ of the geometric fit. There are alternatives to the geometric regularization term, e.g. the number of triangles $|T|$ can be a cost term.

The MDL principle requires projection $proj(p, M)$ to the nearest surface triangle applied just as in Equations 2 and 3. The projectional encoding length $\Phi_t(.)$ is based on the parameterization of each triangle $t \in T$ separately:

$$\Phi_t(proj(\mathcal{M}, M)) = |\mathcal{M}|\overset{1}{lb}(|T|) + 3|T|\overset{2}{lb}(|\mathcal{M}|) + \quad (7)$$
$$+2|\mathcal{M}|lb(\overset{3}{\sqrt{|M|}}/\epsilon),$$

where $|M| = \sum_{t \in T} area(t)$ is the total surface area of the object M. The term 1 of 7 refers to allocating of $|\mathcal{M}|$ points to $|T|$ triangles. The term 2 is the encoding of the triangle information; each triangle consists of 3 points. The term 3 is about the encoding of each point by two local planar coordinates of some triangle $t \in T$.

For a practical implementation, one has to limit the shapes of triangles $t \in T$. The current attempt has triangles selected from a subset $\mathcal{M}_T$ of points $\mathcal{M}$ while so called Delaunay property [6] is being enforced. Details of this approach are still a topic of research.

3. TARGETING OF MDL TESTS

The site selection process is described in Figure 3). The first box generates a stoniness likelihood [1] map, which consists of $20 \times 20\, m^2$ pixels. This likelihood map can be produced at areas where the ground hits exceed 60-70% of all ALS cloud points. This condition holds on most of the northern Finland. The first phase uses sparse ALS data and second phase dense UAV or ALS data (where available). The criterion for starting the MDL check at a specific point is based on the Gaussian curvature of the ground triangulation to have approximately constant value at neighboring ground points. Details are omitted in this presentation, and will be published later in an expansion paper. Red dots at Figure 2 represent prominent places to perform the MDL test. The quality of the local stone map shown in Figure 2 can be improved by neighborhood voting. Details of this process are still being developed. There are many possibilities and the goal is to develop a non-parametric adaptive process.

4. EXPERIMENTS AND RESULTS

A $220 \times 320\, m$ test site with high-density point cloud has been used. The site is located at Harakkakallio at Turku, Finland. The coordinates are: $60.44^o\, N$, $22.2^o\, E$.

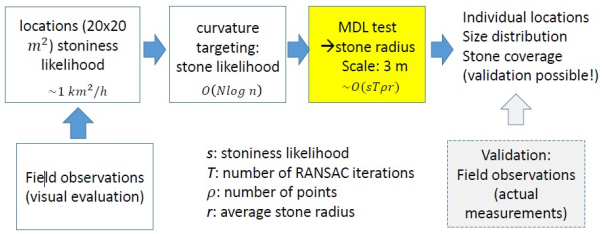

Figure 2. Point cloud pre-processing pipeline. Likely locations are scanned by a Gaussian curvature filter to pinpoint the MDL test spots. The scale of operations proceeds from 500 km to 20 m and 3 m.

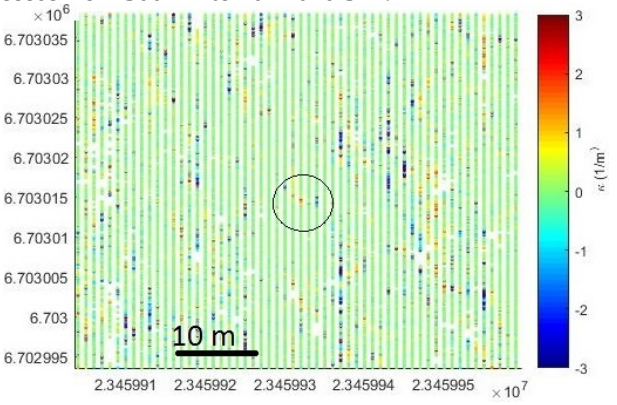

Figure 3. Unfiltered Gaussian curvature distribution. Neighborhood voting would focus red stripes (possible stones) to red dots. The test site is in the city of Turku, Finland.

A ground detail of this cloud is depicted in Figure 3. The cloud is exceptionally striped limiting the detection and radius estimation of the smallest stones. The stone B at Figure 3 has radius of $r = 0.4m$, and it extends over one stripe only. The stone C is larger ($r = 1.2m$), but the sample suffers from sparse ground hits because of the local thick canopy. There are some other small stones visible in the Figure 3.

The following treatise compares the root mean square error (RMSE) values e to the average MDL encoding lengths $\phi = \Phi(\mathcal{G}|...)/|\mathcal{G}|$ with different point cloud densities. The upper density limit $\rho = 120\, m^{-2}$ is typical for local photometric UAV scans. The site is covered by $\rho = 80\, m^{-2}$ helicopter ALS scan, and the low limit is the nationwide open cloud with $\rho = 0.8\, m^{-2}$. The low densities are simulated by uniform removal of points. In reality low density ALS scanners have larger beam radius and higher power thus having better penetration. This leads to somewhat different ground hit height distributions on each point cloud densities. The test side has s.d. $\sigma_z = 0.01...0.04\, m$. The low range is for an urban grass field and the upper range is met at the thicket areas with high ground vegetation. Three stones were selected and both MSE and the average MDL encoding length ϕ calculated.

The top part of the Figure 4 shows that the assumption of a plane is simply a wrong one, when the analysis is tightly centered on a stone. But even the planarity assump-

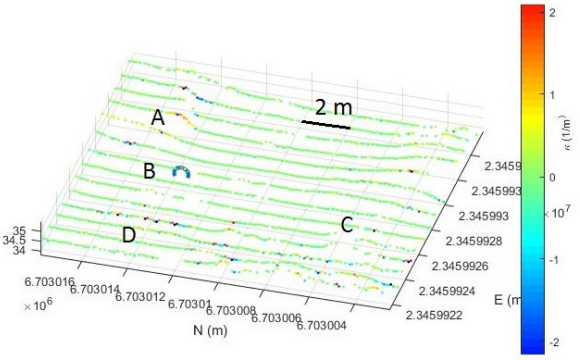

Figure 4. Curvature summary from circled spot of Figure 2. A: a non-stone shape. B: A small stone with radius 0.4 m. C: A large stone on many scan stripes, but deteriorated by canopy hits. D: Noisy curvature by vegetation.

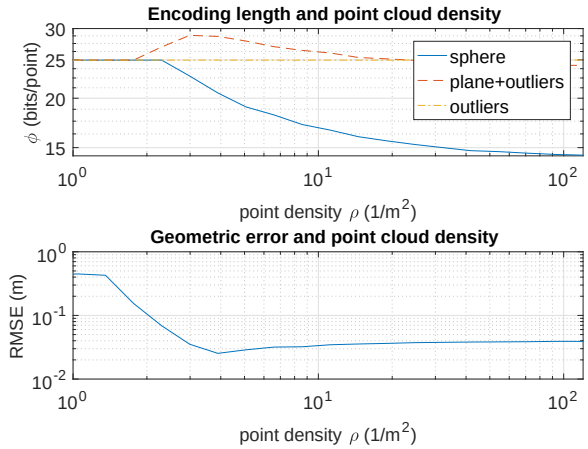

Figure 5. The encoding length per point on different point cloud densities ρ. A set of three stones of radii $r \in \{0.5, 0.8, 1.3\}\, m$ encoded by 3 different assumptions: sphere and outliers, plane and outliers, outliers only.

tion $\Phi(\mathcal{G}|P)$ is better than encoding all points as outliers. The MDL test for sphere is usable at densities $\rho > 3\, m^{-2}$ when the average encoding length $\phi < 0.9\, \phi_3$.

The lower part of Figure 4 shows the root mean square error (RMSE) of a spherical fit from the same three stones. There are many possible geometric model tests, see e.g [7]. RMSE is closely related to maximum likelihood estimation (MLE) approach, that is why RMSE was used as a measure of geometric fit. Initially, at low densities the spherical fit is not logically possible. The geometric fitting is much too optimistic at densities $\rho = 2...3\, m^{-2}$, where cases of 4 hits per stone (and perfect fit!) are common. The MDL principle is more conservative and begins to detect stones only later, when it is meaningful. The curves are averaged and s.d.'s are not indicated, but the behaviour of MDL test is systematic in this respect by the very definition of MDL principle.

5. CONCLUSION

This study is only rudimentary probing the possibilities of combining the existing stoniness likelihood pipeline of [1], preliminary narrowing of MDL test by neighborhood voting and the usability of the MDL test compared to MLE. Results, although preliminary, indicate that ground hit densities $\rho \geq 3\, m^{-2}$ can be analyszed by the existing pipeline and MDL principle.

Another preliminary result is the size limit of the individually detectable stones in Table 2. To extend this treatise to have concrete scientific significance, the following tasks must be completed:

- Locating and measuring a set of stones at the test site for validation purposes.
- Choosing between various neighborhood voting methods to improve the likely stone locations seen in Figure 2.

Also, one has to experiment with the generic smooth shape M either for detection of prominent ground objects or for generating an information theoretically justifiable alternative to the current triangulated ground model produced by SAF.

6. ACKNOWLEDGMENTS

Turku city arranged dense data sets from an accessible locality. Especially efforts of GIS engineer Asmo Leskinen are gratefully acknowledged.

7. REFERENCES

[1] Paavo Nevalainen, Maarit Middleton, Raimo Sutinen, Jukka Heikkonen, and Tapio Pahikkala, "Detecting terrain stoniness from airborne laser scanning data ," *Remote Sensing*, vol. 8, no. 9, pp. 720, 2016.

[2] Peter Grünwald, "A tutorial introduction to the minimum description length principle," in *Advances in Minimum Description Length: Theory and Applications*. 2005, MIT Press.

[3] Michael Ying Yang and Wolfgang Förstner, "Plane detection in point cloud data," Tech. Rep. TR-IGG-P-2010-01, 2010.

[4] Marco Zuliani, "Ransac for dummies," Jan 2012.

[5] Keenan Crane, Fernando de Goes, Mathieu Desbrun, and Peter Schröder, "Digital geometry processing with discrete exterior calculus," in *ACM SIGGRAPH 2013 Courses*, New York, NY, USA, 2013, SIGGRAPH '13, pp. 7:1–7:126, ACM.

[6] Mark de Berg, Otfried Cheong, Marc van Kreveld, and Mark Overmars, *Computational Geometry: Algorithms and Applications*, Springer-Verlag TELOS, Santa Clara, CA, USA, 3rd ed. edition, 2008.

[7] Kenichi Kanatani, "Geometric bic.," *IEICE Transactions*, vol. 93-D, no. 1, pp. 144–151, 2010.