# **3D OCTREE BASED WATERTIGHT MESH GENERATION FROM UBIQUITOUS DATA**

L. Caraffa , M. Brédif, B Vallet

Université Paris-Est, IGN/SR, MATIS, 73 avenue de Paris, 94160 Saint Mandé, France firstname.lastname@ign.fr

## Commission III, WG III/5

KEY WORDS: Surface reconstruction, Octree, Graph-cut, Lidar

## **ABSTRACT:**

Despite of the popularity of Delauney structure for mesh generation, octree based approaches remain an interesting solution for a first step surface reconstruction. In this paper, we propose a generic framework for a octree cell based mesh generation. Its input is a set of Lidar-based 3D measurements or other inputs which are formulated as a set of mass functions that characterize the level of confidence on the occupancy of each octree's leaf. The output is a binary segmentation of the space between *occupied* and *empty* areas by taking into account the uncertainty of data. To this end, the problem is then reduced to a global energy optimization framework efficiently optimized with a min-cut approach. We use the approach for producing a large scale surface reconstruction algorithm by merging data from ubiquitous sources like airborne, terrestrial Lidar data, occupancy map and extra cues. Once the surface is computed, a solution is proposed for texturing the mesh.

## 1. INTRODUCTION

Surface reconstruction from point cloud is an important topic that has already been greatly studied. One of the reasons is the constantly increasing number of applications that use surface reconstruction: digital elevation model computation for flood simulation, 3D modelling, robot path planning, etc., and technical acquisitions: Lidar, images, etc. Despite of the important number of approaches proposed in the state of the art, the ability to handle the acquisition ubiquity of the data remain a challenging task. It is even more true in a large-scale context. It is within this context that (Nexa et al., 2015) has recently provided a benchmark that covers different relevant tasks as camera calibration and dense reconstruction on challenging data. This benchmark is decomposed in 3 scenes that contains thousands of images on both airborne and terrestrials camera coupled with millions of 3D points from Laser. In the context of data merging, (Fuhrmann and Goesele, 2011) proposes a method for merging depth maps with multiple scales, this method takes various depth maps at different scales and produces a piecewise surface.

Thanks to efficient optimization methods such as graph-cuts, global optimization formulation became very popular for producing watertight surfaces. The problem can be seen as a volumetric segmentation of a discretization of the space where the resulting surface is represented by the interface between inside and outside areas. It has been first introduced in (Snow et al., 2000) with a photometric approach.

In this paper, we propose an approach based on the Dempster Shafer theory that allows, with an arbitrary number and type of inputs, a large scale reconstruction of the space. We apply this model to the lidar point cloud reconstruction case. We also propose a formulation for texturing the segmentation based on extra information like laser reflectance.

Figure 2 shows the proposed scheme of the surface reconstruction in 2D. The input is a set of occupancy function that are merged with the Demptser Shafer theory. The final function is then segmented into a binary indoor/outdoor partition. In this case, each pixel is a leaf of the octree. Figure 1 shows the final 3D result of the algorithm.

## 2. SURFACE RECONSTRUCTION FORMULATION

The aim of the proposed method is to find, for a set of inputs I, a segmentation of the space as *occupied* or *empty* following a suitable model. Inputs are given by a set of mass functions  $m_i$  defined by (see section 2.3):

$$m_i(P) \to \begin{bmatrix} 0, 1 \end{bmatrix}^3$$
  

$$P \mapsto (e, o, u) \text{ s.t. } e + o + u = 1$$
(1)

where e, o and u are the occupancy masses for a 3D point  $P \in \mathbb{R}^3$  to be *empty*, *occupied* or *unknown*. The mass functions are merged to have a global mass function

$$m(P) = \bigoplus_{i \in I} m_i(P) \tag{2}$$

m(P) is the final mass that gives, for a point  $P \in \mathbb{R}^3$ , the confidence to be *empty*, *occupied* or *unknown* for the set of mass function I.

The problem is then defined as a binary segmentation of the space regarding to m where the label set is  $L = \{0, 1\}$  where l = 1 represents the state *occupied* and l = 0 the state *empty*. The space is discretized with an octree, where V is the set of octree's leaves and  $F \subset V^2$  the set of interface between adjacent leaves. We note  $l_v$  the label of the leaf  $v \in V$  and  $l_V = (l_v)_{v \in V}$  the labelling  $l_v$  of all leaf of the octree V.

Finally, the goal is to find, for each leaf, a label  $l_v$  that encodes the fact in being either *occupied* or *empty* for a given set of input mass functions  $m_i$ . To this end, the problem is formulated as an energy minimization framework composed of two terms: the data term  $E_{data}(l_V)$ , which represents how a label is close to the global mass function m and the smoothness term  $E_{prior}(l_V)$ , which pe-

This contribution has been peer-reviewed. Editors: M. Brédif, G. Patanè, and T. Dokken doi:10.5194/isprsarchives-XL-3-W3-613-2015



Figure 1: Result of the proposed methods : first the resulting surface, then textured with the regularized reflectance of the data.

nalizes solutions with large surfaces. The final energy is :

$$\min_{l_V \in L^V} E_{data}(l_V) + \lambda_s \ E_{prior}(l_V) \tag{3}$$

where  $\lambda_s$  is the weight balancing the prior and data terms. Now the two terms will be described in details.

## 2.1 Data term

The data term models how a label l fits to the global mass function m defined in equation 2. The confidence of a label l to be occupied is decomposed into 2 functions  $p_e(l)$ ,  $p_o(l)$  defined as:

$$p_i(l) = \begin{cases} 1-l & \text{if i=e} \\ l & \text{if i=o} \end{cases}$$
(4)

Finally, the data term is the integral of the difference between the function  $p_i(l)$  and the mass function  $m_i$  for each label  $i \in (e, o)$  over the volume of the leaf, which gives:

$$E_{data}(l) = \sum_{v \in V} \sum_{i \in \{e,o\}} \int_{v} |p_i(l_v) - m_i(x)| dx$$
 (5)

## 2.2 Prior term

The prior term generally forces solutions with a small surface. This leads to the following term:

$$E_{prior}(l) = \sum_{t_1, t_2 \in F^2} \operatorname{area}(t_1 \cap t_2) |l_{t_1} - l_{t_2}|$$
(6)

where  $\operatorname{area}(t_1 \cap t_2)$  returns the area of the interface  $t_1 \cap t_2$ . When  $l_{t_1}$  is equal to  $l_{t_2}$ , the score is equal to 0. Inversely, it is equal to 1 when  $l_{t_1}$  is different to  $l_{t_2}$ .



Figure 2: 2D example of the main scheme. The captured scene (a). The function related to the space occupancy : from a database (b), from ground and airborne laser information (c) (Red is outside, green is inside). Merging of source (d). inside/outside segmentation (e).

#### 2.3 Mass function definition

All that remains is to define the function  $m_e, m_o$  and  $m_u$  for an input set of data. With lidar data, when the center of the sensor is known, the visibility prior tells that cells crossed by the beam formed by the sensor center and the 3D point have great confidence to be *empty* and those right after the 3D point have great confidence to be *occupied*. This idea is encoded here in the independent definition of a mass function for each input measurement, which will be merged into a single over function using DST.



Figure 3: Mass function in  $R^2$  for a beam defined by the sensor center O and the 3D point Q. The parameters are: the incertitude on the angle  $\sigma_{\theta}$ , on the data  $\sigma_n$  and the thickness of the scene  $\sigma_t$ .

**2.3.1 Beam function** A ray is defined by the 3D detected point Q and the sensor center O. First, an occupancy mass along the ray direction is defined  $(e_r, o_r, u_r)$ . Let  $P \in R^3$  be the measured 3D point and P' the orthogonal projection of P on  $\overrightarrow{OQ}$ .

This contribution has been peer-reviewed. Editors: M. Brédif, G. Patanè, and T. Dokken doi:10.5194/isprsarchives-XL-3-W3-613-2015  $\overrightarrow{OQ}, \theta, r$  represents the relative position of the point P where

$$\theta = \arccos \frac{\overrightarrow{OQ}.\overrightarrow{OP}}{||\overrightarrow{OQ}||.||\overrightarrow{OP'}||} \tag{7}$$

is the angle formed by  $\overrightarrow{OP}$  and  $\overrightarrow{OQ}$  and

$$r = ||\overrightarrow{OP'}|| - ||\overrightarrow{OQ}|| \tag{8}$$

is the signed distance of the projected point P' to Q. When r < 0, P is between O and Q, this is *empty*. Consequently, when |r| is important, the mass  $e_r$  is equal to 1 and the *occupied* mass  $o_r$  is equal to 0. Inversely, when r > 0, P is behind Q, this is *occupied*, then the mass  $e_r$  is equal to 0 and the *occupied* mass  $o_r$  is equal to 1. When P is close to O,  $o_r$  and  $e_r$  tends to 0.5. When  $r < -|\overrightarrow{OQ}|$ , r is behind O, this is *unknown*, then  $e_r = o_r = 0$ .

Finally, behind Q,  $o_r$  is represented by decreasing function from 1 to 0 that models the thickness of the scene. The resulting mass functions when  $r > -|\overrightarrow{OQ}|$  are :

$$r > 0: e_r = 0.5 \ e^{-\left(\frac{|r|}{\sigma_n}\right)^2}; o_r = (1 - 0.5 \ e^{-\left(\frac{|r|}{\sigma_n}\right)^2}) \ e^{-\left(\frac{|r|}{\sigma_t}\right)^2}$$
$$r < 0: e_r = 1 - 0.5 \ e^{-\left(\frac{|r|}{\sigma_n}\right)^2}; o_r = 0.5 \ e^{-\left(\frac{|r|}{\sigma_n}\right)^2}$$
(9)

where  $\sigma_n$  is the scale of the uncertainties related to the sensor and  $\sigma_t$  the scale of the scene thickness prior.

The mass of occupancy is important on the ray while decreasing as it moves away from the ray. Following this assumption, the occupancy function is represented as

$$f(\theta) = e^{-\left(\frac{\theta}{\sigma_{\theta}}\right)^2} \tag{10}$$

where  $\sigma_{\theta}$  is the scale of the angle uncertainty. The mass function of a 3D point for a ray is then defined by:

$$m_i(P) = \left\{ \begin{array}{c} e \\ o \\ u \end{array} \right\} = \left\{ \begin{array}{c} f \cdot e_r \\ f \cdot o_r \\ 1 - e - o \end{array} \right\}$$
(11)

If the normal  $\overrightarrow{n_q}$  at Q is known, it can be added as an extra parameter. r is then redefined as the orthogonal distance from P' to the plane defined by the point Q and the normal  $\overrightarrow{n_q}$ . The figure 3 shows the mass function for a beam in 2D.



Figure 4: Example of fusion. First, four rays with an important uncertainty. The confidence at the intersection is much higher. Then, a fusion with both a strong confidence to be *empty* and *occupied*. The confidence at the intersection is equal to 0.5 for both input, the *unknown* label remains small.

**2.3.2** Global mass : Dempster-Shafer theory The DST is used to combine all inputs and gives the final mass m. Let  $m_1$   $m_2$  two mass functions, the fusion can be computed with the DST

as following:

$$\left\{ \begin{array}{c} e_1\\ o_1\\ u_1 \end{array} \right\} \oplus \left\{ \begin{array}{c} e_2\\ o_2\\ u_2 \end{array} \right\} = \frac{1}{1-K} \left\{ \begin{array}{c} e_1 \cdot e_2 + e_1 \cdot u_2 + u_1 \cdot e_2\\ o_1 \cdot o_2 + o_1 \cdot u_2 + u_1 \cdot o_2\\ u_1 \cdot u_2 \end{array} \right\}$$
(12)

where  $K = o_1 \cdot e_2 + e_1 \cdot o_2$  represents the conflict term. It means that when K is equal to 0, there is no conflict between inputs.

The combination rule is commutative and associative, enabling an order-independent aggregation of the measurements. At a location P for a set of mass functions  $m_i \in I$ , the overall mass is computed following the equation 2. Figure 4 shows two examples of mass merging.

## 3. REFLECTANCE ESTIMATION FORMULATION

Extra information as reflectance can be used for texturing the resulting mesh. Contrary to the image intensity, there is no straightforward way to texture the mesh with the lidar reflectance. To tackle this problem, we propose to labelize each leaf of the octree with the reflectance of the data and use a global formulation approach in order to regularize the reflectance where no information is provided. Finally, the reflectance of a facet is the reflectance of leaves on the surface that are labelized as inside.

The problem is defined as a multi-label segmentation of the space regarding to the reflectance r of the data where the label set is  $L = \{0, ..., n\}$ . The actual reflectance of our data is between -20 and 40. We thus do a linear mapping between the reflectance intensity and the label set.

We note  $lr_v$  the reflectance label of the leaf  $v \in V$  and  $lr_V = (l_v)_{v \in V}$  the labelling  $lr_v$  of all leaves of the otree V.

Two terms are used: the data term  $E_{data.refl}(lr_V)$ , which represents how a label is close to the reflectance of a leaf r(v) and the smoothness term  $E_{prior}(lr_V)$ , which penalizes solutions with a large reflectance difference in the neightbourhood. The energy is :

$$\min_{lr_V \in L^V} E_{data\_refl}(lr_V) + \lambda_r \ E_{prior\_refl}(lr_V)$$
(13)

## 3.1 Data term

The data term is the difference between the reflectance label lr and the reflectance information in each leaf r. For each leaf v, the reflectance average r(v) of all points that remain in the same cell is computed. The data term is then :

$$E_{data.refl}(l) = \sum_{t \in V^*} |l_t - r(v)| \tag{14}$$

Where  $V^*$  is the set of leaves where the reflectance r is defined.

#### 3.2 Prior term

The prior term generally forces solutions with closes neighbors labels. This leads to the following term:

$$E_{prior.refl}(l) = \sum_{t_1, t_2 \in F^2} \operatorname{area}(t_1 \cap t_2) |l_{t_1} - l_{t_2}| \qquad (15)$$

Thank to the metric function on the quadratic term, this energy remains sub-modular and thus, can be optimized in polynomial time.

This contribution has been peer-reviewed. Editors: M. Brédif, G. Patanè, and T. Dokken doi:10.5194/isprsarchives-XL-3-W3-613-2015

# 4. ALGORITHM

The algorithm is decomposed into four steps : First, the scene is segmented with the octree. Then for each leaf of the octree, the equation 5 is computed. The energy 3 is minimized with graphcut approach to compute the indoor/outdoor labeling. Finally, the reflectance of each leaf is computed and the solution regularized.

## 4.1 Octree

Thanks to the previous formulation, any 3-space partitioning could by used to discretize the aggregated mass function in order to yield a tractable combinatorial optimization problem. Ideally, this tessellation should be driven directly by the behaviour of the mass function itself, with lower densities where it is homogeneous and facets orthogonal to its gradients. For that, the space is discretized with an octree. For each 3D point, the octree is sub-divised until the maximum depht at the 3D point location is reached. The octree is also subdivided in the 8 neighbors of the point leaf in order to allows a good local estimation of the surface.

## 4.2 Score computation

Equation 5 requires to compute the integration of the mass function on each octree's leaf.

The integral  $\int_t |p_i(l_t) - m_f(s)| ds$  can be approximated using MC integration as  $\sum_{P \in S} \frac{\text{Volume}(v) * |p_i(l_t) - m_f(P)|}{|S|}$  where S is a set of |S| uniformly sampled points in the leaf v. Volume(v) is the volume of the leaf v. The mass function of each point P is computed with the equation 2.

### 4.3 Surface optimization

As mentioned in (Boros and Hammer, 2002), in our case where the quadratic function on the two labels is a metric, the equation 3 can always be solved globally in polynomial time in the binary case with a *min-cut* reduction. We use the graph-cut code introduced in (Boykov and Kolmogorov, 2004) to solve the binary problem.

## 4.4 Reflectance optimization

According to (Boykov et al., 2001), the equation 13 with a metric function on the prior term 15 is sub-modular and thus, the global minimum can be reached in polynomial time. One of the most efficient ways to solve this problem is the  $\alpha$ -expansion algorithm first introduced in (Boykov et al., 2001). It is fast and converges in practice on a good local minimum which is guaranteed to have an energy at most twice the globally optimal energy.

## 5. RESULT

We tested the proposed method with lidar data where 3D positions and reflectance are available. Figure 5 shows the result on different scene. Each scene has a  $128 \times 128 \times 128$  meter bounding box. The deepest leaf has a size of 0.03m which leads to a 12 levels of depth octree.  $\lambda_r = \lambda_s = 0.02$ . We can see that where 3D data are available, the surface is composed of small facets and are well estimated. Areas that have not been scanned are regularized with bigger leaf because of the lack of data. Challenging objects like gate and wire are also well estimated thanks to the higher level of detail. The regularized reflectance allows a good visualisation of the current mesh. The volumetric estimation of the reflectance allows modification like carving without extra computation.

## 6. CONCLUSION

A octree based framework is proposed for the watertight surface reconstruction of ubiquitous data in the same scene. The problem is formalized as a global optimization problem that is efficiently solved with a graph-cut approach. Results on real data shows the efficiency of the proposed method on challenging scenes.

## REFERENCES

Boros, E. and Hammer, P. L., 2002. Pseudo-boolean optimization. Discrete Appl. Math. 123(1-3), pp. 155–225.

Boykov, Y. and Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(9), pp. 1124–1137.

Boykov, Y., Veksler, O. and Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(11), pp. 1222–1239.

Fuhrmann, S. and Goesele, M., 2011. Fusion of depth maps with multiple scales. In: Proceedings of the 2011 SIGGRAPH Asia Conference, SA '11, ACM, New York, NY, USA, pp. 148:1–148:8.

Nexa, Gerkeb, Remondinoa, Przybillac, Bumkerc and Zurhorst, 2015. Benchmark for multi-platform photogrammetry. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-3/W4, 2015.

Snow, D., Viola, P. and Zabih, R., 2000. Exact voxel occupancy with graph cuts. In: Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, Vol. 1, IEEE, pp. 345–352.

### ACKNOWLEDGMENTS

This material is based upon work supported by the FP7 IQmulus Integrating Project, FP7-ICT-2011-318787, A High-volume Fusion and Analysis Platform for Geospatial Point Clouds, Coverages and Volumetric Data Sets.



Figure 5: Result of the proposed method. Left column : the resulting mesh. Right column : the mesh textured with the regularized reflectance of the data.