HAL
open science

3D watertight mesh generation with uncertainties from ubiquitous data

Laurent Caraffa, Mathieu Brédif, Bruno Vallet

To cite this version:

Laurent Caraffa, Mathieu Brédif, Bruno Vallet. 3D watertight mesh generation with uncertainties from ubiquitous data. Asian Conference on Computer Vision (ACCV), Nov 2016, Taipei, Taiwan. pp.377–391, 10.1007/978-3-319-54190-7_23 . hal-01882625

HAL Id: hal-01882625

https://hal.science/hal-01882625v1

Submitted on 27 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3D watertight mesh generation with uncertainties from ubiquitous data

Laurent Caraffa, Mathieu Brédif, Bruno Vallet

Université Paris-Est, IGN Recherche, SRIG, MATIS

Abstract. In this paper, we propose a generic framework for watertight mesh generation with uncertainties that provides a confidence measure on each reconstructed mesh triangle. Its input is a set of vision-based or Lidar-based 3D measurements which are converted to a set of mass functions that characterize the level of confidence on the occupancy of the scene as *occupied*, *empty* or *unknown* based on Dempster-Shafer Theory. The output is a multi-label segmentation of the ambient 3D space expressing the confidence for each resulting volume element to be *occupied* or *empty*. While existing methods either sacrifice watertightness (local methods) or need to introduce a smoothness prior (global methods), we derive a per-triangle confidence measure that is able to gradually characterize when the resulting surface patches are certain due to dense and coherent measurements and when these patches are more uncertain and are mainly present to ensure smoothness and/or watertightness. The surface mesh reconstruction is formulated as a global energy minimization problem efficiently optimized with the α-expansion algorithm. We claim that the resulting confidence measure is a good estimate of the local lack of sufficiently dense and coherent input measurements, which would be a valuable input for the next-best-view scheduling of a complementary acquisition.
Beside the new formulation, the proposed approach achieves state-of-the-art results on surface reconstruction benchmark. It is robust to noise, manages high scale disparity and produces a watertight surface with a small Hausdorff distance in uncertainty area thanks to the multi-label formulation. By simply thresholding the result, the method shows a good reconstruction quality compared to local algorithms on high density data. This is demonstrated on a large scale reconstruction combining real-world datasets from airborne and terrestrial Lidar and on an indoor scene reconstructed from images.

1 Introduction

Surface reconstruction from point cloud is an important topic that has already been studied extensively. The first reason is the ill-posed characteristic of the problem. The second reason is the constantly increasing number of applications that require surface reconstruction: digital elevation model computation for flood simulation, 3D modeling, robot path planning, etc. Furthermore 3D acquisition

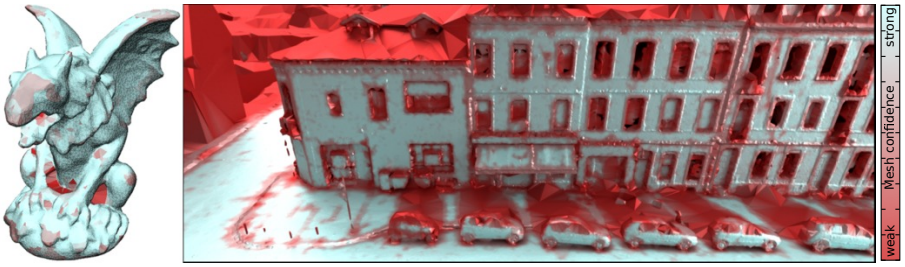


Fig. 1: Result of the proposed method on scanned object and city environment. Areas with strong confidence are blue while occluded or conflict areas are red.

techniques such as Lidar and multi-view stereo are now mainstream and currently harvest vast amounts of data on a daily basis. These two aspects result in multiple and various methods based on different assumptions related to the characteristics of the acquisition device and of the scene and to the user applications. This observation is pointed out by [1] that lists most of the commonly used assumptions for a given situation: smoothness prior, shape primitive reconstruction where the output is composed of a combination of primitives, visibility prior that uses the sensor information in order to detect empty areas, watertightness, etc. Because of the implicit relation between the surface definition and the smoothness prior, this assumption has been historically the most studied. Despite the important diversity of methods, the smoothness prior can be broadly categorized as local, global or piecewise. Local methods ensure that the output surface is smooth where the point cloud density is large enough. Global methods produce a watertight surface on the whole data range by filling empty holes regardless of the density. This fact is emphasized by the benchmark introduced in [2] where global methods based on indicator functions are robust with a small Hausdorff distance to the ground truth, while local methods produce a locally accurate reconstruction when the quality of the input data is good. When the density is large enough, it indicates that both approaches produce good results. Based on this consideration, we consider that rather than trying to improve the quality of the resulting mesh, we should instead characterize its quality and produce a qualified mesh where uncertain areas are localized. To this end, we use a global approach and pose the surface reconstruction problem as a volumetric segmentation of a 3D confidence space.

1.1 Background

The volumetric segmentation approach has been first used in [3] for surface reconstruction. In [4], it is used with a Delaunay triangulation and the visibility assumption for labelling tetrahedra crossed by a ray as outside and tetrahedra behind the 3D detected point as inside the surface to be reconstructed. Unlabelled tetrahedra are regularized in a global optimization framework related to

a surface smoothness prior. [5] shows that this approach is mature and produces high quality results when the amount of data is large enough in the multi-view reconstruction context. In [6], the visibility prior is improved to better reconstruct surfaces having a lack of information. This approach has been greatly extended in various cases like hybrid primitive mesh reconstruction [7] or structure preserving approach where coherent crust, plane and corner are detected for an extra sampling of the point cloud [8]. In [9], a similar approach is used to produce a watertight surface from airborne Lidar with planar assumption.

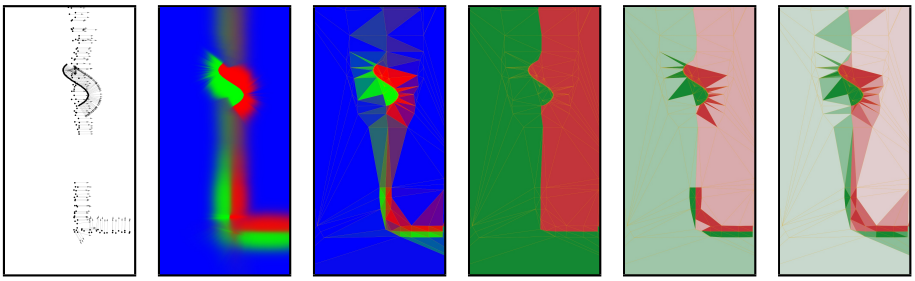
When the normal orientation or the visibility prior are not available, [10] proposes a method that defines an unsigned function related to the detection of crusts around high density point, which allows to reconstruct the surface without any normal information. In [11], a method is based on the winding number that produces a piecewise mesh when a significant number of facet orientations are available. The advantage is the definition of a function in the space that tells, on each point, its probability to be *occupied* or *empty*.

Recently, from the perspective to merge multiple source data, [12] propose an approach that takes into account multiple-scale datasets. It produce high quality surface in high detailed area embedded in higher scale reconstruction.

Even if modern algorithms are robust enough to handle bad quality inputs, some conflicts will always be impossible to prevent in uncontrolled environments like moving pedestrians or cars. Outliers and conflict detection from multiple point clouds has been recently studied. [13] use the Dempster-Shafer Theory (DST) for modeling the space occupancy as *empty*, *occupied* or *unknown*.

1.2 Proposed approach

One may argue that surface reconstruction in easy cases is now mature and many methods produce good results thanks to different priors in multiple cases ([5] for urban reconstruction or [14] for scanned objects). However, the amount and types of input are constantly increasing such that [1] points out acquisition ubiquity as a challenging task for the future. It is to that extent that we propose a new perspective for surface reconstruction: instead of choosing a strong surface prior and specific output, the method produces a segmentation of the space related to the confidence to be *empty* or *occupied*. The input of our algorithm is a set of mass functions that describe the occupancy of the scene with the Dempster-Shafer Theory. Instead of producing a binary segmentation, the result is a multi-label segmentation of the space where each label is a confidence to be inside or outside an object. The final surface is defined as a level set of this confidence at a given threshold of the interface between the mostly-*occupied* and mostly-*empty* volumes. Thanks to this multi-label formulation, the confidence of the output can be characterized as the gradient in occupancy confidence between the neighboring volumes of each output facet. Mesh patches composed by facets with greater confidence may be seen as patches featuring enough data support which result mainly from a local smoothing. By contrast, in uncertain areas, this confidence measure properly documents the fact that corresponding facets are



(a) (b) (c) (d) (e) (f)

Fig. 2: Example. First, two pointsets merged with their local descriptor (fig. 2a), the combined mass function of the two sets of points (fig. 2b), the normalized integral of the mass function on each cell (fig. 2c). The resulting segmentation with 2,4 and 8 labels (figs. 2d, 2e and 2f).

only present to ensure smoothness and watertightness with little data support as it may be the case in any global surface reconstruction approaches.

In this paper, we propose two models, one based on a local descriptor introduced in [15], the second on the visibility prior for surface reconstruction when the sensor position is known for each 3D measurement. The result of the proposed method is shown in the figure 1 (The redder the color, the smaller the confidence). One may note that the confidence increases with the point density. However, areas that have not been scanned or with an important scale remain red (i.e uncertain). To our knowledge, producing a multi-label result, according to the surface confidence has not been explored in the surface reconstruction literature.

Figure 2 shows an example of the proposed approach in 2D. First, a local descriptor is computed on each point cloud for extracting normal and local point cloud distribution(fig. 2a). The space is then discretized with a 3D Delaunay triangulation (as in [4]), the mass function (fig. 2b) is then integrated on each cell (fig. 2c). Finally, a multi-label segmentation is performed on each tetrahedron where a label relates to the confidence of the result to be *occupied* or *empty* (fig. 2d, 2e, 2f). This labelling relies on a global optimization that penalizes large surfaces. For the problem to be tractable, a mapping from labels {*occupied*, *empty*, *unknown*} to labels {*empty*, *occupied*} is performed. This allows to formalize the problem with a sub-modular energy that can be optimized globally in polynomial time. The multi-label segmentation is then conceptually cast as a binary *empty*/*occupied* segmentation by thresholding the label confidence at e.g. 50%, which enables the well-known extraction of the resulting mesh as the set of triangles separating an *empty*-labeled tetrahedron from a *occupied*-labeled one.

Our main contributions are:

- A method that takes an arbitrary number and type of input measurements defined by a confidence function.
- A multi-label segmentation that produce a watertight surface and enables the confidence characterization of each resulting mesh triangle.

For that, we will first define in section 2 the problem statement of the multi-label segmentation of the mass function. The energy formulation of the problem is defined, then the mass function for the case of the point cloud reconstruction is introduced. In section 3, technical aspect of the implementation is discussed: the space is quantification, the mass function computation and the optimization of the resulting energy. Then, the method is evaluated and compared to state of the art algorithms in section 4. Section 5 shows applications and results on mesh generation on both the ability to merge heterogeneous datasets and the confidence characterization of the resulting mesh.

2 Problem 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 2.2):

$$m_i(P) = (e, o, v) \in [0,1]^3 \text{ s.t. } e + o + u = 1 \tag{1}$$

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

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

where $m_f(P)$ is the overall mass that gives, for a point $P \in \mathbb{R}^3$, the confidence to be *empty*, *occupied* or *unknown* for the set of mass functions $(m_i)_{i \in I}$.

The problem is then defined as a multi-label segmentation of the space regarding to m_f where the labels set is $L = \{0, \frac{1}{(N-1)}, \frac{2}{(N-1)}, ..., \frac{(N-2)}{(N-1)}, 1\}$. Each label $l \in L$ represents the occupancy confidence where $l = 1$ represents a absolute confidence to be *occupied* and $l = 0$ a absolute confidence to be *empty*. The space is discretized with a 3D Delaunay triangulation, where T is its set of tetrahedra and $F \subset T^2$ the set of facets represented by pairs of adjacent tetrahedra such that the set of facets is $\{(t_1 \cap t_2) s.t. (t_1, t_2) \in F\}$. We denote as l_t the label of the tetrahedron $t \in T$ and $l_T = (l_t)_{t \in T}$ the labelling l_t of each tetrahedron in the triangulation T. Following this formulation, a per-facet confidence is defined as

$$\text{confidence}(t_1 \cap t_2) = |l_{t_1} - l_{t_2}| \tag{3}$$

Thus, what is required is to find, for each tetrahedron of the triangulation, a label l_t that encodes the confidence of the tetrahedron 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 $E_{data}(l_T)$, which represents how a solution is close to the overall mass function m_f and the smoothness term $E_{prior}(l_T)$, which penalizes solutions with large interfaces between labels. The final energy minimization is:

$$\min_{l_T \in L^T} E_{data}(l_T) + \lambda \, E_{prior}(l_T) \tag{4}$$

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

2.1 Energy terms

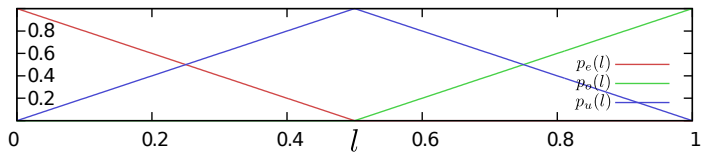


Fig. 3: Example of the related confidence of the 3 mass functions.

The data term models how a label l fits to the overall mass function m_f defined in equation 2. To take into account the *unknown* label, the confidence of a label l to be *occupied* is decomposed into 3 functions $p_e(l), p_o(l)$ and $p_u(l)$ defined as:

$$p_j(l) = \begin{cases} p_e(l) = max(0, 1-2l) \\ p_o(l) = max(0, 2l-1) \\ p_u(l) = 1 - p_o(l) - p_u(l) \end{cases} \tag{5}$$

Figure 3 shows the mapping between labels and the 3 functions. When $l = 0.2$, the corresponding scores are $p_e(l) = 0.6$, $p_o(l) = 0$, and $p_u(l) = 0.4$. With this decomposition, the case $p_e(l) = 0.6$, $p_o(l) = 0.4$ and $p_u(l) = 0.4$ are mutually exclusive solutions for the input $l = 0.6$. It tells that the solution is a conflict or a fuzzy zone. This allows to separate the *conflict* case from the *unknown* case. However, one can imagine another cost function for conflict, change detection, accuracy measurement.

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

$$E_{data}(l) = \sum_{t \in T} \sum_{j \in \{e,o,u\}} \int_t |p_j(l_t) - m_f(v)| dv \tag{6}$$

A prior term is added that generally enforces solutions with a small surface. This idea is extended to the multi-label case by minimizing the surface multiplied

by the distance between label probabilities, which leads to the following term:

$$E_{prior}(l) = \sum_{(t_1,t_2)\in F^2} \text{area}(t_1 \cap t_2)|l_{t_1} - l_{t_2}| \tag{7}$$

where $area(t_1 \cap t_2)$ returns the area of the triangle $t_1 \cap t_2$.

2.2 Mass function definition

All that remains is to define the function m_e, m_o and m_u for an input set of data. When the oriented normal is known, tetrahedra in the oriented normal direction have great confidence to be *empty* and those right after the 3D point have great confidence to be *occupied*. In the case of 3D reconstruction when the center of the sensor is known, the visibility prior tells that tetrahedra crossed by the beam formed by the sensor and the 3D point have great confidence to be *empty* along the ray. This idea is encoded here in the independent definition of a mass function for each input measurement, which will be merged into a single overall function using DST.

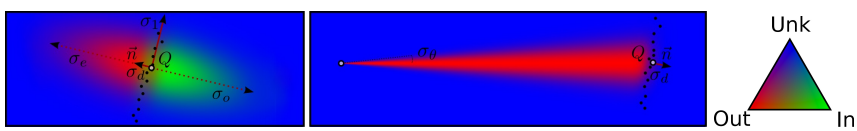


Fig. 4: Mass function in $\mathbb{R}^2$. (Left) For a single point Q, model parameters are: a spanning parameter σ_1 and the noise parameter σ_d. σ_e and σ_o are the thickness parameter. (Right) The mass with the visibility prior in $\mathbb{R}^2$ for a point Q when the sensor center is known, model parameters are: the incertitude on the angle σ_θ and the noise parameter σ_d.

In this section, two mass functions representing the space occupancy in $\mathbb{R}^3$ are defined: one representing the space occupancy around the 3D detected point with normal, the second representing the visibility prior when the sensor center is known.

Local mass definition: For a 3D detected point Q with a normal $\overrightarrow{n}$, an occupancy mass along the normal is defined by (e_r,o_r,u_r). Let $P \in \mathbb{R}^3$ be a 3D point and r its signed distance to the plane $(Q, \overrightarrow{n})$. When $r > 0$, P is on the same side as the normal, it is *empty*, e_r should tends to 1 and o_r to 0. Inversely, when $r < 0$, P, it is *occupied*, then the mass e_r should tends to 0 and o_r to 1. When P is close to Q, the probability to be *occupied* or *empty* o_r and e_r should be 0.5. Finally, when P is far from Q, o_r and e_r should decrease from 1 to 0 for modeling the thickness of the scene and the empty area. One mass related to this behavior is :

$$\begin{aligned} r < 0 : e_r &= 0.5\, e^{-(\frac{|r|}{\sigma_d})^2} \;;\; o_r = (1 - 0.5\, e^{-(\frac{|r|}{\sigma_d})^2})\, e^{-(\frac{|r|}{\sigma_o})^2} \\ r > 0 : e_r &= (1 - 0.5\, e^{-(\frac{|r|}{\sigma_d})^2}) e^{-(\frac{|r|}{\sigma_e})^2} \;;\; o_r = 0.5\, e^{-(\frac{|r|}{\sigma_d})^2} \end{aligned} \tag{8}$$

where σ_d is the noise parameter, σ_o the scale of the scene thickness prior and σ_e the thickness of the empty area.

Space occupancy mass:

In the 3D case, the mass of occupancy on the plane is large while decreasing as it moves away from the point, introducing a spanning parameter, the occupancy function is represented as

$$f(P, P^1, P^2) = e^{-(\frac{||P^1-P||}{\sigma_1})^2-(\frac{||P^2-P||}{\sigma_2})^2} \tag{9}$$

where $(\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{n})$ is a standard basis defined from the normal vector. P^1 and P^2 are the orthogonal projection on $\overrightarrow{v_1}$,$\overrightarrow{v_2}$. σ_1,σ_2 are the scale of the spanning parameter on the basis.

Adding the visibility prior: When the sensor center O is known, the mass of occupancy is large on the ray while decreasing as it moves away from it, for that we define

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

where θ is the angle formed by the ray $\overrightarrow{OQ}$ and $\overrightarrow{OP}$, σ_θ is the scale of the angle uncertainty. In this case, the scale of the *empty* area σ_e is set to ∞ and $o_r = 0$ when $r < 0$ in equation 8.

Mass of an input: Finally, the mass function of a 3D point is defined by:

$$m_i(P) = \begin{Bmatrix} e \\ o \\ u \end{Bmatrix} = \begin{Bmatrix} s_c \cdot f \cdot e_r \\ s_c \cdot f \cdot o_r \\ 1 - e - o \end{Bmatrix} \tag{11}$$

where e_r and o_r models the uncertainty along the normal and f the uncertainty as it moves away from the data. $s_c \in [0, 1]$ is the scale coefficient. The figure 4 shows an example of the two mass functions in 2D.

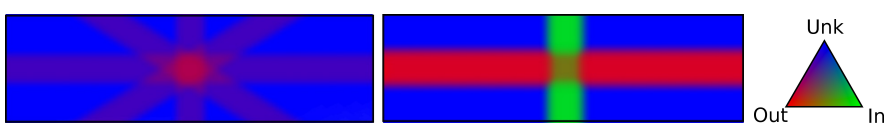


Fig. 5: Example of fusion. First, four rays with a large uncertainty. The confidence at the intersection is much higher. Then, a fusion with both a strong confidence to be *empty* and *occupied*. The confidence at the intersection is equal to 0.5 for both *empty* and *occupied*, the *unknown* label remains small.

Overall mass: The DST is used to combine all inputs and gives the overall mass m_f. Let m_1 m_2 two mass functions, the fusion can be computed with the DST as following:

$$\begin{Bmatrix} e_1 \\ o_1 \\ u_1 \end{Bmatrix} \oplus \begin{Bmatrix} e_2 \\ o_2 \\ u_2 \end{Bmatrix} = \frac{1}{1-K} \begin{Bmatrix} e_1 \cdot e_2 + e_1 \cdot u_2 + u_1 \cdot e_2 \\ o_1 \cdot o_2 + o_1 \cdot u_2 + u_1 \cdot o_2 \\ u_1 \cdot u_2 \end{Bmatrix} \tag{12}$$

where $K = o_1 \cdot e_2 + e_1 \cdot o_2$ represents the conflict term. 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 5 shows two examples of mass merging in 2D.

2.3 Setting parameters with local descriptor

The mass function needs local information like normal, noise level and scale. When a parameter is not provided, it can be extracted by computing the local descriptor proposed in [15]. Based on a local Principal Component Analysis (PCA), it provides 3 eigenvectors v_1, v_2, v_3 with their eigenvalues $\sigma_1, \sigma_2, \sigma_3$. Assuming a planar surface, its normal direction $\overrightarrow{n}$ is given by v_3, σ_3 denotes its noise level. σ_1 and σ_2 are the spanning scale for each scale of standard basis component.

The scale of the point is represented by its noise level $s_c = min(\frac{\epsilon}{\sigma_3}, 1)$ where ϵ is the targeted level of details. $\sigma_3 < \epsilon$ denotes that the highest level of details is reached.

3 Algorithm

3.1 Triangulation

Thanks to the previous formulation, any 3-space partitioning could by used to discretized the aggregated mass function in order to yield a tractable combinatorial optimization problem. Ideally, this tessellation should be driven directly by the behavior of m_f, with less refined tessellation where it is homogeneous and facets orthogonal to its gradients. To limit the computing time and be robust to noise, the number of points is reduced using a K-means approach on the point cloud to approximate the tessellation. To adapt the mesh accuracy to the noise level, points are added while the distance of closest plane of the triangulation in the normal direction for each input point is above $3\sigma_d$. This approach helps to produce mesh with large facet in noisy area and small facet in detailed areas according to the noise level. We currently use the Delaunay triangulation as implemented in the Computational Geometry Algorithms Library (CGAL, http://www.cgal.org).

3.2 Score computation

Equation 6 requires to compute the integration of the mass function on each tetrahedron. The integral $\int_t |p_j(l_t) - m_f(v)| dv$ can be estimated using Monte Carlo integration as $\frac{\text{Volume}(t)}{|S|} \sum_{P \in S} |p_j(l_t) - m_f(P)|$ where S is a set of $|S|$ uniformly sampled points in the tetrahedron t. Volume(t) is the volume of the tetrahedron t. The mass function of each point P is computed with the equation 2.

3.3 Optimization

According to [16], the equation 4 with a metric function on the prior term 7 is sub-modular and thus, the global minimum can be reached in polynomial time. Many methods exist for optimizing sub-modular functions. When the number of labels is small and the scene composed by large homogeneous label areas, the α-expansion algorithm first introduced in [16], converges in practice quickly on a good local minimum which is guaranteed to have an energy at most twice the globally optimal energy.

The α-expansion consists in starting from an arbitrary set of labels l_t in each tetrahedra $t \in T$ and solve iteratively equation 4 with a proposition $\alpha \in L$. It refers to a Boolean optimization problem studied in quadratic pseudo-Boolean optimization (QPBO). As mentioned in [17], in our case where the quadratic function on the two labels is a metric, the equation 4 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 [18] to solve the binary fusion problem.

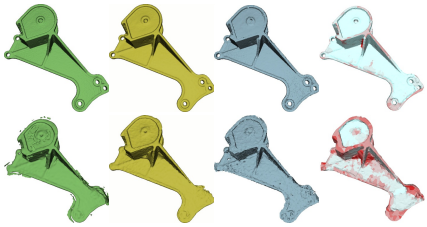
3.4 Parameter settings

The proposed formulation requires to choose an arbitrary number of labels. For surface reconstruction, an even number of labels should be chosen in equations 5: an odd number of labels results in the presence of the label $l = 0.5$ in L which hinders a decision between an *empty* and an *occupied* volume. With a small number of labels, the solution is smoother according to λ. With a large number of labels, extra areas with low confidence appear thanks to the small cost between the occupied and empty volumes in ambiguous regions. In our tests, 6 labels appears to be a good compromise between the time computation, the accuracy of the global result and the number of outliers.

4 Evaluation on benchmark

The method is evaluated on the benchmark introduced in [2], it allows to evaluate the surface according to both error metrics and topological aspects. The benchmark is dedicated to surface reconstruction from point cloud with normal. It is composed of 5 datasets generated from an implicit function with a synthetic scanner. For each dataset, 48 acquisitions are generated with different variations of the scanner's parameters such as noise or different camera positions which lead to occlusions. In order to show the advantage of computing a multi-label segmentation, a single result is computed for each file. For that, we use a confidence threshold T_c. A facet is removed when the confidence defined in equation 3 is lower than T_c. The resulting surface is non manifold. The larger T_c, the larger the certainty on the surface.

The result is then evaluated with two different confidence threshold $T_c = 0$ (full area) and $T_c = 0.5$ (confident area) that can be computed on the fly. In this test, the scale is set to the maximum level of detail ($s_c = 1$ for each point), the thickness to $\sigma_o = \sigma_e = 0.3$.



algorithm	comps	bnd	mani	genus
apss	47.37	140.86	0.50	1.82
fourier	1.54	**0.00**	**1.00**	**0.49**
scattered	1.90	214.21	**1.00**	7.47
wmwu ($T_c = 0$)	1.00	**0**	**1**	0.61
wmwu ($T_c = 0.5$)	2.72	77.08	0.02	12.45

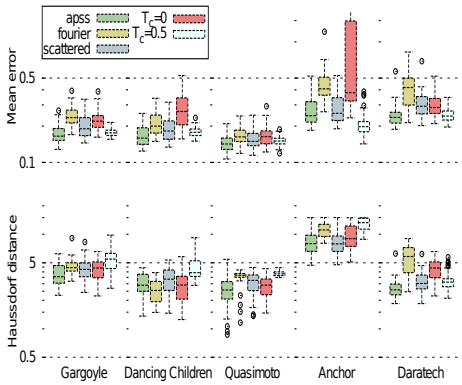


Fig. 6: Result of 4 algorithms on the benchmark: APSS, Fourier, Scattered and the proposed approach: the global surface (wmwu $T_c = 0$) and with the confident threshold set to 0.5 (wmwu $T_c = 0.5$). First line shows an easy case, the second case is much more noisy with occlusions. The error mean and the Hausdorff distance are compared on the 5 datasets. Black dots show outliers when the box plot show the median and quartiles. The table shows the average shape properties over the full benchmark: *comps* refers to number of connected components, *bnd* is the length of boundary components, *mani* is whether or not a mesh is manifold, 1 being manifold and 0 otherwise, *genus* refers to the amount which deviates from the actual genus. The first section refers to the state of the art algorithm evaluated in [2], the second section shows the score of the proposed methods with both thresholds.

4.1 Evaluation on the dataset

Figure 6 shows the result of the proposed method (wmwu) with the two thresholds ($T_c = 0$ and $T_c = 0.5$) compared to APSS [14], Fourier [19] and Scattered [20]. Two results are shown: on the first one, the resulting mesh of the proposed approach is highly confident (blue color) and results in an high quality mesh as other methods. The second case is much more uncertain because of a noisy sensor and occlusions (red areas) where other algorithms work also poorly. The global surface remains close to the ground truth. This fact is emphasized by the benchmark: when thresholding the mesh to $T_c = 0.5$, the mean error is very close to state-of-the art algorithms such as APSS or Scattered. The distribution of the means has globally a small variance and lowest maximum error on all the dataset that reflect the robustness of the proposed approach and the ability to detect uncertain facets. Like Scattered, the number of components remains low compared to APSS. Thanks to the multi-label global optimization framework that enforces the confident surface to be close, the boundary distance is 77.08 compared to 140.86 and 214.21 for APSS and Scattered.

The global mesh displays more error, this is the consequence of shapes produced by extra labels in unknown areas. The result remains good compared to

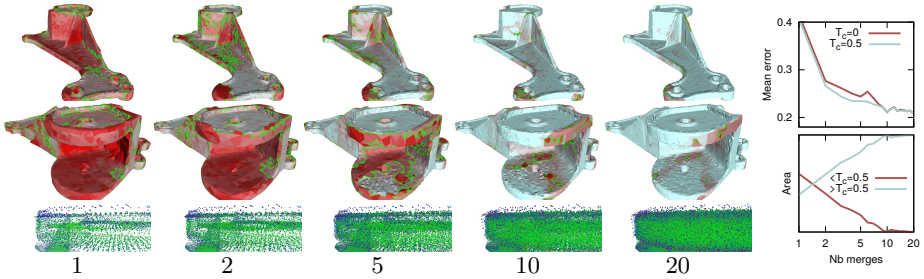


Fig. 7: Merging many data points. First and second line: the results on the union of 1, 2, 5, 10 and 20 point clouds. Third line: the accumulated point cloud. The confident threshold is set to $T_c = 0.5$. Facets in the confidence region that are on the boundary are green. Last column, the error of the global surface and the confidence surface. Then the area of regions upper and lower the threshold.

other algorithms in the benchmark. The advantage of extra labels makes sense by analyzing the Hausdorff distance which is very good compare to other algorithms. The extra labels allow to better explore the unscanned area as [6] thanks to a small regularization where classic graph-cut based approaches tend to over-smooth the solution and produce holes in missing data areas. This is confirmed by the characteristic that the global surface is 100% manifold and watertight with only one component on the whole benchmark.

4.2 Merging

The aim of the proposed approach is to merge a large amount of data. Figure 7 shows the result up to 20 point clouds processed at the same time on the Daratech dataset. For this test, the coefficient of each point cloud mass is 0.1 in order to reach the confidence threshold $T_c = 0.5$ only with a consequent number of data. We clearly show the confidence increasing while the number of files increases. Badly fitted surfaces caused by sensor drift on first iterations with high uncertainty are progressively removed and replaced by surface with high confidence. Statistics show that the confidence area is constantly increased along iterations while the mean error decreases until converging to 0.20 which is equal to the lower Quartile of best algorithms on this dataset. This test confirms the ability of the DST on managing uncertainty.

5 Results

Airborne and terrestrial Lidar reconstruction: This example shows how the DST helps for modeling different types of data. Figure 8 shows an urban scene reconstruction with both airborne and terrestrial Lidar data. Two overlapping

airborne Lidar acquisition stripes cover a large area. A terrestrial point cloud is included at the center of the scene for a final 5.2M point cloud.

Local model's parameters are fixed by the dimensionality descriptor (sec:2.3). Airborne and terrestrial Lidar point clouds are both represented as beams. The angle's scale of the visibility prior is set to 0.001 for the airborne and for the terrestrial data. The thickness parameter is equal to 0.07m for terrestrial Lidar to reconstruct thin structures such as poles, cars or advertising hoarding. It is set to 30m for airborne. This allows to reconstruct high buildings only with points recorded on the top. λ is equal to 0.005. The targeted scale is set to $\epsilon = 0.02$.

Results show that both datasets are taken into account to produce a mesh that combines both the fine details from the ground based Lidar and a global watertight surface from the airborne data. The large thickness parameter on airborne data allows to reconstruct high building. Thanks to the small thickness and extra points on the terrestrial data, pillars are well meshed. The close up view shows that the merging between airborne and terrestrial points is well managed. We clearly see how the point cloud density affects the confidence of the resulting mesh where the confidence scale is reached. In the ground scanned place, the quality of reconstruction of the building's facade is large whereas the areas behind cars remain unknown and call for extra acquisitions.

Fig. 8: Result of the proposed algorithm with terrestrial and airborne data combined.

Structure from motion data: We tested the proposed approach on indoor datasets from 3D reconstruction. A dataset of the full scene of 3.4M point is merged to a 5.4M points dataset representing a mechanical object. This scene has the particularity to have both high quality, noisy and unscanned areas. The targeted scale is set to $\epsilon = 0.03$. Figure 9 show the result of the proposed approach. Results show high quality reconstruction with high confidence in high detailed areas. On the contrary, uncertain areas are well smoothed with small confidence, unscanned areas are flat with the lowest confidence. The proposed approach is compared to FSSR [12], two appropriate scales are affected to each dataset. Where FSSR produces artifacts or holes, the proposed approach provides a low confidence and fills empty holes with the smallest confidence. The quality of the reconstruction is comparable in good quality areas.

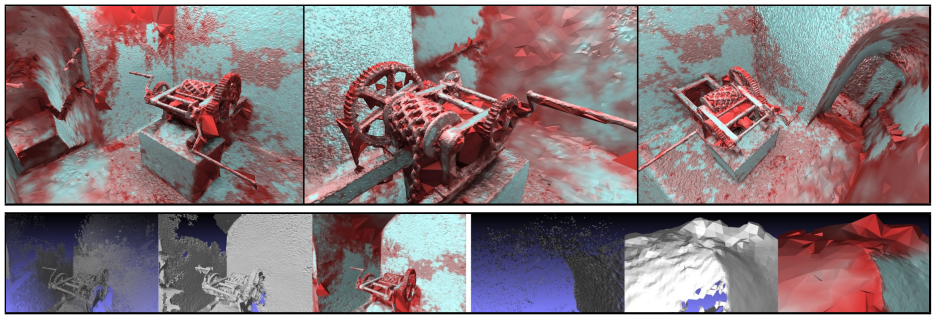


Fig. 9: Result of the proposed algorithm on structure from motion data. First line, 3 results of the proposed algorithm. Second line, two results with: the dataset, FSSR [12] and ours.

6 Limitations

The benchmark shows that the resulting mesh after thresholding is not watertight and manifold anymore in many cases. More complex energies could be explored in order to manage the watertightness of the mesh at different thresholds. The algorithm is actually slow, the time computation can be highly improved by parallelizing the triangulation construction, the sampling and the score computation. Parallel heuristics can be used for the multi-label optimization.

7 Conclusion and perspectives

A generic framework for watertight mesh generation with uncertainties is proposed for the surface reconstruction of ubiquitous data of the same scene. A confidence criterion is given on the resulting mesh. The problem is formalized as a global optimization problem that is efficiently solved with the α-expansion algorithm. The proposed approach shows results close to state of the art on both benchmark datasets and real cases like urban reconstruction from Lidar or structure from motion data. We advocate that the resulting confidence measure is a good estimate of the local lack of sufficiently dense and coherent input measurements, which would be a valuable input for the next-best view scheduling of a complementary acquisition or to detect unexplored area for terrestrial or airborne robot path planning.

Acknowledgments

This work is partly supported by the EU FP7 Project N. ICT- 2011-318787(IQmulus).

References

1. Musialski, P., Wonka, P., Aliaga, D.G., Wimmer, M., van Gool, L., Purgathofer, W.: A survey of urban reconstruction. In: EUROGRAPHICS 2012 State of the Art Reports. EG STARs, Eurographics Association (2012) 1–28
2. Berger, M., Levine, J.A., Nonato, L.G., Taubin, G., Silva, C.T.: A benchmark for surface reconstruction. ACM Trans. Graph. **32** (2013) 20:1–20:17
3. Curless, B., Levoy, M.: A volumetric method for building complex models from range images. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '96, New York, NY, USA, ACM (1996) 303–312
4. Labatut, P., Pons, J.P., Keriven, R.: Robust and efficient surface reconstruction from range data. Comput. Graph. Forum **28** (2009) 2275–2290
5. Vu, H.H., Labatut, P., Pons, J.P., Keriven, R.: High accuracy and visibility-consistent dense multiview stereo. IEEE Trans. Pattern Anal. Mach. Intell. **34** (2012) 889–901
6. Jancosek, M., Pajdla, T.: Multi-view reconstruction preserving weakly-supported surfaces. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. CVPR '11, Washington, DC, USA, IEEE Computer Society (2011) 3121–3128
7. Lafarge, F., Keriven, R., Brédif, M., Vu, H.H.: A hybrid multiview stereo algorithm for modeling urban scenes. IEEE Trans. Pattern Anal. Mach. Intell. **35** (2013) 5–17
8. Lafarge, F., Alliez, P.: Surface reconstruction through point set structuring. Rapport de recherche RR-8174, INRIA (2012)
9. van Kreveld, M.J., van Lankveld, T., Veltkamp, R.C.: Watertight scenes from urban lidar and planar surfaces. Comput. Graph. Forum **32** (2013) 217–228
10. Hornung, A., Kobbelt, L.: Robust reconstruction of watertight 3d models from non-uniformly sampled point clouds without normal information. In: Proceedings of the Fourth Eurographics Symposium on Geometry Processing. SGP '06, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association (2006) 41–50
11. Jacobson, A., Kavan, L., Sorkine-Hornung, O.: Robust inside-outside segmentation using generalized winding numbers. ACM Trans. Graph. **32** (2013) 33:1–33:12
12. Fuhrmann, S., Goesele, M.: Floating scale surface reconstruction. ACM Trans. Graph. **33** (2014) 46:1–46:11
13. Xiao, W., Vallet, B., Paparoditis, N.: Change detection in 3d point clouds acquired by a mobile mapping system. ISPRS Annals Photogramm Remote Sensing Spat Inf Sci II (5/W2) (2013) 331–337
14. Guennebaud, G., Gross, M.: Algebraic point set surfaces. In: ACM SIGGRAPH 2007 Papers. SIGGRAPH '07, New York, NY, USA, ACM (2007)
15. Demantké, J., Mallet, C., David, N., Vallet, B.: Dimensionality Based Scale Selection in 3d LIDAR Point Clouds. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **3812** (2011) 97–102
16. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. **23** (2001) 1222–1239
17. Boros, E., Hammer, P.L.: Pseudo-boolean optimization. Discrete Appl. Math. **123** (2002) 155–225
18. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004) 1124–1137

19. Kazhdan, M.: Reconstruction of solid models from oriented point sets. In: Proceedings of the Third Eurographics Symposium on Geometry Processing. SGP '05, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association (2005)
20. Ohtake, Y., Belyaev, A., Seidel, H.P.: An integrating approach to meshing scattered point data. In: Proceedings of the 2005 ACM Symposium on Solid and Physical Modeling. SPM '05, New York, NY, USA, ACM (2005) 61–69