On Fast Surface Reconstruction Methods for Large and Noisy Point Clouds

Zoltan Csaba Marton, Radu Bogdan Rusu, Michael Beetz

Intelligent Autonomous Systems, Technische Universität München

{marton,rusu,beetz}@cs.tum.edu

Abstract— In this paper we present a method for fast surface reconstruction from large noisy datasets. Given an unorganized 3D point cloud, our algorithm recreates the underlying surface's geometrical properties using data resampling and a robust triangulation algorithm in near realtime. For resulting smooth surfaces, the data is resampled with variable densities according to previously estimated surface curvatures. Incremental scans are easily incorporated into an existing surface mesh, by determining the respective overlapping area and reconstructing only the updated part of the surface mesh. The proposed framework is flexible enough to be integrated with additional point label information, where groups of points sharing the same label are clustered together and can be reconstructed separately, thus allowing fast updates via triangular mesh decoupling. To validate our approach, we present results obtained from laser scans acquired in both indoor and outdoor environments.

I. INTRODUCTION

Autonomous mobile robots need continuous surface models of real-word scenes for navigation and manipulation. The accuracy of these models (often referred to as *maps*) is crucial for 3D collision detection while moving along a planned path, but also for supporting environment modelling and scene classification. Moreover, they need to have low computational requirements, as a mobile robotic platform has limited computational power capabilities which need to be shared between its different software components (e.g. perception, navigation, task planning, manipulation, etc).

Due to the latter constraint, most of the robotic navigation systems today rely on 2D discrete representations of the world, such as large scale 2D occupancy maps. This is in contrast with applications in computer graphics – like the reconstruction of architectural heritage [1] or 3D watertight objects [2] – which build extremely accurate models whose appearance has to be as close as possible to the original surface, but at a much smaller scale.

Recent advances in 3D sensing devices made their deployment possible on mobile robotic platforms, and enabled the acquisition of large 3D point cloud datasets (PCD). However, these datasets represent discrete, sampled representations of the world, so a transformation into continuous surfaces is required. To this end, we investigate the following computational problem: given a noisy 3D point cloud model of a real-world scene as depicted in Figure 1, create a continuous triangular mesh surface representation which reliably represents the scene, as fast as possible.

Our approach for creating the triangular mesh is based on the *incremental surface growing* principle (as devised by [3]), because it can deal with all types of surfaces where points

Fig. 1. Top: point cloud dataset of an indoor kitchen environment (15 millions of points) shown in grayscale intensity. Bottom: its continuous mesh representation (downsampled for visualization).

on two different layers can be distinguished by a distance criterion. Our algorithm selects a starting triangle's vertices and connects new triangles to it until either all points are considered or no more valid triangles can be connected. In the second case a new seed triangle is placed in the unconnected part and the triangulation is restarted.

In contrast to well known surface reconstruction methods from computer graphics [4], [5], [6], [7], our methods works reliably on noisy 2.5D data scans acquired by mobile robots navigating in indoor and outdoor environments.

The proposed surface reconstruction algorithm is partially based on the triangulation algorithm presented in [4]. The method has the advantage that it is a greedy type approach, where edges are written directly and are never deleted, making it fast and memory efficient, yielding the first correct triangulation.

Triangulation is done incrementally: i) for each point p, a k-neighborhood is selected by searching for the point's nearest k neighbors in a sphere with radius $r = \mu \cdot d_0$ that adapts to the local point density (d_0 is the distance of p to its closest neighbor and μ is a user-specified constant); ii) the neighborhood is projected on a plane that is approximately tangential to the surface formed by the neighborhood and ordered around p; iii) then the points are pruned by visibility and connected to p and to consecutive points by edges, forming triangles that have a maximum angle criterion and an optional minimum angle criterion.

The projection along the estimated surface normal occurs in the last two steps previously mentioned. In [4], this normal is estimated as the average of the normals of the triangles that have p as one of their vertices. This simplification makes sense only if the underlying surface is assumed to be smooth, that is the deviation between the normals of any two incident triangles on a vertex is less than 90°. The projection of the

neighborhood along the computed normal is then ordered around p, and the points that have to be connected to it are identified based on their visibility with respect to previous triangle edges (here only the current boundaries of the mesh have to be considered). The remaining points are connected to p in consecutive pairs, forming triangles.

The advantage of this method is that it preserves all points and makes no interpolation, over which we want to have greater control in our algorithm (see Section III). However, for noisy datasets the smoothness and locally uniform sampling constraints presented in [4] do not hold. In our implementation, this problem is solved, and we adapt the resulting triangle sizes to the surface's properties. A similar method is presented in [5], though it suffers from the limitation of not being able to deal with sharp features, and it only works with densely sampled point sets.

Our hole filling approach is similar to the one presented in [8], but we identify holes automatically and grow the triangle mesh into it by the use of robustly fitted vertices.

The main contributions of our surface reconstruction method include the following ones:

- *adaptability to variable densities*. Our method does not assume smooth surfaces or locally uniform neighborhoods, and can adapt to the point density variations present in 2.5D datasets. See Section III.
- *near realtime performance*. We employ fast kd-tree searches for nearest neighbor computations [9] and optimized lookups of previously computed triangle edges, to achieve extremely fast visibility checks (i.e. tests for the intersections of triangle edges). See Section VI.
- *noise robustness*. Estimated surface normals at a point are computed using a robust Weighted Least Squares method on the full neighborhood (see Section II), rather than averaging the normals of the current adjacent triangles as done in [4].
- *memory efficiency*. For each point we propagate a front wave, updated continuously, containing the point's associated advancing triangular edges. This has the advantage that we do not have to remember all the triangles in the entire dataset, and the lookup of edges that have to be checked for visibility is performed locally (as opposed to [4] – see Section II).
- *fast updates for incremental scans*. By determining the overlapping area between each new registered data scan and the existing surface mesh, our method is able to grow the existing model without recreating the entire triangular mesh. See Section IV.
- *supporting point labels*. The proposed surface reconstruction framework is flexible with respect to a local triangulation stopping criterion, in the sense that previously determined point labels can be used to decouple triangular meshes. An example is presented in Figure 7, where a distance connectivity criterion was used to segment and label objects lying on a horizontal planar surface, and thus resulting in separate triangular meshes for each object. See Section V.

The remainder of this paper is organized as follows. The next section introduces the theoretical foundations for our surface reconstruction method and discusses differences and improvements over similar work. In Section III we present our adaptive and robust resampling technique. Section IV discusses our improvements for achieving fast updates in incremental scans, while Section V presents a special case for triangular mesh decoupling using point labels. In Section VI we present results obtained on noisy indoor and outdoor datasets. Finally, we conclude and give insight on our future work in section VII.

II. THEORETICAL CONSIDERATIONS

The input to our algorithm is an unorganized 3D point cloud P obtained from laser sensing devices, representing real-world scenes or objects, with no constraint on them being watertight or complete. The data can contain occlusions and may be a combination of several registered partial views. In particular, in our work, we use an autonomous personal robotic assistant operating in household environments [10], and acquire point clouds by sweeping motions of an arm with a 2D laser sensor mounted on it (see Figure 1). The obtained data represents both large environmental models such as complete rooms, but also smaller segmented objects lying on horizontal planar surfaces (e.g. tables, cupboard shelves, counters) which may be manipulated by the robot. These particularities of our application scenario, require a set of constraints on our surface reconstruction system's behavior:

- since the robot acquires data in stages (i.e. partial scans of the world), whenever new data scans are available, the current existing mesh needs to be updated efficiently (see Section IV);
- to support dynamic scenes, such as objects being moved around from one place to the other, a mechanism for decoupling and reconstructing the mesh as fast as possible needs to exist, and separate objects should be triangulated separately (see Section V);
- since the input data contains high levels of noise, measurement errors, and holes, and to ensure that there are enough points in the neighborhood to capture all the relevant details in the reconstructed surface, a resampling step with variable vertex densities needs to be available (see Section III).

The last step is especially useful for large datasets when a compact representation is needed for high variations in surface curvature, requiring higher vertex densities on sharp edges and lower on flat surfaces.

In order to deal with the special requirements posed by our application and to achieve near real-time performance, we employ several techniques and optimizations. An important difference between our approach and [4] is that our method works directly on the 3D point cloud data as we facilitate fast nearest neighbor searches with kd-trees, and avoid creating an additional dexel structure, which requires a two step search through the entire dataset.

Additionally, since we deal with large datasets, we optimize the lookup of edges that have to be checked for pruning points by the visibility criterion in a k-neighborhood.
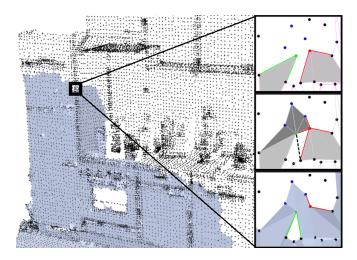
Fig. 2. Two fronts touched in a single point (red) resulting in multiple incident front edges for that point, which was solved by adding an extra triangle (black dashed line) – see picture in right-bottom. The green point is the current source for triangulation, the blue and red points are the points in its neighborhood, the green lines are the front edges of the green point, while the red ones are for the red point. The light gray triangles are the ones originally connected to the green and red points, while the dark ones would be the only new triangles if no correction would occur, producing four front edges connected to the red point – see picture in right-middle.

Whenever a new triangle is generated, the two neighbors that are connected to each fringe point (i.e. a point on the boundary of the current mesh) through the edges of the advancing front are saved. This creates a redundancy that is required for cases when only one of the two vertices is in a neighborhood, but speeds up the selection of the edges since we find them locally, by looking at the fringe points.

This has the advantage of automatically providing a closed loop of boundary points (i.e. points lying on the boundary of the underlying surface) on the sides of the surface and around its internal holes. We use these points to fill small holes in the data, as explained in Section III.

The disadvantage of this optimization is that it creates special cases which have to be solved locally when the advancing fronts touch in a single point (this happens at $\approx 6\%$ of the points in a dataset). We solved this problem by connecting additional points until each point has only two front edges, thus avoiding to run into the same situation again (see Figure 2 right, from top to bottom).

Due to the amount of noise present in our datasets, the modelled surface is not smooth, and averaging normals of incident triangles (as done in [4]) in a point $p \in P$ to obtain an approximation of the surface normal gives unsatisfactory results. To account for noise, we compute a weighted least squares plane in the p's k-neighborhood, and use the normal of the plane as an estimate of the true surface normal. We assemble a weighted covariance matrix from the points p_i of the neighborhood (where $i = 1...k$):

$$C = \sum_{i=1}^{k} \xi_i \cdot (p_i - \overline{p})^T \cdot (p_i - \overline{p}), \ \overline{p} = \frac{1}{k} \cdot \sum_{i=1}^{k} p_i$$

followed by the eigenvector V and eigenvalue λ computation $C \cdot V = \lambda \cdot V$. The term ξ_i represents the weight for point p_i: $\xi_i = exp(-d_i^2/\mu^2)$ – where μ is the mean distance from the query point p to all its neighbors p_i, and d_i is the distance from point p to a neighbor p_i. Optionally, we refine the normal by fitting a bivariate polynomial to the neighborhood (as in [11]). This extra step is computationally more expensive than averaging triangle normals, but this way triangles don't have to be stored and looked up, rendering our approach memory efficient.

Note that the pruning of small triangles is mandatory for our application because noisy real world scenes are not smooth, and thus vertices of a small triangle could appear as "flipped" in different neighborhoods, creating small inconsistencies in the triangulation.

Finally, we are using our RMLS algorithm [10] to achieve locally uniform sampling where needed by resampling the underlying surface while preserving its edges. The level of smoothing can be adjusted and it produces adaptive vertex densities for triangulation, as presented in Section III.

III. VARIABLE VERTEX DENSITIES

The triangulation method is capable of correctly adapting to locally uniform sampling. If the point density is variable, but there are smooth transitions between the different densities, the local neighborhood search radius adapts to it and makes sure that no point is left out. When distances between the scan lines are substantially different than the distances between the points in a scan line, the data does not satisfy this sampling criterion (see Figure 3).
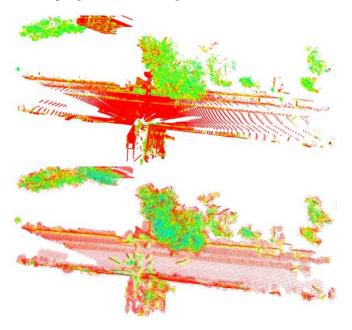

Fig. 3. Top: 3D scan of an urban scene with 497220 points (top view) and non-uniform local point density; bottom: resampled scene with 289657 points and even sampling distribution and filled holes. The colors represent the estimated surface curvatures (red meaning low, and blue high curvature).

To ensure a correct triangulation, extra points need to be interpolated between the existing ones. We are using our

Robust Moving Least Squares algorithm [10] to resample the point data, by combining a robust sample consensus plane computation with bivariate polynomial fits. Figure 4 presents the results obtained after resampling an outdoor urban scene to achieve a local uniform density using RMLS.
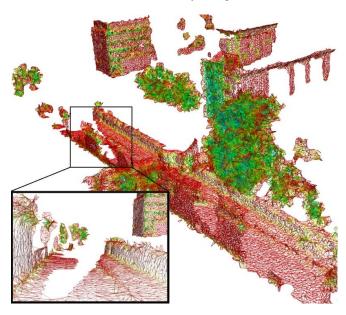

Fig. 4. Triangulation of the urban scene from Figure 3 (downsampled for visualization) with a closeup on the far side of the street, where the triangulation fails in the original dataset. Colors represent estimated surface curvature (red meaning low, and blue high curvature).

To efficiently generate the set of vertices for triangulation, each point $p \in P$ is snapped to a grid and within a fixed radius, points are generated around it to cover small holes the datasets might have. The snapping and a discretization of the density make this approach efficient, as points that were already generated can be easily filtered.

Resampling has another advantage that we can exploit: by influencing the density of the generated vertices based on the estimated curvature of the underlying surface [12], more points can be fitted to regions where more detail is needed. The density for the average estimated surface curvature in the dataset is specified by the user and for other values it is linearly interpolated with a given slope (see attached video for details). This ensures that the reconstruction will be finer in these areas (see Figure 5), and thus *adaptive* with respect to the underlying surface curvature.

These vertices are projected onto the plane D computed robustly with a sample consensus method [13], thus bringing them in the proximity of the underlying surface. Each point q from this projected set Q is fitted to the surface which approximates P by a bivariate polynomial height function along the normal of D.

If u, v and n are coordinates in the local neighborhood, we can define a number of nr functions f^i which approximate the surface in the proximity of q such as:

$$n_{(u,v)} = \sum_{i=1}^{nr} c_q^i \cdot f_{(u,v)}^i$$

To compute the coefficients c_q^i, we minimize the error function as in [8]. The functions are selected as members of bivariate polynomial functions of second order: $f_{(u,v)}^i = u^a \cdot v^b$, where $a, b \in \{0, 1, 2\}$. Depending on the required accuracy, higher orders can be used at the cost of slower computation time. However, our experience shows that orders greater than 3 have only a minimal effect on the result.
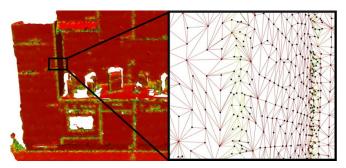

Fig. 5. Triangulation of a surface with adaptive vertex density. Left: triangulated surface colored based on estimated surface curvature (red meaning low, and green high curvature). Right: close-up on an edge where the vertices generated by RMLS are denser than on planar areas.

An additional hole filling strategy is to use the known holes from the internal boundary loops that result from triangulation. We automatically fit extra vertices to the covered area if its size is smaller than a user-defined threshold, and advance the boundary loop further to close the hole.

IV. INCREMENTAL DATA ACQUISITION

To construct complete models from partial views, the existing model needs to be updated by aligning it with newly acquired data scans. In our previous work, we have presented an efficient method for point cloud registration [14]. These scans are overlapping with the current surface mesh, and all the new points have to be incorporated into the model. However, neglecting all the previously constructed triangulations and creating a new mesh containing all the points each time a new scan is acquired does not constitute a feasible solution.

To solve this problem, our method triangulates only the new vertices which influence a change *into* the existing mesh. By determining the overlapping area between each new data scan and the current surface mesh, and removing those points which are close to existing triangle faces, we can grow the existing model without recreating the entire triangular mesh.

In order to include points that are in the vicinity of a triangle into the mesh, the corresponding triangle is *neglected*, and is treated as a hole containing the new points inside it, by initializing the advancing front with its vertices. The method is similar to the normal initialization, where a seed triangle is *filled* and the advancing front is initialized in the same way, with its vertices. This guarantees that no change is required in the actual triangulation algorithm – it is basically a restart of the previously finished step. The points in the new parts of the scene are also straightforward to include, as the previous step's boundary loop has to be treated as the advancing front that will expand into the new area.
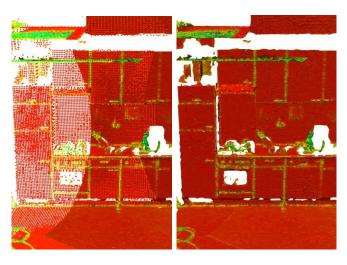
Fig. 6. Left: a new registered data scan has been acquired and the mesh model needs to be updated; right: the results after incremental triangulation, by preserving the mesh in the parts where no data update exists, and triangulating the new vertices. The colors represent the estimated surface curvature (red meaning low, and green high curvature).

planar areas based on their estimated surface curvatures or generating separate meshes for separate objects in the environment (see Figure 7).

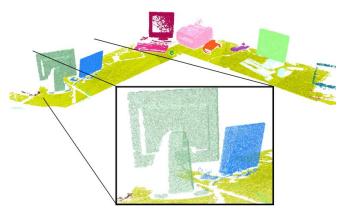Fig. 7. Separate meshes generated for different objects supported by planar horizontal areas (tables) in a point cloud dataset.

Figure 6 presents the result of these re-initialization methods in the kitchen scene which was extended with a new scan performed by the robot. To obtain a more compact representation, the high density of points can be reduced by adaptively resampling the surface in the overlapping area.

V. TRIANGULATION FOR DYNAMIC SCENES

In our previous work [10], [15] we have presented a system that can label different parts and regions of an indoor environment, based on their geometric and functional properties. For applications where fast model updates are required, for example while the robot is manipulating objects in the environment, our surface reconstruction methods make use of these additional labels to check which parts of the surface model are relevant for update.

For example, by detecting big horizontal planar areas and labeling them as tables [10], all the remaining point clusters that are supported by them can be labeled as possible objects. These segmented object candidates usually have the property that they are movable, so it would be useful if their mesh representation would be decoupled from that of the environment, thus making future updates of their position/orientation insensitive with respect to the connectivity of the vertices.

This is achieved by adding two different filters into the surface reconstruction algorithm:

1) first, the seed points that can be used as the vertices of a starting triangle on a new surface are filtered based on their point labels (e.g. a point label can be an attribute which classifies the primitive surface that supports this point [15]);
2) second, the points in a neighborhood are filtered based on their similarity to the current seed point.

Implementing the above filters results in the possibility to stop the current triangulation at an arbitrary condition, depending on the user's requirements, like triangulating only

VI. DISCUSSIONS AND EXPERIMENTAL RESULTS

To evaluate our system's efficiency, we performed several tests using multiple datasets. Since the input datasets are comprised of millions of points with dense neighborhoods at small scales, it would be extremely inefficient to reconstruct the surface using the complete data. To obtain sensible data resolutions, the datasets are resampled with RMLS using variable vertex densities. The results are presented in Figure 8.

Dataset	points	triangles	time
Kitchen front (Fig. 5)	24378	45828	0.816 sec
Kitchen 360° (Fig. 1)	40242	71436	2.347 sec
Urban scene (Fig. 4)	65646	114452	8.983 sec
Radiohead (Fig. 9)	329741	546516	17.857 sec

Fig. 8. Performance evaluation of our surface reconstruction method for multiple datasets.

The timings presented in Figure 8 represent the actual time our methods took to compute the surface mesh for each dataset, excluding any I/O operations. The hardware used was a standard Centrino Core2Duo notebook at 2Ghz with 2GB of RAM, running Debian Linux. Depending on the desired mesh accuracy, the input dataset could potentially be downsampled even more resulting in faster surface mesh computation times.

The bottom left part of Figure 9 presents the correct triangulation by our algorithm for a scan where the density of points in a scan line is much higher than the distance between the scan lines. The triangulation method presented in [4] failed to correctly triangulate the dataset since the locally uniform sampling criterion was not met. Even with $\mu = 5$ and a running time of 1134.36 seconds the result is not satisfactory as seen in the bottom right.

The datasets presented in Figure 8 are all formed of one part and were all reconstructed using the entire data as a

whole, thus leaving little space for further optimizations, like for example a parallelization of the triangles computation. In contrast, the dataset presented in Figure 7 is comprised of 22 parts with 252719 points total, where a part is defined by a subgroup of points which share the same attribute or label. A triangulation of the entire dataset without using the point labels results in a total computation time of 14.842 seconds. However, by using the extra information present in the data, we can create each triangle mesh individually, and thus parallelize the results. By performing the reconstruction on 4 different parts at once, the average computation time was brought down to 3.905 seconds. These results are encouraging for the usage of additional point labels not only for dynamic scenes as presented in Section V, but also for generating the surface meshes faster by partitioning and merging them, using the approach presented in Section IV.
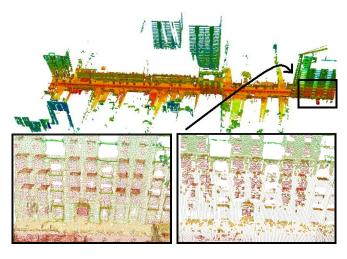

Fig. 9. Scan of an approximately 500m long stretch of an urban street resampled and triangulated (top) and closeup (bottom left). The triangulation presented in [4] failed because the dataset does not respect the local uniform sampling criterion (bottom right).

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a fast surface reconstruction method for large and noisy point cloud datasets. Our approach combines robust resampling methods with a fast projection-based triangulation method to achieve adaptability to variable point densities, and noise robustness. To cope with incremental data scans, we derive the overlapping areas between existing surface meshes and the new datasets, and update the existing model without recreating the entire mesh.

The performance of our approach was evaluated on datasets coming from both indoor and outdoor environments, and the results are promising for online surface generation for 3D path planning and collision detection on a mobile robotic platform.

Additionally, our reconstruction method supports extra information present in the dataset, such as point labels or attributes which can be grouped to cluster subsets of points together. This concept was validated in Section V and

provides a good starting point for online, fast mesh updates for dynamic scenes.

Future plans include improving the robustness of our algorithm by better handling special triangulation cases due to the high degree of noise. In this respect, we are targeting the usage of our methods on point cloud data generated by TOF (Time Of Flight) cameras.

REFERENCES

[1] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk, "The Digital Michelangelo Project: 3D Scanning of Large Statues," in *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 2000, pp. 131–144.

[2] M. Gopi, S. Krishnan, and C. T. Silva, "Surface Reconstruction Based on Lower Dimensional Localized Delaunay Triangulation," in *Computer Graphics Forum (Eurographics 2000)*, vol. 19(3), 2000.

[3] R. Mencl and H. Müller, "Interpolation and approximation of surfaces from three-dimensional scattered data points," in *State of the Art Reports, Eurographics '98*, 1998, pp. 51–67.

[4] M. Gopi and S. Krishnan, "A Fast and Efficient Projection-Based Approach for Surface Reconstruction," in *SIBGRAPI '02: Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing*, 2002, pp. 179–186.

[5] C. E. Scheidegger, S. Fleishman, and C. T. Silva, "Triangulating point set surfaces with bounded error," in *SGP '05: Proceedings of the third Eurographics symposium on Geometry processing*, 2005, p. 63.

[6] N. Amenta, S. Choi, and R. K. Kolluri, "The power crust," in *SMA '01: Proceedings of the sixth ACM symposium on Solid modeling and applications*. New York, NY, USA: ACM Press, 2001, pp. 249–266.

[7] T. K. Dey and S. Goswami, "Tight cocone: a water-tight surface reconstructor," in *SM '03: Proceedings of the eighth ACM symposium on Solid modeling and applications*, 2003, pp. 127–134.

[8] J. Wang and M. M. Oliveira, "A Hole-Filling Strategy for Reconstruction of Smooth Surfaces in Range Images," *XVI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)*, vol. 00, p. 11, 2003.

[9] S. Arya and D. M. Mount, "Approximate Nearest Neighbor Searching," in *Proc. of 4th Ann. ACM-SIAM Symposium on Discrete Algorithms (SODA'93)*, 1993, pp. 271–280.

[10] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, "Towards 3D Point Cloud Based Object Maps for Household Environments," *Robotics and Autonomous Systems Journal (Special Issue on Semantic Knowledge)*, 2008.

[11] M. Alexa, S. Rusinkiewicz, M. Alexa, and A. Adamson, "On normals and projection operators for surfaces defined by point sets," in *In Proceedings of Eurographics Symposium on Point-based Graphics, Eurographics*, 2004.

[12] M. Pauly, M. Gross, and L. Kobbelt, "Efficient Simplification of Point-Sampled Surfaces," in *Proceedings of IEEE Visualization*, 2002.

[13] P. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, pp. 138–156, 2000.

[14] R. B. Rusu, N. Blodow, Z. C. Marton, and M. Beetz, "Aligning Point Cloud Views using Persistent Feature Histograms," in *Proceedings of the 21st IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Nice, France, September 22-26*, 2008.

[15] R. B. Rusu, Z. C. Marton, N. Blodow, and M. Beetz, "Learning Informative Point Classes for the Acquisition of Object Model Maps," in *Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision (ICARCV), Hanoi, Vietnam, December 17-20*, 2008.