University of New Mexico

Adaptive Mesh Based Surface Reconstruction For Noisy and Incremental Point Cloud Data Sets

A thesis proposal submitted in partial fulllment for the degree of Masters of Science

Lucas Chavez
lucasc@unm.edu

Department of Mechanical Enginering
University of New Mexico

Advior
Dr. Ron Lumia

December 2012

Contents

1 Goal

The goal of this work is to design a method which can create a reliable representation of the environment from sequential registered noisy point cloud data sets. The method must be computationally feasible for online applications and have a low memory requirement. In addition, the method must adapt the representation to new measurements of revisited parts of the environment.

2 Problem

A rich representation of an environment is essential for most autonomous systems because it allows the agent or operator to have an increased situational awareness of the world. The methodology to build this representation is a continuously evolving subject in the field of robotics. The origins of the research into this problem date back roughly 25 years. Since then the methods and the representations themselves have continued to evolve at an impressive rate. The main catalyst behind this growth is the advancement of sensing technologies over the same time period. In general, sensors have continued to generate more measurements at higher rates over the years. This has provided an amazing opportunity to build richer and more useful representations of the environment.

In robotics the map building in an unknown environment is referred to as the Simultaneous Localization and Mapping (SLAM) problem. This label describes the fact that a methodology which solves the SLAM problem must simultaneously locate the robot in the environment as well as mapping the environment. The methodology by which the representation is built is called mapping and is the focus of this work. Early mapping methods represented the environment as a set of landmark locations. The result was a sparse set of points usually on a 2D plane. This allowed research to show that their SLAM solutions worked but it soon became clear that a richer representation of the world was needed. In response several methods were developed using various other representations. The resulting representations can be seen on Table 1.

	Adaptability	Computationally Inexpensive	Low Memory Requirement	SA: Robot	SA: Human
Landmark Locations	x	x	x	-	-
Point Clouds	-	x	-	-	-
Surfels	x	-	-	-	x
Implicit Functions	x	-	-	x	x
Mesh	x	-	x	x	x

Table 1: Characteristics of current forms of representation

Table 1 compares the characteristics of the representations which are currently used by mapping methods. Adaptability describes the ability of the representation to correct itself given new information. Computational expense describes how difficult it is to create and maintain a representation. Memory requirement describes how large much memory a method must use to run. Situation Awareness (SA) describes how well suited a representation is for both robot and human decision making. Robot decision making requires a representation that can be used for such problems as obstacle avoidance. Human decision making requires a method that can be allow an operator to intuitively understand the state of the robot given the map. The Table is supposed to reflect what a representation is capable of and not necessarily where the state-of-the-art is.

A mesh based representation is arguably an extremely good choice in comparison to the other representations. It has been used as sole representation by the gaming community because it is the best for representing large environments with the minimum needed memory. Also, this sort of representations works well to increase the SA of a robot because methods for performing physical simulations such as obstacle collision already exist. In addition, a mesh based environment is a very natural and comfortable method to display information to a human operator.

Currently, the problem with mesh-based environmental mapping techniques is that they are greedy. Once the mesh is in place there is no mechanism to adapt to newer measurements. The problem of adapting a mesh to new information is a very well studied problem in computer graphics, but these methods were not designed with large scale environmental mapping in mind. The biggest questions are: How can we quickly decide which measurements should be used to adapt what part of the mesh? How can we quickly detect new and removed objects? How can we robustly deal with noise and obtain a methodology that makes use of the new measurements of an already existing part of the representation? So the real question is can we develop a methodology that can solve all of the above questions and still have a manageable memory requirement and be computationally feasible?

3 State of the Art

To understand the cutting edge in algorithms which generate a representation from input sensor data we must consider two large fields of research. The first we will consider are the surface reconstruction algorithms from the field of computer graphics. Next we will investigate the question of SLAM from the field of robotics to see how the methodologies and the representations have evolved. Finally, we will show how ideas from surface reconstruction are starting to be used in some of the best available SLAM solutions. Finally, we will discuss the limitations of the currently available SLAM solutions.

Surface Reconstruction gained a lot of interest as the capabilites of computer graphics grew. It became a need to represent real-world items digitally in the computer. Some of the most important applications of this technology comes from medical imaging. For example, the typical way to view an organ was to look at several 2D scans of the organ. An algorithm was needed which could take several scans a give a surface reconstruction of the organ. This digital reconstruction could then be rendered and manipulated on ever evolving graphics cababilites of computers. Another desirable application was just to turn any given object in a digital representation. Instead of medical imaging techniques, this usually meant using a laser scanner or several camera images of the desired object. Surface reconstruction techniques can be classified into four classes following taxonomy of Seitz et al. [4].

4 Contribution

This work will add to the state of the art by creating a new methodology to leverage the depth image of the RGBD sensor to quickly identify the measurements of the depth image which are new and which are supporting a preexisting surface. The method to perform this categorization has never been performed on the depth image. The advantage of performing it on the depth image is that computationally efficient machine imaging techniques can be used. This method also quickly assigns the new measurements with the corresponding part of the mesh that needs to be adapted. In addition the adaptation procedure leverages QEM methods to perform the adaptation. Previously, this method had not been applied to adapting a preexisting mesh to new data. This methodology will avoid the trouble of over smoothing the mesh which are inherent in other methods such as virtual springs. The QEM adaptation is a computationally efficient method to create and adapt a piecewise smooth representation while preserving sharp details.

5 Impact

There are two main autonomous applications which will benefit from a richer representation of the environment.

A methodology for generating a rich representation mesh based representation of the environment would be useful for all autonomous systems which require situational awareness. There are many classical robotic applications which seek to reason about the environment in which the robot is working in. One major example would be object manipulation with a robotic arm. There are many already developed methods which use physical simulation methods to perform grasp selection and trajectory planning. In addition, mobile robotics can use a mesh map to perform path planning and obstacle avoidance. In general, a mesh a useful representation to plan actions and simulate those actions because such simulation engines already exist due to the gaming community. Planning those actions correctly requires a correct mesh. The proposed adaptive mesh procedure can give the best representation using the prior measurements. It can also work in a dynamic environment which is needed for real world applications.

In addition, an adaptive mesh based mapping procedure is needed for Urban Search and Rescue (USAR). The work of Bruemmer et al. [1] and other similar works [2, 3] has shown that the most effective user interface is one which display information in a similar way to many popular First-Person Shooter (FPS) video games such as Quake and Half-Life. The most natural way to build these types of interfaces is by rendering a mesh. Mesh rendering techniques, hardware, and software are already optimized for using meshes. Other types of representations can be displayed as a mesh, such as implicit functions, but they require a post processing step. The proposed methodology will build and maintain the map directly. The main impact of this work on USAR applications is that a much better mesh will be shown to the user than a mesh created through greedy methods which is the current state of the art. The mesh presented will adapt to true environment over time and with enough measurements. In addition, it will be able to operate in dynamic environments which have new or removed objects. In this way, the operator will always have a better awareness of the real environment.

6 Approach

Each time step of the algorithm will consit of two moajor processes. The results of these processes will then be used in a final map update procedure. After the map has been updated the next time step will begin. We can see in Figure blah a overview of the algorithmic structure. We will cover the details of

We will approach the problem of mantaining a mesh representation of the environment by designing two seperate proccesses. The outputs of which will then be utilized in an update procedure. Referring

The overall algorithmic design of the proposed system can be seen system flow diagram of Figure 1. The general shape of the algorthim consists of two independent proccesses which provide information which is used to update the current mesh. One of the proccesses provides an initial triangulation using the current depth image. The second proccesses categorizes the current depth image measurements using the most current estimation to generate an expected depth image and compare that to the actual acquired depth image. The categorization will determine how each region of measuremnts in the depth image will be used in the update procedure. If new elements must be added the triangulation found in the process "Find Connectiviity" will be used to define new elements in the mesh model. We will describe in detail each step of the two independent proccesses and then describe how the output of the proccesses are used to update the mesh.

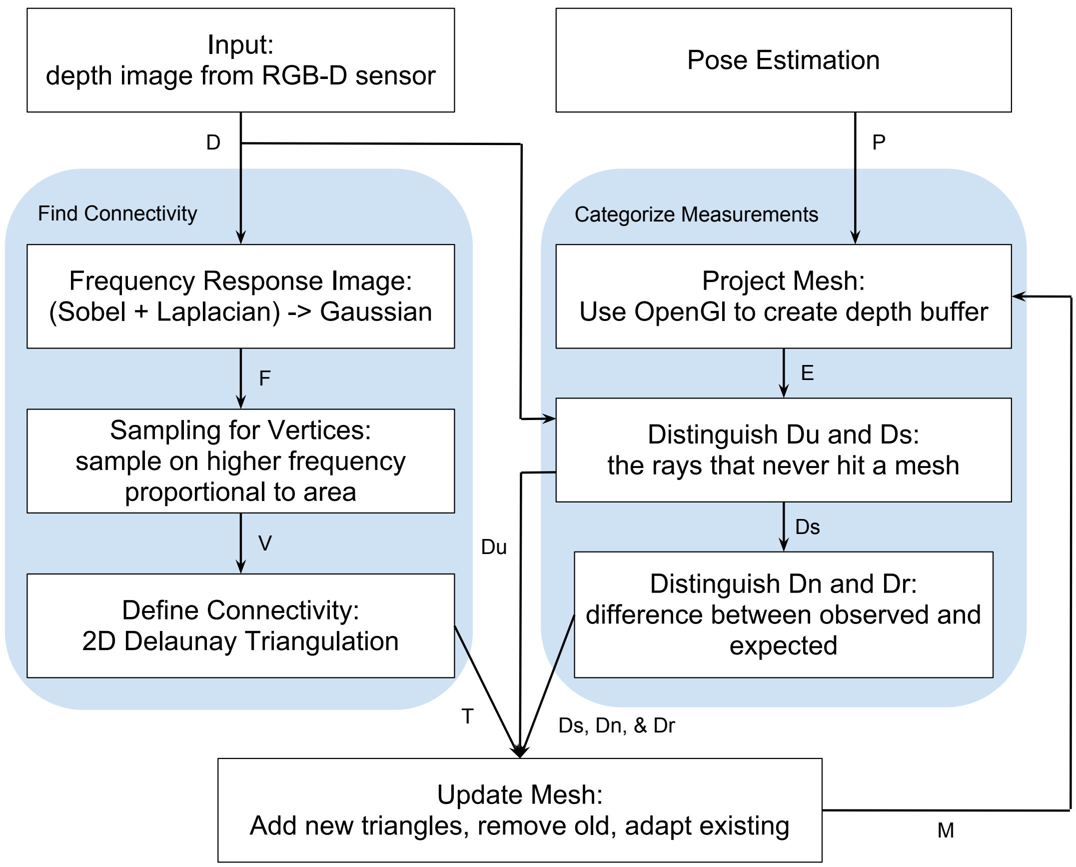

Figure 1: System Flow Diagram

6.1 Current Triangulation

6.1.1 Frequency Response Image

The objective of this step is to use image convolution techniques to design a sequence of image operations which will create an image that responds to portions of the depth image with high frequency information. The equation for the convolution can be seen in Equation **??**. These operators are not generally used on depth images, and in fact the best way to get an estimate of the curvature is to estimate the normal at each point in the pont cloud and calculate the change in the normal using neighboring point. The drawback of that method is that estimating the normals at each point is has a large computational cost. By utilizing several imaging operators we can create an estimate of the regions of the depth image which have high frequency information. Referring back to Equation

?? we can see that the depth image is convolved with a small Sobel operator K_S which is a discrete apporximation of the first derivative, this convolution will produce and image with high response at boundaries, such as corners. The depth image is also convoled with a Laplace operator K_L with a discrete approximation of the second derivative of the depth image, this creates an image with a high response on parts of the image with high curvature, such as a sphere. The absolute value is taken of the result from both the K_S and K_L operators. This is due to the fact that the derivative and second derivative estimates can produce negative values, however we are only interested in the magnitude of the response. The result of the convolutions of using both operators is then added, the idea is to produce an image which has a high response at boundaries and corners. A large Gaussian operator is then passed over the summation in order to produce an image which is a more conservative guess. This is inline with our goals because the idea is to quickly provide a rough guess of the frequency content. These imaging operations lend themselves well to optimization tecniques and are readily parallizeable in a GPU.

$$F = (\mid D * K_S \mid + \mid D * K_L \mid) * K_G \tag{1}$$

K_S – small Sobel operator

K_L – small Laplace operator

K_G – large Gaussian operator

6.1.2 Histogram and Sample for Vertices

6.2 Categorize Regions of Depth Image

7 Validation

8 Tasks

9 Gantt Chart

Table 2 shows my plan for completion of the research.

Thus, I plan to defend my thesis in XXX XXXX.

10 Committee

Timeline	Work	Progress
	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	completed
Nov. xxxx	XXXXXXXXXXXXXXXXXXXXXXXXXXX	ongoing
Jan. xxxx	Thesis writting	
Feb. xxxx	Thesis defense	

Table 2: Plan for completion of my research

References

[1] D.J. Bruemmer, D.A. Few, R.L. Boring, J.L. Marble, M.C. Walton, and C.W. Nielsen. Shared Understanding for Collaborative Control. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 35(4):494–504, July 2005.

[2] J.L. Drury, J. Scholtz, and H.A. Yanco. Awareness in human-robot interactions. In *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference*, volume 1, pages 912–918. IEEE, 2003.

[3] M. Waleed Kadous, Raymond Ka-Man Sheh, and Claude Sammut. Effective user interface design for rescue robotics. In *Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction - HRI '06*, page 250, New York, New York, USA, March 2006. ACM Press.

[4] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06)*, 1:519–528, 2006.