Oct- Tree Approach for Scattered Data Visualization

Addanki Adithya
Department of Computer Science, The University of Akron,
Akron, Ohio-44304. aa207@zips.uakron.edu

Abstract

The project aims to observe the effect of using an Oct-tree based approach for converting scattered data into scientific data for the purpose of Visualization. When we try to interpolate the scattered data into scientific data; in our case using Shepard method, we consider all the given scattered data points and calculate the value at all the grid nodes (given the dimensions N_x, N_y and N_z). But, the issue with this is when the points are scattered all over in the 3D Space, or even worse if the points in the scattered data are biased to a single side, we might want to calculate the value at a particular grid node based only from the scattered data points that fall into a particular bound range of (x, y, z) co-ordinates. We discuss the results of our analysis using Pollution dataset.

1. Introduction

Visualizing Scattered Data is a critical challenge as it tries to correctly model sample scattered data to accurately render the data variation throughout the volume of interest [3]. Scattered data samples are encountered in almost all the disciplines of science and engineering. Thus studying these techniques might help us to gather more interest in this area.

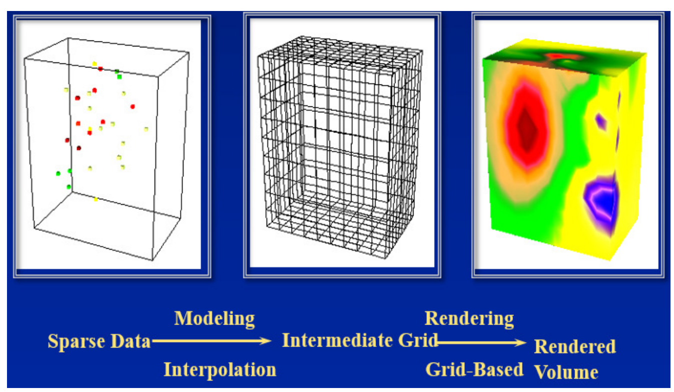

Figure 1.1: Scattered Data Visualization [7]

[Figure 1.1] The Primary step is to convert the sparse data through interpolation into scientific data which is then rendered using grid. The analysis we perform using Oct-Trees is a preprocessing step for studying the effect of scattered sample points that lie far from the point on the grid node.

2. Background

2.1 Visualization

Visual Communication is the fundamental process for exploring and dissemination of information through visual aids such as diagrams, sketches, photographs, charts, video, animations etc., thus improving the quality of comprehension, memory and inference. Being able to store exabytes of information is of little value by itself, understanding the data to draw useful inferences from and make decisions is vital. The strength of visualizations lies in the design principles which connect the visualizations with viewer's perception and cognition of underlying information that is conveyed. The key to good visualizations use the best practices to either emphasize the relevant information or de-emphasize irrelevant information [5]. Data Visualization is a critical and important tool to understand and make inferences about the observed phenomena [4].

2.2 Oct Tree

Oct Tree is a data structure similar to Quad-trees, except for that Quad-trees are used in 2D and Oct-Trees are used in 3D. The whole space is divided into eight octants (as the space we consider is 3D) and then the process is continued until the lowest possible or required granularity is obtained. [Figure 2.2.1] depicts how a sample space is split into octants. Then, points which fall into the octant range of the grid nodes are considered for calculating the Shepard value and rest are ignored; the process is continued until all the values for structured grid are obtained.

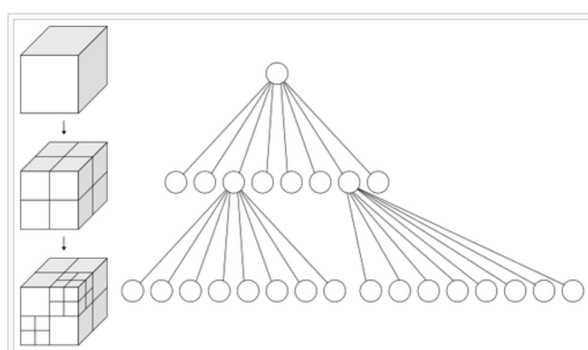

Figure 2.2.1: Oct Tree Split Demonstration [1]

3. Design

3.1 UML Diagrams

We build models to

- To communicate the desired structure and behavior of our system.
- To visualize and control the systems architecture.
- For better understanding of the system we are building; often exposing opportunities for simplification and reuse.
- To manage risk.

3.1.1 Use case diagram

A Usecase diagram is used to illustrate the static design view of the system. Given below is the use case model for our system. Representations and relationships used inside of the system could be checked through references [6].

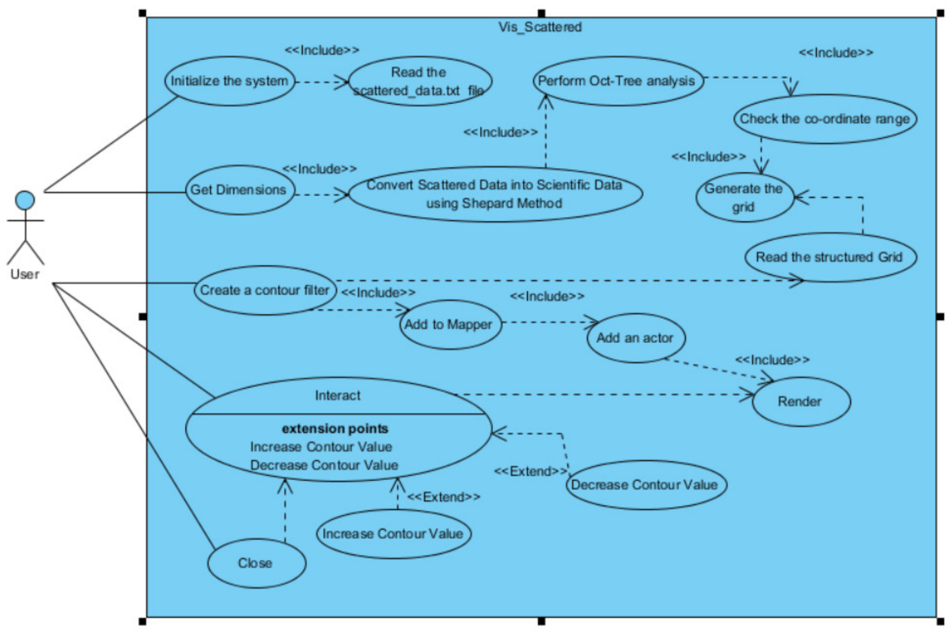

Figure 3.1.1.1: Use Case Diagram

The use case diagram illustrates the requirements on a given system. It gives an understanding of how the model is implemented and the actions the user can perform over the system through his interaction. In other words it represents the behavioral view of the system.

3.1.2 Class Diagram

A class diagram depicts a static view of the system, the way different classes interact with each other, their relationships and dependencies between the classes.

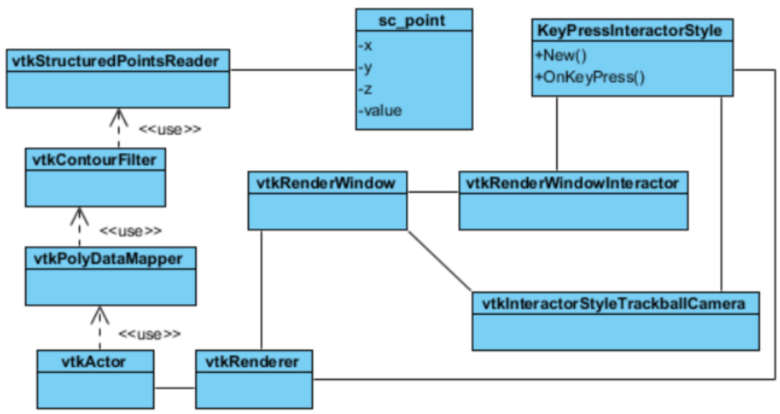

Figure 3.1.2.1: Class Diagram

As shown in the [Figure 3.1.2.1] we have the standard VTK classes as well as custom classes for event listeners that allow the user to interact with the visualized model and render again on itself, when required.

3.2 Implementation

3.2.1 Dataset

Pollution concentration dataset, has 138 scattered points with the X, Y, Z and V values falling in the range of [0, 1]. We interpolate this scattered data into scientific data using Shepard Method and create a visualization that illustrates the pollution concentrations in 3D.

3.2.2 Algorithm Implementation

Interpolation from scattered data to scientific data could be done in three ways:

- Local: based on the points lying near by the point interpolation.
- Global: based on all the points in the dataset.
- Exact: the interpolation function could exactly reproduce the points in the dataset.

Oct-Tree Analysis: Based on the required granularity by the user, the implementation divides the 3D space into octants and continues the split until the octant range of the points on the grid is known. Once the range is identified the algorithm checks for the sample points that lie in this interval range and considers them for interpolation (Local Interpolation) using Shepard Method.

Shepard Method: We used Shepard method as it is one of the oldest method used for converting scattered data into scientific data for the purpose of visualization. An algorithmic overview of Shepard method is given below. It uses an inverse distance based weighted average of the sample values, resulting in normalized values.

$$f(x,y,z)=\left[\sum_{i=1}^{n} v_i d_i^{-\alpha}\right] / \left[\sum_{i=1}^{n} d_i^{-\alpha}\right].$$

Figure 3.2.1.1: Shepard Method

Here n is the number of sample data points, x,y,z are the co-ordinate points on the grid, d is the distance from the grid node to the sample point, α is the user defined constant(our case 2, i.e., distance inverse squared).

These values are written into a file which is then read by the vtkStructuredPointsReader and visualized to the user, the user is allowed to interact with the displayed visual object.

3.2.3 Filters Used

Contour Filtering is used to render the ISO surfaces for different contour values, user is allowed to increase or decrease the value with key press; *"i"* to increase the contour value, and *"d"* to decrease the contour value.

4. Results

4.1 Expected Result

A better and more accurately calculated scientific data, (represented in 3D using VTK). Programming is done in C++. As the user presses *"i"* the ISO surface becomes smaller and smaller as fewer regions have high pollution concentration and vice versa.

But unlike before where the concentrations are to be observed at a particular contour value, the Oct-Tree approach should help us even out the effect of scattered points that are far away from the point under consideration. The actual results and screenshots of our implementation are given in the next section.

4.2 Observed Results

4.2.1 Prior to using Oct-Tree Analysis

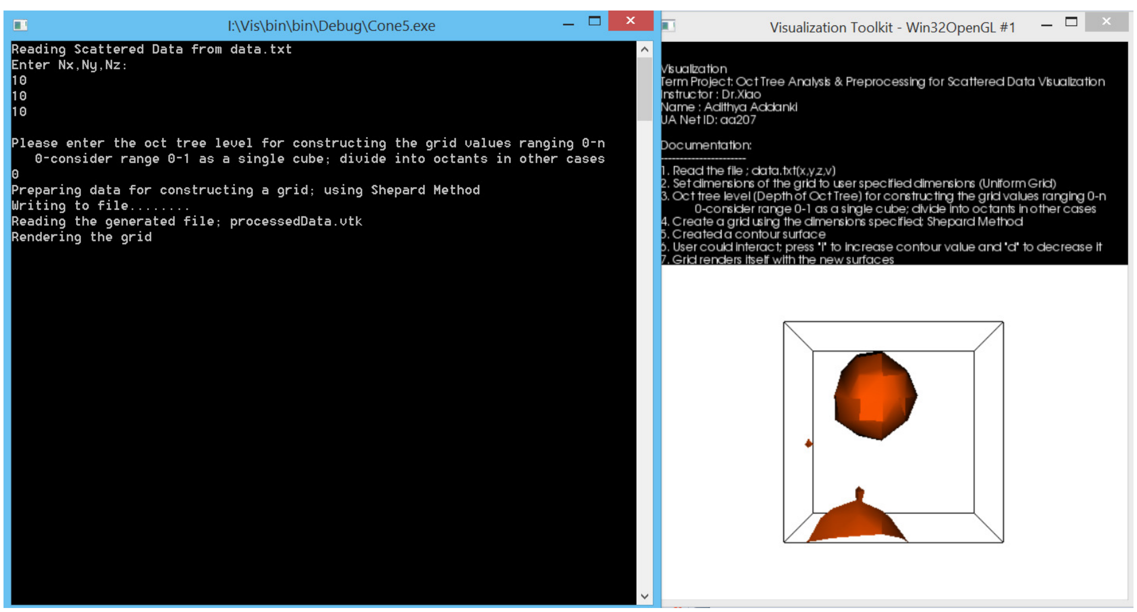

Figure 4.2.1.1: Output generated by vtk after application of Shepard method.

4.2.2 After Application of Oct-Tree based Approach

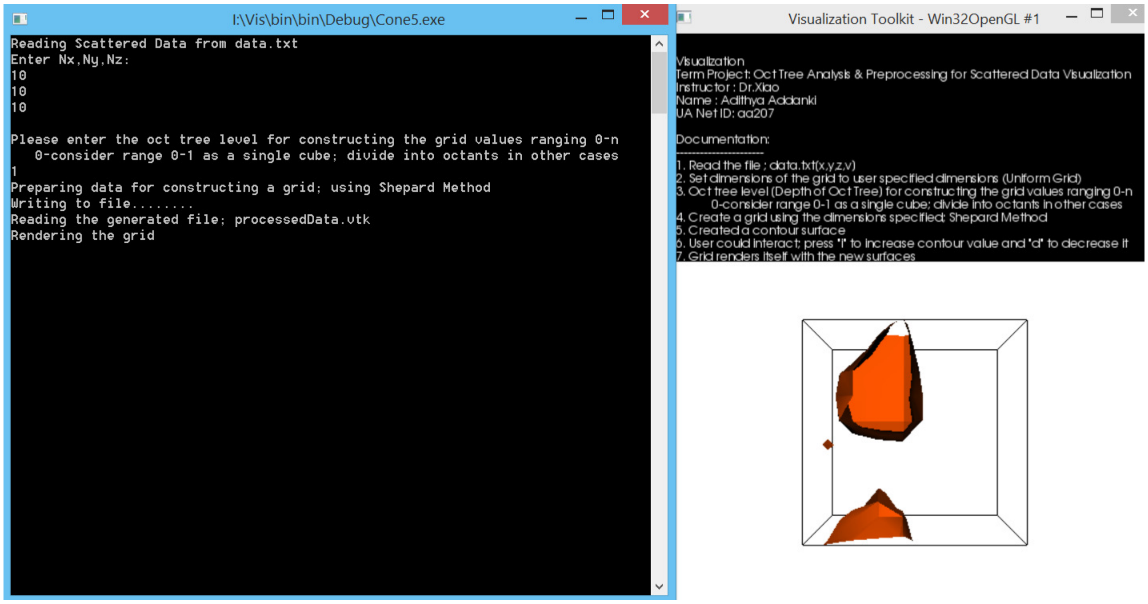

Figure 4.2.2.1: Oct-Tree Level-1; 0-1 co-ordinate range is divided into 8 octants.

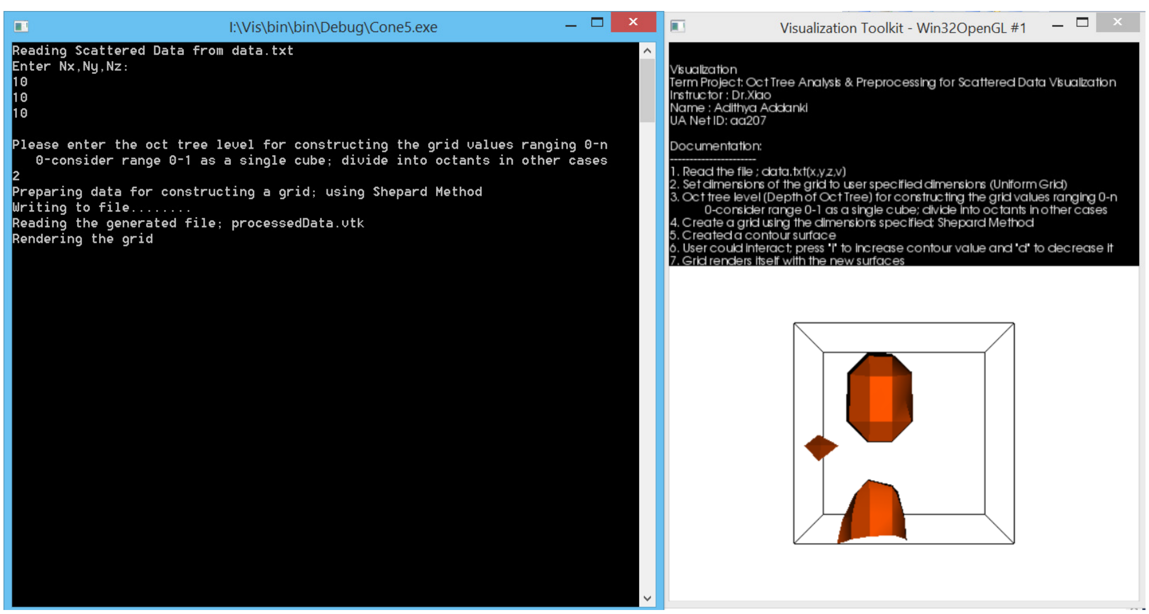

Figure 4.2.2.2: Oct-Tree Level-2; 0-1 co-ordinate range is divided into 8 octants, repeated once.

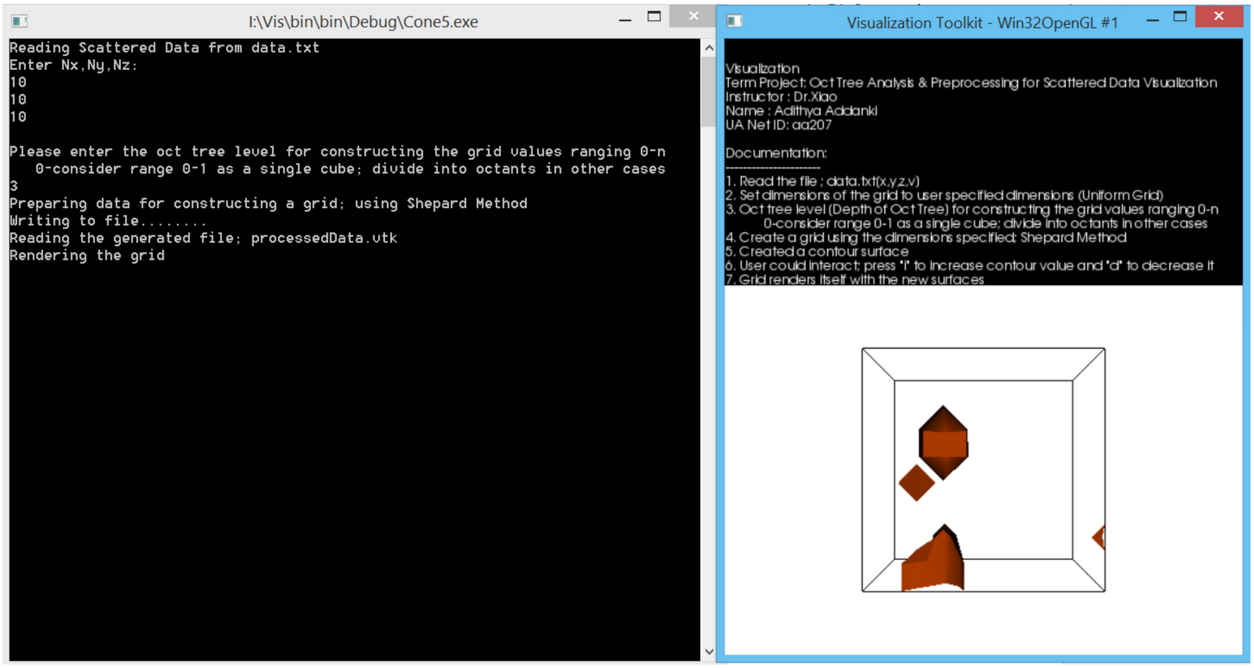

Figure 4.2.2.3: Oct-Tree Level-2; divided into 8 octants, repeated twice.

5. Observations

We observed that using Oct-Trees for preprocessing the data, and determining the local points for a point on grid node helped us understand the effects of scattered sample points that lie far away. Thus Oct-Tree based approach helped us to even out the unrealistic concentrations found in the visualization displayed. The unwanted effect of sample points that are not in the octant range are not considered thus reducing the error of the interpolation function.

6. Conclusion

Oct-Trees have found its use in a wide range of areas ranging from memory space requirements reduction to triangulation. Our project helped us understand that Oct-Trees based approach could be used for understanding and studying the effects of localizing the interpolation in the conversion of scattered data to scientific data. In the future this might be considered as preprocessing step to be applied for various scattered data interpolation methods.

7. References

[1] . http://en.wikipedia.org/wiki/Octree

[2] . BSP Trees, Quadtrees & Octrees; University of Alaska Lecture Notes.

[3] . Xiao, Y., Ziebarth, J. P. (2000). FEM-based scattered data modeling and visualization. Computers & Graphics, 24(5), 775-789

[4] . Mustafa, G., & Abad, A. S. (2007). Modeling and visualization of scattered positive data. Information Technology Journal, (1), 103

[5] . Agrawala, M., Wilmot, L. I., & Berthouzoz, F. (2011). Design principles for visual communication. Communications of the ACM, 54(4), 60-69. doi:10.1145/1924421.1924439

[6] . Grady Booch, James Rumbaugh, Ivar Jacobson. The Unified Modeling Language User Guide (2005).ISBN-10: 0-321-26797-4. ISBN-13: 978-0-321-26797-9.

[7] . Dr. Xiao, Scattered Data Visualization; University of Akron, Lecture Notes.