QUAD TREE DATA STRUCTURE FOR LIDAR DATA -DETAIL AND USES

Prateek Kumar Behera

Emaail : prateekbeh20@iitk.ac.in
Roll no. : 20103082

ABSTRACT:

As we are aware of the data mountains being generated due to variours reasons such as the we take the observation at a frequent interval and also we take redundant data which leads to very large amount of data for which a very large amount of storage space is required. Therefore we use some data structures such as quad-tree to minimize the storage requirement for the large amount of data.

1. INTRODUCTION

With the emergence of the aerial survey and photogrammetry there has been an exponential growth in the data size of the geographical data which can be acquired using different sensors of type such as optical, laser, hyperspectral . LiDAR data is being used to acquire Digital Elevation, Digital Terrain models. We can use the quad tree data structure to reduce the complexity of the irregular sample point clouds. For example, In building ground plans we need to get the outline of the area of the building and we can eliminate too small edges of the building . The Quad tree data structure can also be used to reduce the storage requirements for the data mountains which are being generated by the aerial LiDAR data survey. As the size is being reduced of the LiDAR data the precision of the data is reduced as we are trying to represent a variable number of points using a single class/value.

Some of the parameters for the Quad Tree Data structure are the levels of the quad tree.The root of the quad tree has the level as zero and the level is increased by one level for each subsequent child of the tree. Another properties of the quad tree data structure is that the nodes at the same level have the same level of coarseness. We also discuss about other uses such as how we can use the Quad Tree Data Structure for the determination of large scale digital elevation model in any area.

2. QUAD TREE STRUCTURE

Similar to binary tree data structure which has two atmost two children in case of the tree, in case of quad tree data structure we have 4 children for a node.This type of structure basically represents a 2-D plane where we can bifurcate or divide the data in 4 different quadrants and recursively keep dividing the node until a termination condition is reached or in other terms we can visualize it that every child of the node represents a quadrant of the plane.Similar to quadtrees we can also use oct-trees for implementation of 3-D plane where it has 8 children for any particular node.

In case of geo-spatial data we can basically divide the four children into 4 zones of NW(North West),NE(North East),SW(South West),SE(South East) as four different quadrants. and define some terminating conditions for which we get tree at different levels.

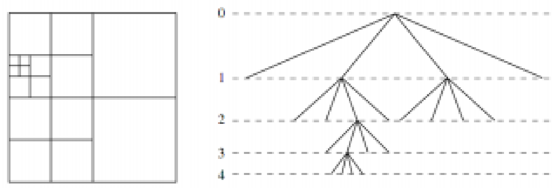

Figure 1. On left side we have representation of how quad tree structure is being implemented for a image and on the right we can see the quad tree structure with the levels and each node having either zero children or 4 children .

2.1 Construction of Quad Tree Data Structure

The algorithm which we can use for the construction of the quad tree data structure are as follows:

1. In case of LiDAR data we iterate through all the points in the data file.
2. If any four data points satisfies some desired properties we merge it into a single node.
3. We keep on merging the cells until we reach the root node.
4. This process is also known as bottom-up neighbour finding.

2.2 Space Optimization and Time Complexity

We will discuss about the space optimization using the example below.

In the figure 2 the A or the root node depicts the entire data.We keep on dividing the data until the constraint equation is being satisfied. As in the above diagram we could see a region quadtree with values as 0 and 1 and we define a constraint can be that we need a blocks of data which contain 0 or 1. Then in this case we divide the A or the root node into 4 children or quadrant. After dividing the tree we get B and D quadrants as shown as the quadrants which satisfy the condition that all the values in the quadrant should be either 0 or 1 but for remaining children i.e C and E the condition is not satisfied and we divide the tree further into its children. After dividing the node C and E we get the nodes F,G,H,I,J,K,L,M as shown in the figure

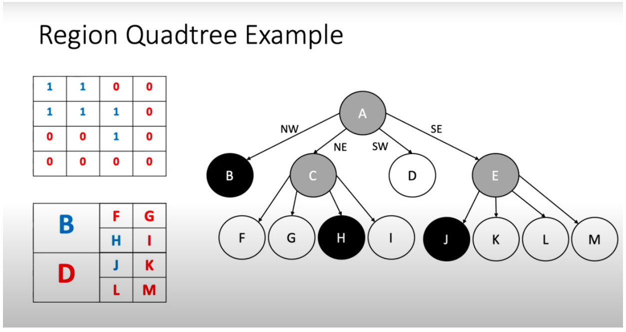

Figure 2. We use the above rasterized data to compute the space complexity and the time complexity for the quad tree data structure .

which satisfy the condition that either blocks of all 0's or 1's is present for the quadrant.

To compute for the space complexity we can try taking two cases: In 1st case we try to take values for each data point i.e for each cell we get 16 data points for the above example but if we use the quad tree data structure and try to store it we see that we only have 13 nodes which we need to store but the storage optimization leads to reduced level of accuracy because we are not able to visualize the data in a more granular way.

To compute the time complexity in case of LiDAR data it will be O(n) where n is the no. of data points. In case of rasterized data the complexity is $O(n^2)$ where n is the no. of lines because we have to traverse the quad tree row by row.

2.3 Uses of quad tree in LiDAR data

There are many uses of the quad tree. Some of which are mentioned below as:

2.3.1 LIDAR Data for Building Extraction Based on Quad-tree Structure The LiDAR data consists of (X,Y,Z) where (X,Y) gives the coordinates and Z gives the elevation and we also have the intensity value.A quad tree data structure has a threshold value which is used to reduce the size of the tree but the accuracy is being reduced as the generated output using quad tree is no longer a replica of the original data. The threshold is given to simplify the quad tree. The threshold value is a condition which is used to merge the two adjacent nodes into a single node it the difference between them is less than the threshold.

Method for extracting the building from LiDAR data using quad tree

1. To visualize the output of the data structure we take only a sample are from the original image as shown in Figure 3(a)

2. We generate a DEM by eliminating the above ground points which consists of buildings and other artificial structures so that we can combine the slope of the ground with the adjusted elevation value. In the experiment conducted the threshold value for height is taken as 45 and the slope is taken as 3.0.

3. Then we classify the points to different classes by comparing the height and the distance value for every two point . The threshold selected in this experiment is 5.0m.

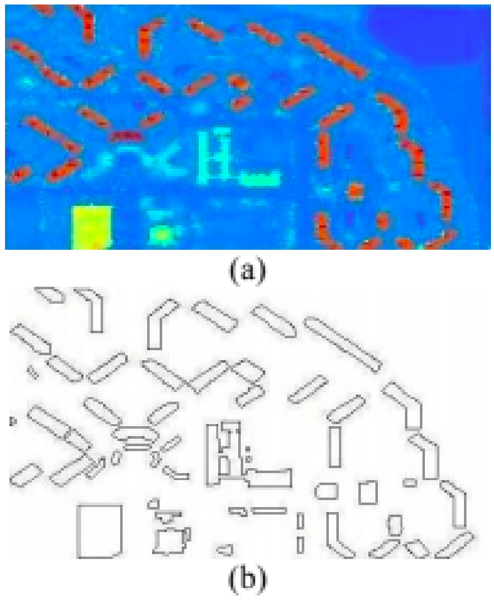

Figure 3. The figure a gives us the sample or the original image and figure(b) shows the image we get after applying the quad tree data structure

4. For every group which is generated from the above step we assign a shape by using a chain algorithm suggested by Andrew's Monotone Making and then coloring the outline of the building.The image of generated image using the quad tree is being displayed in figure 3(b).

2.3.2 LIDAR Data for determination of large scale digital elevation model in any area. Acquiring the DEM is essential for many land survey application such as forest management. The quadtree-based iterative algorithm is proposed by Kraus's iterative algorithm Following is the procedure we use to determine the large scale DEM :

1. First we bifurcate the LiDAR data into a grid network on 2-dimensional plane (x,y). Each quadrant can be thought of as the sides depicting NW(North West) , NE(North East) , SW(South West) , SE(South East).

2. Then, Kraus's filtering algorithm is being performed within each initial square unit to remove above ground points for the first time. The statistics of the leftover points will provide the following parameters for the next iteration:
   (a) No. of leftover LIDAR points.
   (b) Residuals of leftover LIDAR points to derived average surface along with altitude direction.

3. According to the above indices, initial unit regions can be divided into four quadrants or subregions.Each box has a dimension of n/2.

4. Condition 1: Number of leftover LIDAR points is greater then the specified threshold.

   Condition 2: Average of residuals is greater than the specified threshold.

5. In next step, we will perform quadtree segmentation within unit regions, which will satisfy the above constraint

6. The above steps are recursive in nature and repeated for whole of the data. The final DEM is generated from TIN, which consists of ground points extracted from LiDAR data.

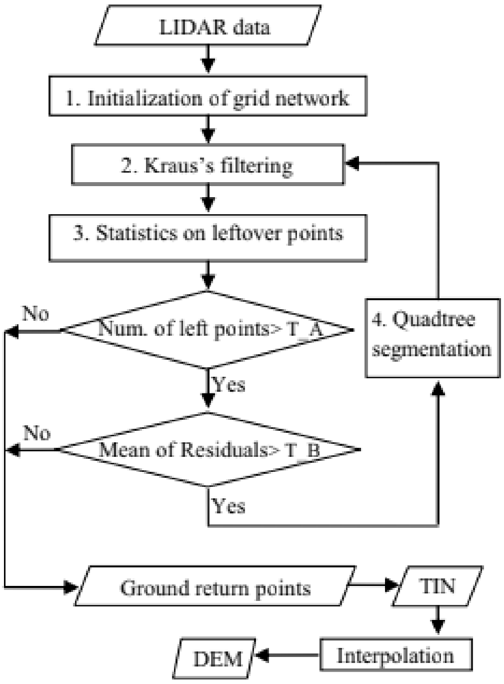

Figure 4. The above flow chart depicts the algorithm we use to determine the DEM for any area

Results

Following is the image representation after we do the segmentation using quad tree.

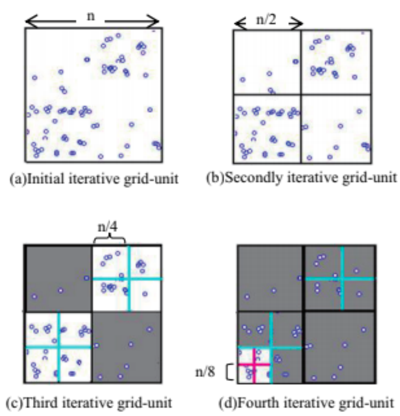

Figure 5. The above flow chart depicts the algorith we use to determine the DEM for any area

REFERENCES

Jianqiao Lia, Haisheng Fan b, Hongbing Ma a, Shintaro Goto b *Determination of Large-Scale Digital Elev. Model in Wooded area with Airborne LiDAR data by applying adaptive quadtree-based iterative filtering method.*.

Ruoyu Du1, Hyo Jong Lee1,2 *Simplification of LIDAR Data for Building Extraction Based on Quad-tree Structure Division of Electronics and Information Engineering, Chonbuk National University 2 Center for Advanced Image and Information Technology*.