Fuzzy SART Clustering for 3D Reconstruction from Irregular LIDAR Data

NICHOLAS SHORTER & TAKIS KASPARIS
Department of Electrical and Computer Engineering
University of Central Florida
4000 Central Florida Blvd.
Orlando, FL 32816-2450, USA
nshorter@mail.ucf.edu & kasparis@pegasus.cc.ucf.edu
http://www.nshorter.com & http://people.cecs.ucf.edu/kasparis/

Abstract: – Fuzzy Simplified Adaptive Resonance Theory (Fuzzy SART) is proposed for clustering the normal vectors of coplanar triangles in a Triangulated Irregular Network (TIN) derived from raw, irregularly spaced LIDAR data. The raw LIDAR data is triangulated using a greedy insertion triangulation algorithm. The algorithm is modified to implement a proposed noise filtering technique which improves Fuzzy SART's coplanar clustering performance. Furthermore, several proposed preprocessing techniques, specifically scaling and translating the LIDAR data, are also implemented and were found to improve clustering accuracy. Then, a proposed multiple or planar regression algorithm further refines the coplanar grouping by removing outliers and merging multiple clusters representing singular planes. Finally, further refinement is achieved by calculating the intersection of the best fit roof planes and moving nearby points close to that intersection to exist at the intersection, resulting in straight roof ridges.

Keywords: – 3D Reconstruction, LIDAR, FSART, Greedy Insertion, Triangulation, Fuzzy ART

1) Introduction

The advent of Light Detection and Ranging (LiDAR) systems has spawned a new genre of reconstruction algorithms which explore using captured LiDAR data as an additional feasible source of information for 3D reconstruction. Most 3D Reconstruction algorithms typically group coplanar points together and then derive a model to approximate those points which yields minimum error from the original points. Rottensteiner and Briese, after distinguishing buildings from terrain via the method described in [10], detect roof planes via a curvature based segmentation technique, and then group those planes into polyhedral building models. Chen et. al. in [2], segment buildings into coplanar sections, which are then processed by a patented Split-Merge-Shape (SMS) method to create building models. Fujii and Arikawa use both LIDAR data and aerial images in [4]. After interpolating the LIDAR data to fixed intervals, the data is analyzed for line segments forming object contours. In [8], Overby et. al make use of a three-dimensional extension of the Hough transform for extracting planes from point cloud data.

Several data driven 3D reconstruction algorithms currently exist in the literature ([9],[6],[8],[7]) that utilize triangulated irregular networks (TINs) to construct model approximations of depicted urban and residential scenes. A TIN is a 3-dimensional depiction of LIDAR point cloud data represented with a series of connected, non-overlapping triangles which have no intersecting edges. Morgan and Habib in [9] use a region growing TIN algorithm, based on least-squares adjustment, to extract building facades from the transformed point cloud data (transformed to the triangulated feature space). Chen et. al., in [8], also use a region growing TIN algorithm, considering both the height difference between triangles and the angle difference between normal vectors of neighboring triangles for merging criterion for planar approximation. Hoffman however uses the clustering approach in [6] to group together triangles in the TIN that contain similar

properties. In [6], the position of each triangle is mapped out in spherical coordinates which are the dimensions of the triangles that are then clustered.

The majority of the previously described algorithms, in addition to numerous other 3D reconstruction algorithms existent in the literature, ([2], [3], [4], [8],[9]) all share one particular trait in common: the algorithms interpolate the LIDAR data to fixed intervals to make use of conventional matrix processing methodologies. As discussed in section 2.2, interpolation of the irregular points has significant drawbacks.

2) Problem Formulation

Characteristics about the data set used are presented in section 2.1. The disadvantages of interpolating LIDAR data are elaborated on in section 2.2. Section 2.3 reports on the difficulties of clustering coplanar LIDAR points with TIN normal vectors.

2.1 Data Set

The Fairfield data set, used for this research effort, covers two square kilometers of both an urban and residential scene in Fairfield, Australia. A LIDAR sensor and other systems used to procure the Fairfield Data Set were able to capture, for each laser point, the longitude, latitude, elevation, first return pulse, last return pulse and returned laser intensity. The procured LIDAR data has an approximate point density of 1 point per $1.3m^2$ and the aerial photography captured has 15 centimeter pixel resolution.

2.1 LIDAR Interpolation Disadvantages

In [3], Clode et. al., report the limits of their building detection technique and how interpolating the LIDAR points only adds to the inaccuracy and limitations of their method and all other methods using interpolated LIDAR data. The accuracy in which a given algorithm can delineate building from non building regions is dependent on the laser footprint uncertainty. However, if the data is interpolated to fixed intervals, then the limitation of the accuracy is worsened. With interpolation to fixed intervals, now a given algorithm's uncertainty is a function of the laser foot print uncertainty and the point spacing interpolation combined.

2.2 Depicting LIDAR with TINS

Several of the aforementioned TIN 3D Reconstruction methods utilize a TIN generated from the LIDAR data and cluster the triangles existent in that TIN into coplanar regions. The normal vectors of the coplanar triangles are very similar to one another and therefore are the attributes clustered. However, the accuracy of even the most modern LIDAR sensor systems have limits of precision in their range to target measurements. Noise is therefore existent in the data and causes havoc for algorithms attempting to cluster triangle normal vectors. An ideal set of coplanar triangles is depicted in [Figure 1]. However, noise introduces slight deviations in the elevation dimension [Figure 2], causing the normal vectors to deviate from their true value, limiting the practicality of clustering an irregular spaced TIN based on normal vectors.

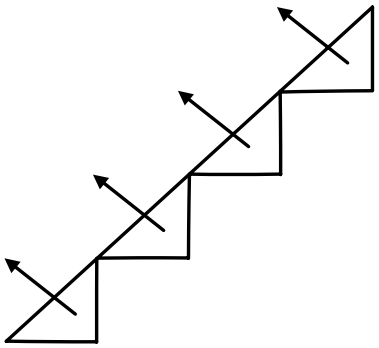

Figure 1 – Ideal normal TIN Vectors

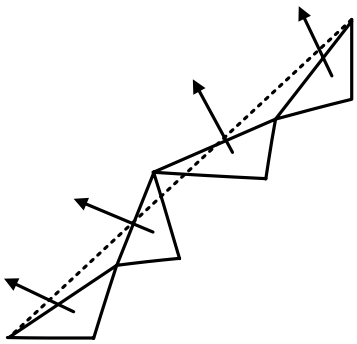

Figure 2 – Actual TIN Vectors

3) Problem Solution

Section 3.1 elaborates on the TIN generation method implemented. Section 3.2 describes what attributes of Fuzzy SART make it desirable for detecting coplanar LIDAR points. Heuristic procedures which improve on the Fuzzy SART clustering presented in section 3.3.

3.1 Modified Greedy Insertion Triangulation

Wang et. al. in [11] benchmark several triangulation processes on LIDAR data. Among all of the triangulation methods tested, they found the sequential greedy insertion algorithm [5] performed the best in terms of accuracy. The sequential greedy insertion algorithm simultaneously optimizes two adaptive cost functions: (1) local Delaunay triangulation; (2) global point insertion. The distances between all unused points and the existing generated TIN are calculated. The point farthest from the generated TIN is then inserted into the TIN (hence the name greedy insertion). After point insertion, all of the triangles surrounding the newly formed triangles are then checked to see if flipping their diagonals with adjacent triangles will further maximize the lesser internal angles of given pairs of triangles (Delaunay triangulation).

It is important to note that Greedy Insertion inserts the points farthest away from a given triangle. Therefore the most defining features in a given triangulated scene, such as break lines and building corner points, are the features triangulated first. The very last points triangulated are points with the smallest errors, points existing on well defined planes.

A novel method is therefore proposed to filter the triangulated data by absorbing points less then a given threshold to an already defined plane. While the longitude and latitude dimensions were preserved, the elevation dimension of a candidate point was modified if it met the following conditions: the perpendicular distance, defined in (1), was less than .2 meters ($D_c \le .2m$) from the containing triangulated plane; and the pitch of the roof , defined in (2) and [Figure 3], was less than 60 degrees ($\theta \le 60^o$).

$$D_c = \frac{\left|a\cdot(x_0 - x_1) + b\cdot(y_0 - y_1) + c\cdot(z_0 - z_1)\right|}{\sqrt{a^2 + b^2 + c^2}} \quad (1)$$

$$\theta = 90 - \sin^{-1}\left(\frac{D_c}{Z_c}\right) \quad (2)$$

Without the constraint imposed on the roof plane pitch ($\theta \le 60^o$), building edge points, which were not yet inserted/triangulated and less than .2 meters perpendicular distance from the building were being merged into the building's edge, distorting the building outline. The second constraint therefore confines points which only exist on a roof plane with a pitch ($\theta \le 60^o$) to become merged with that existent roof plane.

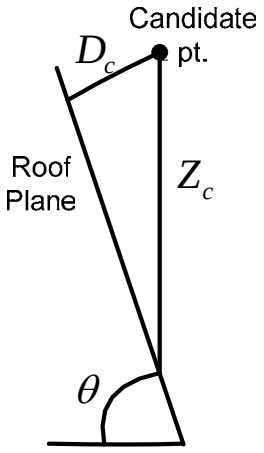

Figure 3 – Pitch of Roof Plane (Theta)

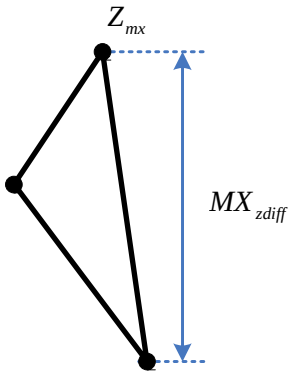

Figure 4 – Triangle Elevation Difference

This filtering technique was found to remove the noise depicted in [Figure 2] and therefore significantly improve normal vector triangulation clustering results. Merging these noisy points' elevation z-dimension reinforced the presence of existing roof plane clusters resulting in an improved clustering performance by Fuzzy SART.

Only the spherical coordinates of the normal vectors of roof triangles were passed to the Fuzzy SART clustering algorithm. Triangles belonging to building walls and terrain were disregarded. In order to distinguish roof triangles from all other triangles the following measures were implemented. For all triangles, the difference in elevation between the two vertices farthest from one another in a given triangle is calculated (MX_{zdiff} in [Figure 4]).

All triangles having MX_{zdiff} greater than 2 meters were isolated. Then the average, Z_{avg}, of the z-dimension (elevation) of the highest vertex in a given triangle, Z_{mx} in [Figure 4], for all triangles with $MX_{zdiff} \geq 2m$ was taken. All triangle centers must have an elevation greater than Z_{avg} in order to be considered as a candidate for a roof plane. The Z_{avg} - restriction implements the assumption that all the buildings are greater than 2 meters or 6 and ½ feet in height.

3.2 Fuzzy SART Clustering

Due to co-planarity of triangles being defined as triangles sharing the same normal vectors, which are expressed in spherical coordinates, those spherical coordinate vectors are what is passed to the unsupervised clustering algorithm. An optimal choice of a clustering algorithm, which will exploit the vectors being represented as spherical coordinates, is Fuzzy SART.

One of the key features of Baraldi and Parmiggiani's Fuzzy Simplified ART (SART) clustering algorithm [1] which makes it choice for this clustering problem is its activation function. The Vector Degree of Match (or activation) function (3) consists of the product of two functions: the Module Degree of Match (MDM) (4) and the Angle Degree of Match (ADM) (5).

$$VDM(T,X) = MDM(T,X) \cdot ADM(T,X) \quad (3)$$

$$MDM(T,X) = \min\{|T|/|X|, |X|/|T|\} \quad (4)$$

$$ADM(T,X) = (\pi - \alpha)/\pi \quad (5)$$

$$\alpha = \cos^{-1}\left[(X \circ T)/(|X| \cdot |T|)\right] \quad (6)$$

Both of these functions, (4) and (5), have values that range from 0 to 1 corresponding to their input component similarity. In other words, MDM approaches unity as the two vectors inputted to the function approach equal magnitude. As the inputs, the template vector T, or vector representing a given roof plane, and the input vector X, or vector representing a given triangle's normal vector, approach the same orientation and direction, the ADM approaches unity.

Fuzzy SART Contains two user defined parameters: Tau, the approximate time it takes Fuzzy SART to learn patterns; and the Vector Degree of Math Threshold (VDMT). The VDMT controls how conservative Fuzzy SART is when absorbing inputs to preexisting patterns. The VDMT ranges from 0 (where all inputs belong to the same pattern) to 1 (where all inputs are assigned to different patterns). The VDMT was set to 0.7 to avoid Fuzzy SART from falsely accepting an input to a plane in which that input did not belong to (false positive).

Assuming the Vector Degree of Match Threshold (VDMT), also called the vigilance parameter, is held constant, the VDM function, applied to a vector pair T and X, defines a hyper-volume in bi-dimensional feature space [Figure 5].

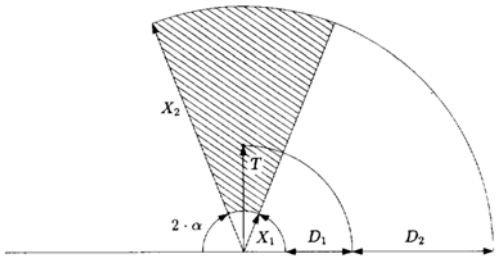

Figure 5: Fuzzy SART Pattern Encoding Region

The template T represents a cluster center or in this case an established roof plane vector. The angle α is derived from equations (5) and (6). The hyper arc volume encoding regions in which Fuzzy SART encodes its patterns with make it an optimal choice for clustering spherically represented normal vectors. Coplanar normal vectors with slight deviations will accurately be captured by the pattern encoding regions.

Note that the cluster encoding regions formed are all circular about the origin. All of the original building coordinates were originally all positive values. It was found that the performance of Fuzzy SART for clustering the normal coplanar vectors improved if the building coordinates were shifted such that the building was centered at the origin. Thus each dimension (longitude, latitude and elevation) was modified as follows:

$$X^d(i) = X^d(i) - \left(X^d_{\min} + \frac{X^d_{\max} - X^d_{\min}}{2} \right) \quad (7)$$

Where i represents the i-th point in the data set, d represents the dimension, and $X^d_{\max}$ and $X^d_{\min}$ represents the maximum and minimum points in those dimensions, respectively.

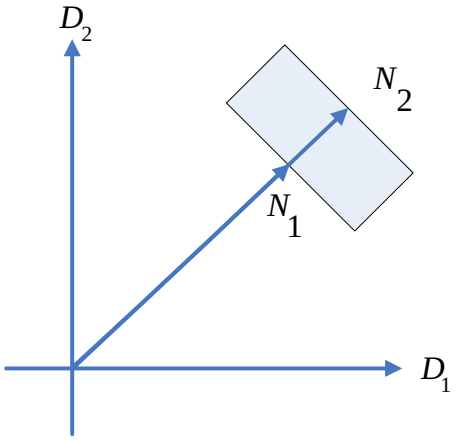

Figure 6 – Not Shifted

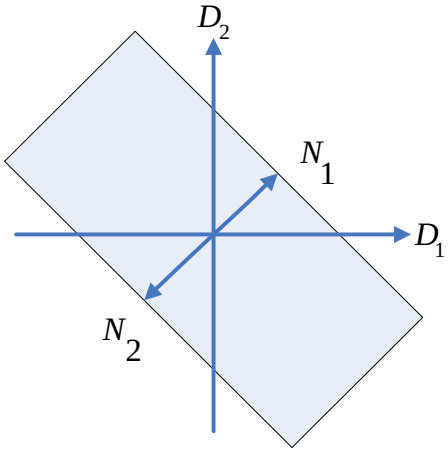

Figure 7 - Shifted

Consider the two building footprints depicted in [Figure 6] and [Figure 7]. Note that the difference between two of the normal vectors, N_1 and N_2, of two of the planes depicted in [Figure 6] is a difference in their magnitude, while their orientations are equal to one another. This difference converges to 0 as the building moves infinitely far from the origin. However, if the building is translated to center at the origin, then the two aforementioned normal vectors, N_1 and N_2, now have the same magnitude but are 180 degrees out of phase from one another. This further separates the planes from one another from the coplanar clustering perspective of Fuzzy SART, thus increasing Fuzzy SART's clustering performance accuracy.

After the building coordinates were converted to spherical coordinates, it was found, experimentally, that Fuzzy SART assigned more weight to dimensions with higher values compared to other dimensions when clustering. Therefore, in order for each dimension to receive an equal weight or significance when clustering, all the dimensions were scaled such that they existed within the same range.

3.3 Heuristic Procedures

Fuzzy SART only provides a rough clustering of the LIDAR points. Furthermore, because the VDMT or vigilance parameter was set so high (0.7), Fuzzy SART winds up creating multiple clusters to represent a singular cluster. Also, several outliers, existent in roof planes, must be dealt with. In order to remove the outliers and merge planes depicting the same clusters, a best fit plane was calculated from the largest clusters existent in the clustered TIN. Only clusters containing sufficient minimum number of members and having triangle centers existing above Z_{avg} are considered for best fit plane formation. A multi or planar regression analysis, formulated by minimizing the sum of the squared error, is then done on all of the clusters passing the aforementioned restrictions. The proposed 'planar' regression algorithm solves for the planar coefficients that will construct a best fit plane for the LIDAR data points presented to it. Consider the equation of a plane:

$$a \cdot x + b \cdot y + c \cdot z + d = 0 \tag{8}$$

After solving for z and making substitutions for the coefficients the following is formulated:

$$z = -\frac{d}{c} - \left(\frac{a}{c}\right) \cdot x - \left(\frac{b}{c}\right) \cdot y = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot y \tag{9}$$

The sum of the squared error between an estimate of z and the actual value of z based on n points can be represented as follows:

$$SSE = \sum_{i=1}^{n} \left[z_i - \left(\beta_0 + \beta_1 \cdot x_{i_1} + \beta_2 \cdot x_{i_2} \right) \right]^2 \tag{10}$$

The derivative of the sum of the squared error with respect to each coefficient is taken:

$$\frac{\partial SSE}{\partial \beta_0} = 2 \cdot \sum_{i=1}^{n} \left[z_i - \left(\beta_0 + \beta_1 \cdot x_{i_1} + \beta_2 \cdot x_{i_2} \right) \right] (-1) \tag{11}$$

$$\frac{\partial SSE}{\partial \beta_1} = 2 \cdot \sum_{i=1}^{n} \left[z_i - \left(\beta_0 + \beta_1 \cdot x_{i_1} + \beta_2 \cdot x_{i_2} \right) \right] \left(-x_{i_1} \right) \tag{12}$$

$$\frac{\partial SSE}{\partial \beta_2} = 2 \cdot \sum_{i=1}^{n} \left[z_i - \left(\beta_0 + \beta_1 \cdot x_{i_1} + \beta_2 \cdot x_{i_2} \right) \right] \left(-y_{i_1} \right) \tag{13}$$

Then setting those derivatives equal to 0 yields the following system of equations:

$$n \cdot \beta_0 + \beta_1 \cdot \left(\sum_n x_i \right) + \beta_2 \cdot \left(\sum_n y_i \right) = \sum_{i=1}^{n} z_i \tag{14}$$

$$\beta_0 \cdot \left(\sum_n x_i\right) + \beta_1 \cdot \left(\sum_n x_i^2\right) + \beta_2 \cdot \left(\sum_n x_i y_i\right) = \sum_{i=1}^{n} z_i \cdot x_i$$

$$\beta_0 \cdot \left(\sum_n y_i\right) + \beta_1 \cdot \left(\sum_n x_i \cdot y_i\right) + \beta_2 \cdot \left(\sum_n y_i^2\right) = \sum_{i=1}^{n} z_i \cdot y_i$$

The above system of 3 equations with 3 unknowns, or the planar 3 coefficients, can therefore be solved.

All points, with a perpendicular distance to a given best fit plane, less than a certain threshold, are then merged with the plane. Those points' z-dimension (elevation) is decreased or increased till the points are existent exactly on the defined best fit plane.

The intersections between the best fit roof planes are calculated. All points existing within a certain threshold from that roof plane intersection are then moved to exist on the line formed by the best fit roof plane intersection, resulting in straight roof ridges.

4) Experimental Results

The implemented algorithm was first tested on a simple cube test case and then subsequently on actual buildings. The test case was generated by randomly inserting points with 0m and 10m elevation (with a +/- .1m standard deviation in elevation to simulate noise). The triangulation noise filtering technique removed the .1 standard deviation and segmented the building as shown in [Figure 8].

In all of the buildings presented, all of the roof planes correspond to an arbitrary color other than red. All of the triangles not belonging to a roof plane are colored red. Therefore all triangles in the test case having a dark blue color correspond to the only existent roof plane in that structure.

For all of the following actual building cases, the reconstructed buildings ([Figure 9], [Figure 10], [Figure 11], and [Figure 12]) and the accompanying digital aerial photographs ([Figure 13], [Figure 14], [Figure 15], and [Figure 16]) are presented.

In all of the non building roof triangles (red triangles), the building wall triangles were isolated from the surrounding terrain triangles. The surrounding terrain triangles were then filtered to remove small under sampled objects such as cars, trees, shrubs, etc.

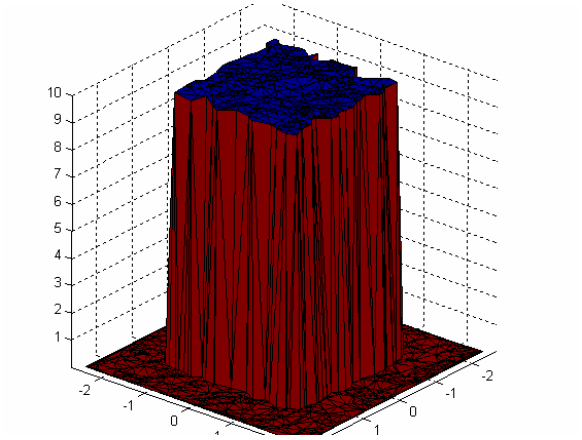

Figure 8 Test Case

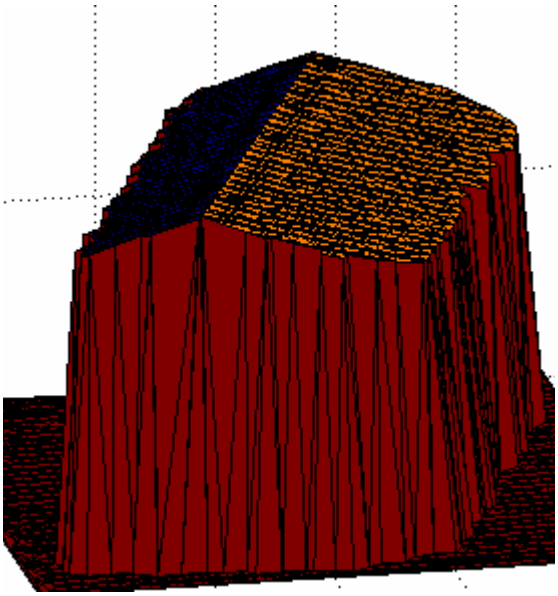

Figure 9 Building #1 3D Reconstruction

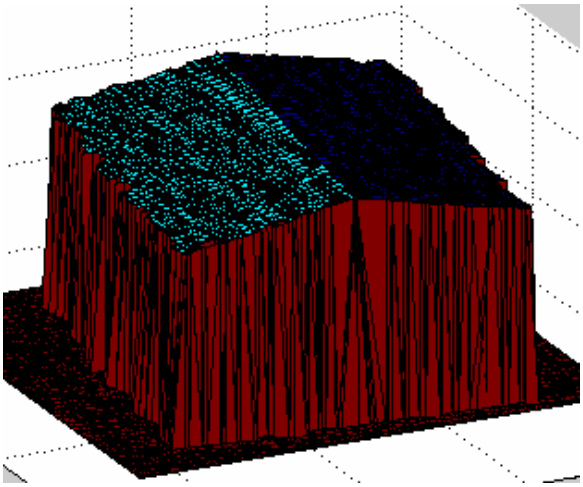

Figure 10 Building #2 3D Reconstruction

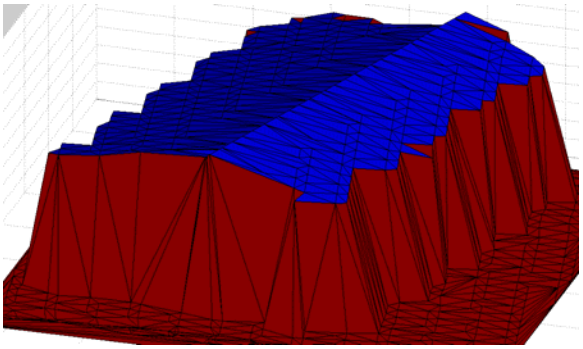

Figure 11 Building #3 3D Reconstruction

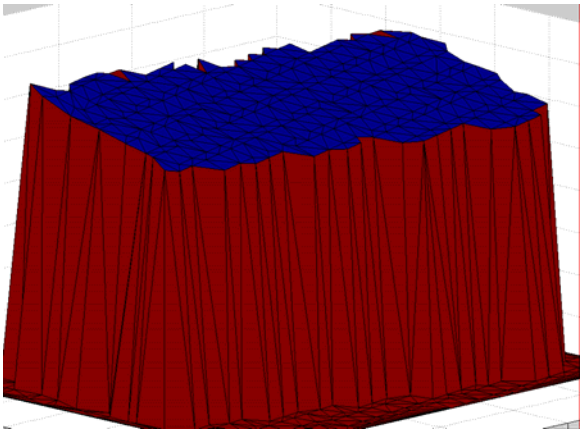

Figure 12 Building #4 3D Reconstruction

Figure 13 Building #1 Aerial Photograph

Figure 14 Building #2 Aerial Photograph

Figure 15 Building #3 Aerial Photograph

Figure 16 Building #4 Aerial Photograph

5) Conclusions and Future Work

The presented algorithm performs fairly well on large buildings with simple roof structures. The algorithm however is unable to accurately reconstruct small houses. The small houses were only fractions in size of the buildings and contain as much as 5 times the number of roof planes. The under sampled houses and consequently roof planes are unable to be accurately delineated by the Fuzzy SART clustering algorithm and therefore 3D reconstruction is not possible. Had the sampling been higher, 3D reconstruction on the smaller houses may have been possible.

In this paper we have proposed a number of preprocessing steps for 3D reconstruction from raw LIDAR data. Preliminary experimental results have shown that merging points with the proposed novel thresholding technique, during the generation of the LIDAR TIN, removes noise which causes the normal vectors to deviate from their ideal direction. This in addition to translating the coordinates to center at the origin and scaling the dimensions such that they exist in the same range, significantly improves Fuzzy SART's coplanar clustering performance. Finally, the process of performing a multi (planar) regression analysis on the clustered planes and then merging all triangle centers within a certain distance from those best fit planes yielded even better results as outlying triangles and small roof top structures, unable to be reconstructed due to under sampling, were merged with those best fit planes.

This paper describes a work in progress. Exploration of further optimization of several portions of the implemented algorithm is currently underway. Future research will be conducted in the following areas. Other unsupervised learning algorithms, variations of ART and other algorithms such as single linkage clustering, ISODATA, etc. will be tested for their accuracy in delineating roof planes. The presented research work concentrates on the 3D reconstruction aspects of recreating a given building. Additional research will be done on distinguishing buildings from one another and autonomously isolating them.

6) Acknowledgements

This research would have not been possible without the generous donations of LiDAR test set data provided from several companies. The authors would like to acknowledge Dr. Simone Clode and Dr. Franz Rottensteiner for providing the Fairfield data set, which was procured by AAMHatch. While only the Fairfield test set is considered in this paper, in future related research works and publications, the other test sets donated by other companies will be incorporated. We thank Mr. John Ellis with AeroMap, Mr. Paul Mrstik with Terra Point, and Mr. Steffen Firchau with TopoSys for their companies' contributions as well.

7) References

1. Baraldi, A.; Parmiggiani, F.; "Fuzzy combination of Kohonen's and ART neural network models to detect statistical regularities in a random sequence of multi-valued input patterns", *Neural Networks,1997., International Conference on*, vol. 1, 9-12, pp.281 – 286, June 1997
2. Chen, Liang Chien; Teo, Tee-Ann; Shao, Yi-Chen; Lai, Yen-Chung; Rau, Jiann-Yeo; "Fusion of LIDAR Data and Optical Imagery for Building Modeling." *International Archives of Photogrammetry and Remote Sensing*, vol. 35, no. B4, pp. 732-737, 2004
3. Clode, S.P.; Kootsookos, P.J; and Rottensteiner, F. "Accurate Building Outlines from ALS Data." *12th Australasian Remote Sensing and Photogrammetry Conference*, 2004
4. Fujii, Kensaku; and Arikawa, Tomohiko; "Urban Object Reconstruction Using Airborne Laser Elevation Image and Aerial Image." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2234-2240, 2002
5. Garland, M.; Heckbert, P. S.; "Fast polygonal approximation of terrains and height fields." *Technical Report, Department of Computer Science*, Carnegie Mellon University, 1995.
6. Hofmann, A.D.; "Analysis of TIN-Structure Parameter Spaces In Airborne Laser Scanning Data for 3-D Building Model Generation." *Geo-Imagery Bridging Continents XXth ISPRS Congress*, 2004
7. Lattuada, Roberto; Raper, Jonathan; "Applications of 3D Delaunay triangulation algorithms in geoscientific modeling." *The Third International Conference/Workshop on Integrating GIS and Environmental Modeling CD-ROM*, January 1996
8. Liang-chien C. , Tee-ann T., Yi-chen S., Yen-chung L., Jiann-yeou R.; "Fusion of Lidar Data and Optical Imagery for Building Modeling." *Geo-Imagery Bridging Continents XXth ISPRS Congress*, 12-23 July 2004
9. Morgan, Michel; Habib, Ayman; "Interpolation of LIDAR Data and Automatic Building Extraction." *ACSM-ASPRS2002 Annual Conference Proceedings*, 2002
10. Rottensteiner, F.; and Briese, Ch.; "Automatic Generation of Building Models from LIDAR Data and the Integration of Aerial Images." *ISPRS*, vol. 34, 2003
11. Wang, Kai; Lo, Chor Pang; Brook, George; Arabnia, Hamid; "Comparison of existing triangulation methods for regularly and irregularly spaced height fields." *INT. J. Geographical Information Science*, vol. 15, no. 8, pp 743-762, 2001