AUTOMATIC ADJUSTMENT OF DISCREPANCIES BETWEEN LIDAR DATA STRIPS - USING THE CONTOUR TREE AND ITERATIVE CLOSEST POINT ALGORITHM

Jaebin Lee, Dongyeob Han, Kiyun Yu, Yongil Kim

The school of Civil, Urban, and Geosystem Engineering, Seoul National University
{dama77, hkyon2, kiyun, yik}@snu.ac.kr

ABSTRACT: To adjust the discrepancy between Light Detection and Ranging (LIDAR) strips, previous researches generally have been conducted using conjugate features, which are called feature-based approaches. However, irrespective of the type of features used, the adjustment process relies upon the existence of suitable conjugate features within the overlapping area and the ability of employed methods to detect and extract the features. These limitations make the process complex and sometimes limit the applicability of developed methodologies because of a lack of suitable features in overlapping areas. To address these drawbacks, this paper presents a methodology using area-based algorithms. This approach is based on the scheme that discrepancies make complex the local height variations of LIDAR data whithin overlapping area. This scheme can be helpful to determine an appropriate transformation for adjustment in the way that minimizes the geographical complexity. During the process, the contour tree (CT) was used to represent the geological characteristics of LIDAR points in overlapping area and the Iterative Closest Points (ICP) algorithm was applied to automatically determine parameters of transformation. After transformation, discrepancies were measured again and the results were evaluated statistically. This research provides a robust methodology without restrictions involved in methods that employ conjugate features. Our method also makes the overall adjustment process generally applicable and automated.

KEY WORDS: LIDAR, strip adjustment, Area-based approach, Contour Tree, Iterative Closest Point algorithm

1. INTRODUCTION

Since the introduction of Light detection and ranging (LIDAR) systems, they have been adapted in a wide range of application areas such as creating digital surface models (DSM) and ortho-photo generation. It is due to its ability to quickly acquire 3D terrain coordinates over target areas. Present researches related to LIDAR systems are seeking ways to improve the collection and analysis efficiency of LIDAR data. Despite of these developments, there are still some noticeable systematic errors in the overlapping areas of LIDAR strips. These errors are generated from inaccurate calibration of the entire measurement system and the limited accuracy of direct geo-referencing with Global Positioning System (GPS) and Inertial Measurement Unit (IMU), including systematic errors while the LIDAR system is flying in multiple overlapping strips in order to completely cover the desired area (Norbert, 2005). Such systematic errors usually result in less meaningful extracted features and questionable quality of the final product. With this increased need to adjust discrepancies, a series of studies has been proposed in recent years using conjugate features, which are called feature-based approaches (Mass, 2002, Filin, 2004, Kager, 2004, Norbert, 2005, Vosselman, 2004, J. Lee, 2005).

In these researches, the conjugate features, irrespective of the type of features, should be detected and extracted for adjustment from raw LIDAR data. It makes the overall process complex and sometimes limits the availability of developed methodologies due to lack of features in overlapping area. Even if there are enough conjugate features in the target area, the automatic detection and extraction of features have been problems to be solved. To circumvent these drawbacks, this paper presents a methodology using the geographical characteristic of terrain itself, which can be called area-based methods. When two neighbouring strips are overlapped, the local height variation of LIDAR data in overlapping area increases if there are discrepancies. It makes the topology of the terrain complex. Thus, one should be able to adjust the discrepancy by finding the transformation function which minimizes the geographical complexity of terrain i.e., removes the discrepancy.

For this purpose, Contour Tree (CT) is used to represent the topology and measure the topological complexity of the terrain. CT is a fundamental data structure in scientific visualization. It is mainly used to capture the topological characteristics of a scalar field which represent data in different application areas like geographic information systems, medical imaging or scientific visualization (V. Pascucci, 2002). A CT consists of a finite set V of objects called vertices, and E of objects called edges. Usually the vertices represent contour lines or height points and the edges represent the adjacent relationship of the two vertices the edge links. (Bollobas, 1998). For every vertex Vi in a contour tree, we can count the total number of neighbouring vertices of

Vj. Especially, when the number of the neighbouring vertices is 1, the vertex Vi is called the leaf. Leaves are usually assumed to be isolated and have locally extreme value of elevation in CT structure (Y. Shinagawa, 1991, V. Pascucci, 2002). Therefore, one can easily imagine that the more critical points in a target area, the more complex the topography. Based on this scheme, the topological complexity was measured by using the number of leaves of CT. When LIDAR data strips are overlapped, the discrepancy existed between strips increases the number of points which have locally extreme value of elevation, that is, the number of leaves of CT in overlapping area. Therefore, one can find the appropriate transformation function between neighbouring strips to minimize the number of leaves of CT in overlapping area. By using this methodology, the adjustment of strips can be performed without the process of detecting and extracting conjugate features.

To find the initial values of parameters of the transformation, ICP algorithm is applied. ICP is widely used in ground range image registration as an accurate and reliable method for free form surfaces (P.J. Besl, 1992). In this paper, the goal of ICP is to find the initial parameters of transformation that best aligns a cloud of points with a cloud of points in a reference strip. The alignment process works to minimize the mean squared distance between points in target and reference strips. At each iteration, the algorithm computes correspondences by finding closest points and then, minimizes the mean square error in position between corresponding points (O.D. Faugeras, 1986). Then, these initial parameters are refined by computing the geographical complexity of terrain using CT. It makes the overall adjustment process of LIDAR data strips more easily and automatically.

In the following section, an overview of a contour tree and the methodology used to create CT in the paper are presented. Using ICP and CT, the methodology for adjustment of the discrepancy is given out in section 3. In section 4, an experiment with real LIDAR data strips which are obtained by a high end LIDAR system is performed to demonstrate the feasibility of this approach. Then, conclusions and future works are discussed in section 5.

2. CONTOUR TREE

The CT was introduced by Boyell (1963) as a summary of the elevation of contours on a map (i.e. in 2-D). Since its introduction, it has been used for image processing and geographic information system. It is a kind of a tree structure and a data structure that represents the relationships between the connected components of the level sets in a scalar field. The display of CT provides the user with direct insight into the topology of the field and reduces the user interaction time necessary to "understand" the structure of the data (Bajaj, 1997). Especially, the points which are usually assumed to be isolated and have locally extreme value of elevation can be detected easily under CT structure. To compute CT, we use the algorithm proposed by (Carr H., 2003), which is an elegant and efficient algorithm for the computation of CT in any dimension. The algorithm consists of three stages: (i) sorting the vertices in the field, (ii) computing the Join Tree (JT) and Split Tree (ST), and (iii) merging the JT with ST to build CT.

For example, from a given 3-D point cloud, we can create the mesh M which consists of vertices and edges (see Figure 1). The 2D Scalar field (terrain) is represented as a Triangulated Irregular Network (TIN) with elevation values associated with each vertex, in which the number means the elevation of each vertex and the number in the blank means the order. The critical points are marked with colour disks: Local maximum in red and minimum in blue. Using this mesh M, CT is created as shown in Figure 2.

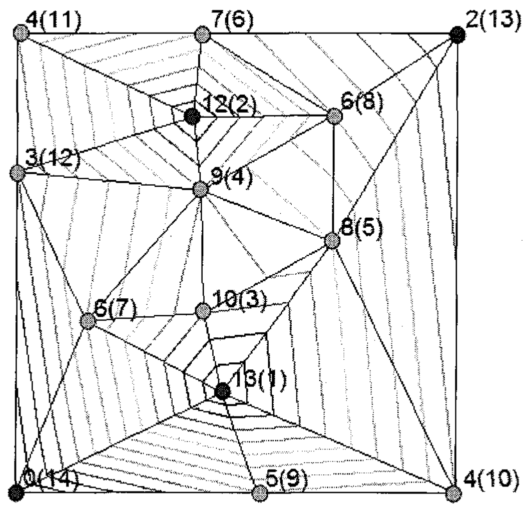

Figure 1. The mesh M and CT

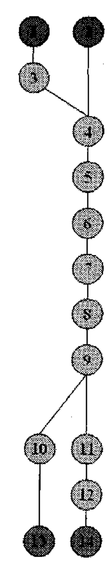

Figure 2. Corresponding CT to the terrain in Figure 1

3. USING ICP AND CT TO ADJUST FOR DISCREPAMNCIES

The goal of adjustment is to find the appropriate transformation T between strips such that the discrepancy between the transformed strip and reference strip is removed. For this purpose, first of all, we applied ICP to find initial values of parameters of transform. Then, we utilize the scheme that when the discrepancy is perfectly removed, the number of leaves of the CT which is computed for overlapping area achieves its minimum. Based on this scheme, the process to find the appropriate transformation goes as follows.

Given two LIDAR data strips S1 (a test strip) and S2 (a reference strip).
Step #1:
Extract the overlapping area from S1 and S2
Apply ICP to find the initial spatial transform T. In this paper, it consists of seven parameters which are scale (S), rotation (ω,φ,κ) and translation (X_T,Y_T,Z_T).
To refine the initial T, determine the range and the cell size of parameters considering initial values of parameters, which are depending on the quality of the approximations of parameters.
Apply the T to S1 and compute the CT and N (CT) of the overlapping area
Step #2:
Repeat Step #1, refining the parameters of T.
Find the parameters of T when N (CT) achieves its minimum.
Step #3:
Decrease the range and cell size of parameters.
Repeat Step #1~Step #3 until the parameters converge to the desired precision.

4. EXPERIMENTAL RESULTS

We applied the developed algorithm to real LIDAR dataset. It was captured using an OPTECH ALTM 2050 laser scanner at a mean flying altitude of 975 m and a mean point's density of 2.24 points/m2. According to the sensor and flight specifications, 0.5 m horizontal and 0.15 m vertical accuracies are expected for this dataset. For the target area, the LIDAR dataset consists of six strips, of which two neighbouring strips were used for this experiment. A total of 1.2 million points were identified from overlapping area between two strips and a center part of the overlapping area which includes 250,000 points was used because of computational efficiency (See Figure 3). In Figure 3, red and blue points represent two overlapping strips respectively and yellow points are the points in overlapping area and green points are used for adjustment process.

In this experiment, a 3-D conformal transformation is chosen as a spatial transformation for adjustment (see (1)).

$$\begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix} = \begin{bmatrix} X_T \\ Y_T \\ Z_T \end{bmatrix} + SR(\omega,\phi,\kappa)\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{1}$$

where, S is the scale factor, $(X_T, Y_T, Z_T)^T$ is the translation vector between the origins of coordinate systems of each LIDAR data strip and R is the 3-D orthogonal rotation matrix, $(X,Y,Z)^T$ are the point coordinates in the test strip and $(X_r,Y_r,Z_r)^T$ are the point coordinates in the reference strip.

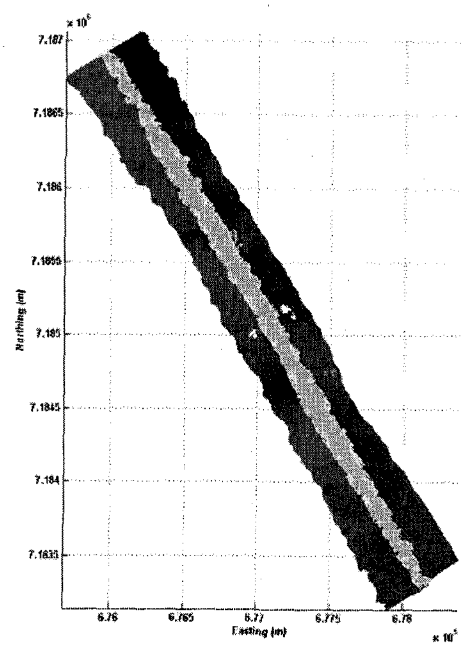

Figure. 3Extracted LIDAR data points for adjustment.

After setting up the spatial transformation, the initial parameters of transformation were determined using ICP. The transformation was applied to the test strip, and then CT and the number of leaves of CT were computed for overlapping area. Within a predetermined range and cell size of accumulator arrays of parameters, seven parameters of transformation were refined until the number of leaves achieves its minimum. Then, we decrease the range and cell size of parameters and repeat the process. Table 1 lists the parameters of transformation which are determined by proposed algorithm.

Table 1. Estimated Parameters

3-D conformal transformation	
S (Scale)	1.000
X_T (m)	0.410
Y_T (m)	0.210
Z_T (m)	0.000
ω (degrees)	0.000
ϕ (degrees)	-0.000
κ (degrees)	-0.000

To test the feasibility of the determined 3-D conformal transformation, a total of 164 pairs of conjugate line features were identified from the two overlapping strips and extracted respectively. These linear features were used to measure the discrepancies before and after applying the transformation (see (J. Lee, 2005) for complete description). To measure discrepancies, a normal vector was calculated from the midpoint of one conjugate linear feature to the other line for every pair of conjugate features. Table 2 lists the overall discrepancies between the two strips before and after applying the transformation.

Table 2. Measurements of Discrepancies

	Before Transform		After Transform	
	Mean	SD	Mean	SD
dx (m)	–0.182	±0.286	-0.028	±0.219
dy (m)	–0.082	±0.315	-0.009	±0.166
dz (m)	0.003	±0.111	-0.015	±0.070

We next conducted a hypothesis test to examine whether the results are statistically significant. Using the paired comparison method, the differences between normal vectors before and after the transformation were examined to determine if they are significantly large from a statistical viewpoint. The test statistic was set up as follows:

$$T = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_{1,i} - X_{2,i}) - \delta_0}{S_D / \sqrt{n}} \tag{2}$$

where $S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}[(X_{1,i} - X_{2,i}) - \frac{1}{n}\sum_{i=1}^{n}(X_{1,i} - X_{2,i})]^2$

indicates the pooled standard deviation, $X_{1,i}$ and $X_{2,i}$ denote the value of i the normal vectors before and after transformation, respectively, and n indicates the number of normal vectors. In this case, δ_0 is zero.

The corresponding hypothesis is:

$$H_0 : \mu_1 = \mu_2 \qquad H_1 : \mu_1 > \mu_2 \tag{3}$$

where μ_1 and μ_2 are the population means of the normal vectors before and after applying the transformation, respectively.

The test results reject the null hypothesis at the 99% significance level. T-values are 7.245 in the X-direction and 7.143 in the Y-direction, respectively, which rejects the null hypothesis at the significance level (99%, at which the t value is 2.364). This result indicates that the discrepancies in the LIDAR data strips had been reduced by a statistically significant amount in X-and Y-direction.

5. CONCLUSIONS

This paper presents an automatic and generally applicable algorithm for adjustment of the discrepancy between LIDAR data strips, which overcomes drawbacks of using conjugate features for adjustment. By using ICP and CT algorithm, the algorithm explicitly formulates step-by-step methodologies to determine the transformation for adjustment. It was applied to real LIDAR dataset and the results clearly demonstrate a statistically significant reduction in discrepancies following the proposed algorithm. Now we are focusing on investigating how the results affected by the point's density of LIDAR data points and the number of points used for adjustment.

References from Journals:

H. Carr, J. Snoeyink, and U. Axen, 2003 Computing Contour Trees in All Dimensions. *Computational Geometry Theory* 24(2), pp. 75-94

H. G. Maas, 2002 Methods for Measuring Height and Planimetry discrepancies in Airborne Laserscanner Data. *Photogrammetric Engineering & Remote Sensing*, vol. 68, no. 9, pp. 933–940

Norbert Pfeifer, 2005 Airborne Laser scanning strip adjustment and automation of tie surface measurement. *Boletim de Ciências Geodésicas*, vol. 11, no. 1

O.D. Faugeras and M. Hebert, 1986 The Representation, Recognition, and Locating of 3-D Objects. *Int'l J. Robotics Research*, vol. 5, no. 3, pp. 27-52

P.J. Besl and N.D. Mckay, 1992 A Method for Registration of 3D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239-256

References from Books:

B. Bollobas, 1998 *Modern graph theory* Springer-Verlag, pp408

References from Other Literature:

C.L. Bajaj, V. Pascucci, and D. R. Schikore, 1997 The Contour Spectrum. IEEE proceeding Visualization 1997

H. Kager 2004 Discrepancies between Overlapping Laser Scanner Strips-Simultaneous Fitting of Aerial Laser Scanner Strips. XXth ISPRS Congress, Istanbul, Turkey

J. Lee, K. Yu, Y. Kim, A. Habib 2005 Segmentation and extraction of linear features for adjustment of discrepancies between ALS data strips. IGARSS 2005 IEEE Int. Conf., Seoul, Korea

R. L. Boyell and H. Ruston, 1963 Hybrid Techniques for Real-time Radar Simulation. IEEE Proceedings Fall Joint Computer Conference 63, Las Vegas, USA

VOSSELMAN G. 2002 On the estimation of planimetric offsets in laser altimetry data. International archives of Photogrammetry and Remote Sensing, Graz, Austria

V. Pascucci and K. Cole-McLaughlin, 2002 Efficient computation of the topology of level sets. IEEE proceedings Visualization 2002