COMBINED SEGMENTATION OF LIDAR POINT CLOUD AND REGISTERED IMAGES

*** Xiangyun Hu[1], Junfeng Zhu[1], Lizhi Ye[1]**
1. 129 Luoyu Road, School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China.
[1]Phone (O): +86 27 68778010
Phone (H): +86 158 02752751
Email: xiangyun.hu@gmail.com

ABSTRACT

By fusing with other sensory data, especially high resolution imagery, Lidar can be good source of information for DEM extraction and feature extraction. Nowadays airborne Lidar system vendors such as Leica and Toposys and others are providing systems (Leica ALS50II, ALS60, Toposys FALCON II) with integrated camera capturing 3D point cloud and high resolution images simultaneously. The full potential of the integrated system has to be explored yet. This paper presents an automatic segmentation method based on the fused data of point cloud and imagery. The method automatically partitions the scene by taking into account spectral, spatial and elevation information of pixels. The segmented regions contain multiple cues of object, which can further be used for feature extraction. The experimental result shows that the combined segmentation is useful for better DEM filtering and object classification than using single source of data (Lidar or imagery).

KEYWORDS: Airborne Lidar; Sensor Integration; Image Segmentation; Information Fusion.

INTRODUCTION

During the last decade, airborne laser altimetry has become a promising method to capture digital elevation data effectively and accurately. In our text we use LiDAR (Light Detection And Ranging) as the abbreviation for the various laser altimetry methods. More and more applications take advantage of the high accuracy potential, dense sampling, and the high degree of automation that results in a quick delivery of products derived from the raw laser data (Schenk and Csatho, 2002). But because of the complexity the automatic data interpretation and the sometimes low number of bare-earth points, LiDAR data's full automatic process meets great limits. (Sithole and Vosselman, 2004)

Photogrammetry and LiDAR have their own advantages and drawbacks for segmentation, DEM extraction and feature extraction. It is interesting to note that some of the shortcomings of one method can be compensated by advantages the other method offers. Hence it makes eminent sense to fuse, or combine as you wish, the methods. A number of researches have shown the approaches of combine data for DEM extraction and feature extraction, e.g., LiDAR and aerial image (Schenk and Csatho, 2002), LiDAR and three-line-scanner image (Nakagawa, Shibasaki and Kagawa, 2002), LiDAR and high satellite image (Guo, 2003).

Advantages of data obtained from integrated sensor have been mentioned in numerous studies yet, so this paper presents an automatic segmentation method based on the fused data of point cloud and imagery. In the second

---

* Primary Author and Presenter

section we briefly discuss the advantages and disadvantages of the two kind of data for segmentation and our motivation for the combined segmentation. In the third section we depict our method for orienting the LiDAR data and the aerial images in to a common reference frame. And then, we introduce two data fusion methods and occlusion problems. At last, we present an automatic segmentation method based on the fused data of point cloud and imagery.

BACKGROUND

In this section we briefly discuss the advantages and disadvantages of the two kind of data for segmentation. We also introduce a number of previous researches and our motivation for the combined segmentation.

Segmentation and Previous Work

Segmentation of LiDAR data is to group points with similar features into segments. In the field of laser scanning usually homogeneous regions (e.g. roof, square and bare-Earth) are segmented. Many researches need this attribute information for DEM extraction and feature extraction, therefore a number of segmentation algorithms have been developed, but they just based LiDAR data.

The first type of segmentations based on region growing. These methods group points based on geometrical relations of neighborhood like height, slope or curvature difference. Lee and Schenk, 2001, introduce a method works on triangles and driven by a robust plane fitting. Filin (2002) proposed a method based on clustering analysis. Vosselman and Dijkman (2001) proposed Hough-transformation to detect planar roof surfaces within the given building boundaries.

Motivation for the Combined Segmentation

As is known to all, the results of the segmentation are connected with the point cloud density. With low density, the run of the process is easier, but the lack of special and texture information makes difficult to find the segment borders. With high density, the procedure is more complex and may produce too many small facets.

LiDAR data provide high accurate 3D points but lack breaklines and texture information. On the contrary, optical imagery with high spatial resolution provides more accurate breaklines and texture information than LiDAR data. By combining LiDAR data and optical imagery most of the disadvantages associated with either method are compensated.

CROSS SENSOR DATA REGISTRATION

Our fusion approach is also known as cross sensor data alignment or registration. That is to say we have to establish a common reference frame for LiDAR data and aerial imagery. It entails a transformation, forward and backward, between sensor data and reference frame. LiDAR point cloud is usually computed in the WGS84 reference frame. Hence it makes sense to reference the aerial images to the WGS84 reference frame. And then we can consider referencing aerial imagery to LiDAR as an orientation problem. Fusion is important because it will affect the accuracy of segmentation. Schenk and Csatho, (2002) proposed a fusion method which used the sensor invariant features to determine the relevant orientation parameters. We use their method to complete our fusion process, the specific process can be found in Schenk and Csatho, (2002).

TWO MANNERS OF FUSED DATA GENERATION

From the discussion of the previous section we have established a common reference frame for LiDAR data and aerial images. In this common reference, we can generate two kind fused data - imagery fused with range image re-sampled from point cloud and point cloud with assigned image pixel attributes.

Imagery Fused with Range Image

In aerial photogrammetry, the collinearity equation has below form:

$$\left.\begin{aligned} X - X_s &= (Z - Z_s)\cdot\frac{a_1 x + a_2 y - a_3 f}{c_1 x + c_2 y - c_3 f} \\ Y - Y_s &= (Z - Z_s)\cdot\frac{b_1 x + b_2 y - b_3 f}{c_1 x + c_2 y - c_3 f} \end{aligned}\right\} \qquad (1)$$

Here, $R = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ refers to the rotation matrix, (X_s, Y_s, Z_s) refers to the coordinate of perspective center in object space coordinate system, (x, y) refers to the coordinate of image point in image plane coordinate system, f refers to the principal distance, (X, Y, Z) refers to the coordinate in object space coordinate system of object space point corresponding to the image point. R and (X_s, Y_s, Z_s) can be acquired from POS.

The formula above is a procedure from 2D (x, y) to 3D (X, Y, Z), so we need to assume an approximation for Z_0 in order to obtain (x, y) from (X, Y, Z). Obtaining (X_1, Y_1), we can get height Z_1 at point (X_1, Y_1) from LiDAR data using bilinear interpolation method. Then calculate (X_2, Y_2) by the formula (1). Iterative like this until the difference between X_n and X_{n-1}, Y_n and Y_{n-1}, Z_n and Z_{n-1} in last two calculation is small enough (for example, 0.01) and we take

$$\left.\begin{aligned} X &= \frac{X_n + X_{n-1}}{2} \\ Y &= \frac{Y_n + Y_{n-1}}{2} \\ Z &= \frac{Z_n + Z_{n-1}}{2} \end{aligned}\right\} \qquad (2)$$

as the coordinate of the ground point corresponding to the image point (x, y). Repeat the above steps and we can add 3D coordinate information to every pixel in the image. We can get the depth image depending on the Z of every pixel, as well. The result was showed in figure 1.

Figure 1. The first one (from left to right) is the original LiDAR point cloud, the second is the original imagery and the last depicts the range image.

Point Cloud with Assigned Image Pixel Attributes

In aerial photogrammetry, the collinearity equation has another form:

$$\left.\begin{aligned} x &= -f \cdot \frac{a_1(X - Xs) + b_1(Y - Ys) + c_1(Z - Zs)}{a_3(X - Xs) + b_3(Y - Ys) + c_3(Z - Zs)} \\ y &= -f \cdot \frac{a_2(X - Xs) + b_2(Y - Ys) + c_2(Z - Zs)}{a_3(X - Xs) + b_3(Y - Ys) + c_3(Z - Zs)} \end{aligned}\right\} \tag{3}$$

Here, $R = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ refers to the rotation matrix, (X_s, Y_s, Z_s) refers to the coordinate of perspective center in object space coordinate system, (x, y) refers to the coordinate of image point in image plane coordinate system, f refers to the principal distance, (X, Y, Z) refers to the coordinate in object space coordinate system of object space point corresponding to the image point. R and (X_s, Y_s, Z_s) can be acquired from POS.

From formula (3), we know that it is an easy procedure from 3D to 2D to add RGB information to every LiDAR point. At this moment, we just need to put (X, Y, Z) of every LiDAR point, high accuracy POS and camera parameter into formula (3) to obtain coordinate (x, y) in image space coordinate system corresponding to every LiDAR point. Then assign the (R, G, B) at coordinate (x, y) in the image to the corresponding LiDAR point. Figure 2 shows the result.

Figure 2. Point cloud with assigned pixel attributes.

OCCLUSION PROBLEMS AND HOW TO SOLVE IT

With the improvement of image resolution and advances in LiDAR technology, the true orthophotos develop rapidly. Compared with the traditional orthophotos, the true orthophotos compensate the occluded portions caused by the sensor tilt and terrain relief. Since the existence of height displacement, may make some LiDAR points invisible in the image. Before compensation, we need to detect these invisible points and the most commonly used method is z-Buffer algorithm proposed by Amhar et al (1998). The z-Buffer method resolves the ambiguity arising from having more than one object point competing for the same image pixel. Among the competing object points, the point closest to the perspective center of an image is considered to be visible, while the other points are considered to be invisible in that image.

In order to improve the speed, we achieved the z-Buffer algorithm using depth-buffer in OpenGL. The workflow is as follows:

1) Build TIN （Triangulated Irregular Network）.
2) Render TIN with polygon mode GL_FILL.
3) Read and store the depth value in the depth-buffer.

Calculate the depth value of every point and compare it with the value in depth-buffer to determine whether the point is visible.

Because vertexes and polygons are not rasterized in exactly the same way in OpenGL, the depth values generated for pixels on a vertex are usually not the same as the depth values for a polygon vertex. In order to ensure the accuracy of the final judgments, we can enable polygon offset and render the point cloud between step 2) and 3). In addition, in order to save memory, we can partition the TIN into blocks after step 1) and process them block-by-block. Figure 3 shows the result.

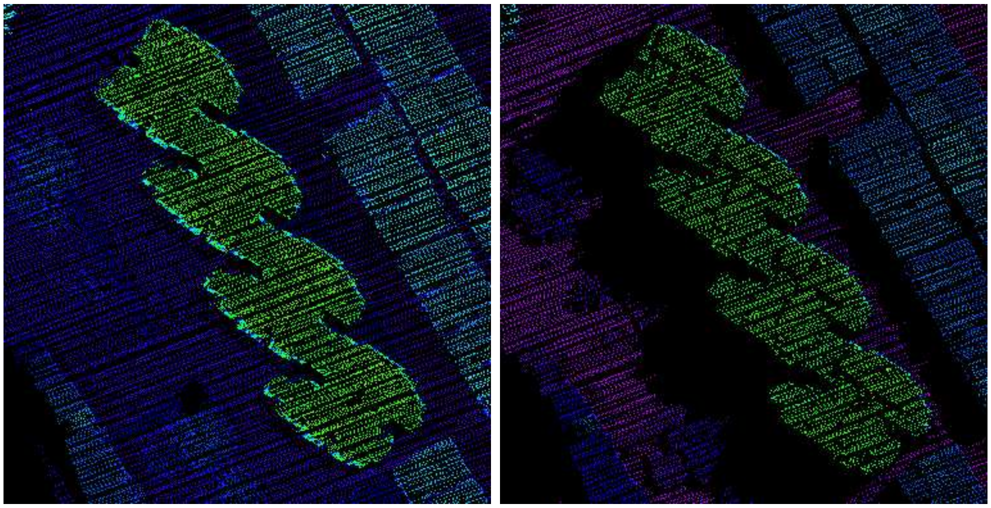

Figure 3. Original LiDAR data(the first one) and the result after the elimination of invisible points.

COMBINED SEGMENTATION

Our segmentation method applied and described here was original developed for terrestrial laser scanner data. It is based on region growing and uses the *n* nearest neighbors of the points and their spectral information captured from aerial imagery. Algorithm realization steps are as follows:

1) Randomly select a seed point and establish a seed bank (seed array);

2) Find the k-nearest neighbors of the current seed point as the candidate points. Check whether the candidate point is to meet the specific criteria (see below) one by one, if satisfied, put the candidate point to the seed bank and marked as belonging to the current region;

3) If the seed bank is not empty, goes to step (2), otherwise the growing procedure is complete.

In determining whether the candidate point is a new seed point, we use all the points belonging to the current region (more than 3 points) to fit a plane, and estimate the plane's normal vector, the specific process can be seen in Roggero (2002). The criteria of determine whether the current candidate point is the seed point is as follows:

- similarity of normal vectors (α),

- distance of candidate point to the adjusting plane (*r*),

- difference of current candidate point's elevation and the region's average elevation (*dz*).

- texture (see blow) differences (*t*).

When a new point is accepted as the current regional point, the plane need to be calculated again. With regards to the texture feature, the Local Binary Pattern/Contrast (LBP/C) distribution is chosen and approximated using a discrete two-dimensional histogram of size 256 by *b*, where b is set to 8. This texture descriptor was proposed by Ojala et al. (1996) and has shown great power in texture discrimination and computation simplicity. You can see a detailed computational process in my previous text ().The process is affected by 3 parameters: α, *r, dz, t.*

Our fused data can compensate for their shortcomings and make full use of their advantages. In our combined segmentation algorithm which is based on multiple layer data group points based on geometrical relations of neighborhood (e.g., height, slope and plane vector) and spectral property captured from high resolution imagery. A result of the segmentation has been shown in figure 4. Different segments are shown in different color. Homogeneous regions have been divided into their own blokes and different parts of the terrain have been split at the respective break lines.

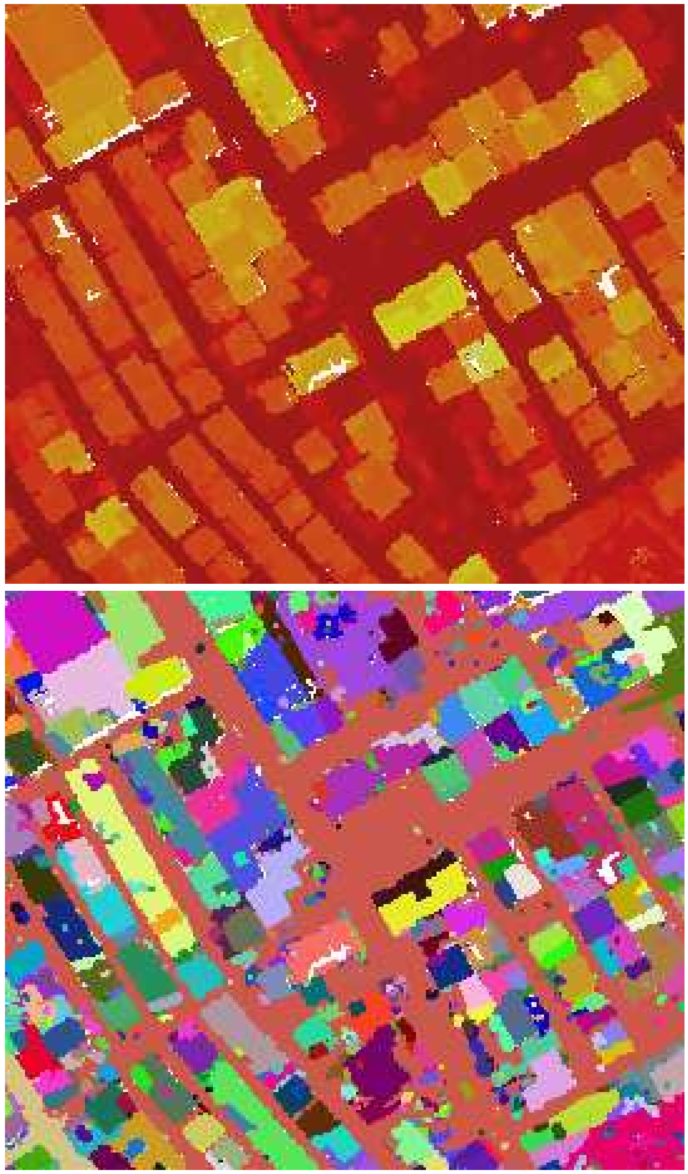

Figure 4: Original laser scanning (above) data and segmentation of laser scanning data.

REFERENCES

Schenk, T., Csatho, B., 2002. Fusion of Lidar data and aerial imagery for a more complete surface description. In: ISPRS. *Photogrammetry and Computer Vision*, Graz, Austria, 9-13 September, pp. A-310 ff.

Sithole, G., Vosselman, G., 2004. Experimental comparison of filtering algorithm for bare-Earth extraction form airborne laser scanning point clouds. ISPRS *Journal of Photogrammetry and Remote Sensing* 59, pp. 85-101.

Nakagawa, M., Shibasaki, R., and Kagawa, Y., 2002, Fusion Stereo Linear CCD Image and Range Data for Building 3D Urban Model, IAPRS, Vol.34, Part4, pp.200-211.

Guo, T., 2003, 3D city modeling using high-resolution satellite image and airborne laser scanning data, Doctoral dissertation, *Department of Civil Engineering*, University of Tokyo, Tokyo.

Amhar, F., et al., 1998. The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM, *International Archive Photogrammetry and Remote Sensing*, 32: 16–22.

Tovari, D., Pfeifer, N., 2005. Segmentation based robust interpolation – a new approach to laser data filtering. In: IAPRS, XXXVI, 3/W19, Enschede, The Netherlands, pp.79-84.

Lee, I., Schenk, T., 2001. 3D perceptual organization of laser altimetry data. In IAPRS, Vol.XXXIV 3/Went, Annapolis, MD, USA.

Filin, S., 2002, Surface clustering from airborne laser scanning data. In IAPRS, Vol.XXXIV 3A, Graz, Austria.

Chen, L. C., Tao, T. A., Lai, Y. C. Rau J. Y., 2004. Fusion of LIDAR data and optical imagery for building modeling. *International Archives of Photogrammetry Remote Sensing and Spatial Information Science,* 35(B2): 586-591.

Vosselman, G., Dijkman, S., 2001. 3D building model reconstruction from point clouds and ground plans. In IAPRS, Vol.XXXIII, 3/W4, Annapolis, MD, USA.

Roggero, M., 2002. Object segmentation with region growing and principle component analysis. In: ISPRS. *Photogrametery and Computer Vision*, Graz, Austria, 9-13 September, pp.A-289.

Ojala, T., Pietikainen, M., 1996. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29:51-59.

Hu, X., Tao, C.V., Prenzel B., 2005. Automatic Segmentation of High-resolution Satellite Imagery by Integrating Texture, Intensity, and Color Features. *Photogrammetry and Remote Sensing*. 1309-1406.