EXTRACTION OF 3D FEATURE FROM LIDAR DATA FUSED WITH AERIAL IMAGES USING IMPROVED MEAN-SHIFT ALGORITHM

Chun Liu[a,b,*], Yunling Zhang[a]

a Dept. of Survey and Geo-Informatics Engineering, Tongji University, Shanghai, China-liuchun@tongji.edu.cn
b Key Laboratory of Advanced Engineering Surveying of SBSM, Shanghai, China

KEY WORDS: LiDAR, Feature Extraction, Mean-Shift

ABSTRACT:

An innovative approach is proposed for the extraction of the complex urban three dimensional feature efficiently and accurately. In this method, firstly, both the LiDAR data and the aerial images are respectively pre-processed and matched using an affine transformation technique. Then, an improved mean-shift algorithm is employed to classify the LiDAR data fused with reflected intensity and spectrum attribute into groups by kinds of feature, such as buildings, vegetation, water etc. The spectral information are extracted from aerial-image to extend the attribute of the LiDAR data. The classification accuracy is evaluated by a confusion matrix. Finally, the 3D model of interested region is quickly constructed based on the classified points and the aerial-image in software SketchUp. During the experiment, the key issue is how to control the results of classification through parameters setting.

1 INTRODUCTION

LiDAR (Light Detection and Ranging) is an active remote sensing system, which utilizes laser beam for detection and measurement to provide three dimensional information of earth surface and object. As its superiority, there are wide fields for LiDAR application, such as 3D city models, urban planning, design of telecommunication networks, vegetation monitoring and disaster management. By contrast with traditional photogrammetry, the 3D urban data capturing using LiDAR is of higher speed, higher vertical accuracy and lower cost(Mei, Bing et al. 2009).

Since nineties of last century, automated or semi-automated 3D reconstruction of urban buildings(Debevec, Taylor et al. 1996; Noronha and Nevatia 2001) from photographs has been widely interested and deeply studied. However, due to the complexity of the technology and the rise of LiDAR, reconstruction with the new data source LiDAR, has been continuously tried and achieved good results(Haala and Brenner 1997; Maas and Vosselman 1999; Rottensteiner and Briese 2002). Meanwhile, LiDAR was experimented in the fields of ecology and topography. Various characteristics of forests were measured using LiDAR, such as stand volume, Canopy structure, biophysical properties etc. (Nilsson 1996; Lefsky, Cohen et al. 1999) The SHOALS(Scanning Hydrographic Operational Airborne LiDAR Survey) system uses LiDAR to remotely measure bathymetry and topography in the coastal zone(Irish and Lillycrop 1999). Since then, feature extraction by integration of multispectral/hyperspectral images, LiDAR or SAR became emerging research focus(Tao and Yasuoka 2002; Rottensteiner, Trinder et al. 2005; Dalponte, Bruzzone et al. 2008; Wulder, White et al. 2009).

In the 3D urban feature extraction field, the vegetation can not be preserved while the buildings are extracted and reconstructed. However, vegetation is one of the key issues in urban environment planning and monitoring all the time (DeCandido 2004). The reasons of non-preserved vegetation are as following: On the one hand, the non-grounded points in LiDAR data mainly consist of building roofs and trees, while the recognition methods can only identify the roof-shaped regular point groups(Rutzinger, Hofle et al. 2006; Sohn and Dowman 2007). On the other hand, the pixels on the images represent buildings or trees usually have close position and spectral signature to a certain extent, while The image segmentation methods are generally prone to have some misclassification error in that case(Herold, Scepan et al. 2002). In this paper, a mean-shift algorithm that always used in image segmentation(Comaniciu and Meer 2002) is employed to classify the LiDAR data fused with reflected intensity and spectrum attribute into groups by feature. Therefore, both the building feature and the vegetation feature can be respectively extracted at the same time.

2 STUDY AREA AND MATERIAL

The data used in this paper included a true color image and a set of LiDAR point clouds. The data was produced by the Woolpert Company for the Ohio Imagery Data Project. (The data can be downloaded from: http://metadataexplorer.gis.state.oh.us/metadataexplorer/explorer.jsp.) It was captured by a Leica ALS digital LiDAR system during March to May 2006. The LiDAR data set includes 62708 points and stores vegetation, bare soil, buildings, and roads. The area of the research region is about 2.3 square kilometres. The aerial image is 900 pixels by 700 pixels and every pixel equals 0.3 m. The LIDAR point data and the aerial image are shown in figure 1.

* Corresponding author. liuchun@tongji.edu.cn

a. Aerial image data

b. Point clouds

Figure 1. Experimental data in the study area.

3 METHODOLOGY AND PRINCIPLE

The 3D feature extraction is divided into three stages: pre-process, classification and feature extraction, and the flow chart is given to explain the used methodology in figure 2.
The advantage of height smooth can eliminate the gross error of raw lidar data by setting the height threshold, and that can help to reduce the probability of misclassification. Then, to achieve data fusion, the registration employs the affine transformation technique by manually selecting several homologous points. Thus, the spectrum value can be extracted from matched image to lidar data, which is a sort of fusion method of lidar and aerial images.

During the classification, before the improved mean-shift algorithm conducted to classify the fused LiDAR data into groups, the employed color space of spectrum needs to be transferred from RGB to L*a*b*. To obtain a meaningful classification, perceived color differences should correspond to Euclidean distances in the color space chosen to represent to the features. The space L*a*b* were especially designed to best approximate perceptually uniform color spaces. The dependence of all three coordinates on the traditional RGB color values is nonlinear.

A computational module based on the mean shift procedure is an extremely versatile tool for feature space analysis and can provide reliable solutions for many task(Comaniciu and Meer 2002). Thus, this algorithm is induced to classify the point clouds in the experiment. Besidse, the key issue of the mean-shift algorithm is appropriate kernel function setting and bandwidth selection.

When an appropriate classification has applied to the point clouds, the ground feature can be extracted through an anylsis of the result of classification. In the end, 3D scene can be quickly modeling based on the extracted feature with the software"SketchUp".

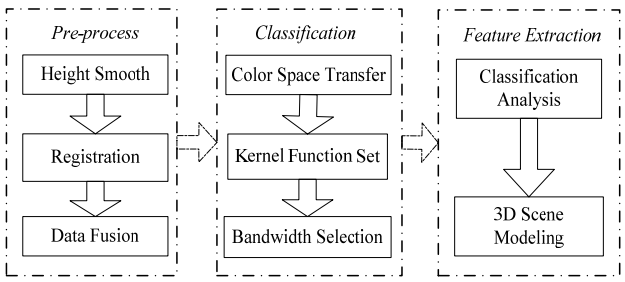


Figure 2. Operation Flowchart

3.1 Mean Shift Principle

Let data be a finite set S embedded in the n-dimensional Euclidean Space, X. Let K be a flat kernel that is the characteristic function of the λ-ball in X.

$$K(x)=\begin{cases}1 & if\ \|x\|\le\lambda\\ 0 & if\ \|x\|>\lambda\end{cases} \tag{1}$$

The sample mean at $x\in X$ is:

$$m(x)=\frac{\sum_{s\in S}K(s-x)s}{\sum_{s\in S}K(s-x)} \tag{2}$$

The difference $m(x)-x$ is called mean shift in Fukunage and Hostetler(Fukunaga et al, 1975). The repeated movement of data points to the sample means is called the mean shift algorithm, as illustrated in figure3. In each iteration of the algorithm, $s\leftarrow m(s)$ is performed for all $s\in S$ simultaneously. The mean shift algorithm has been proposed as a method for cluster analysis, classification, etc.

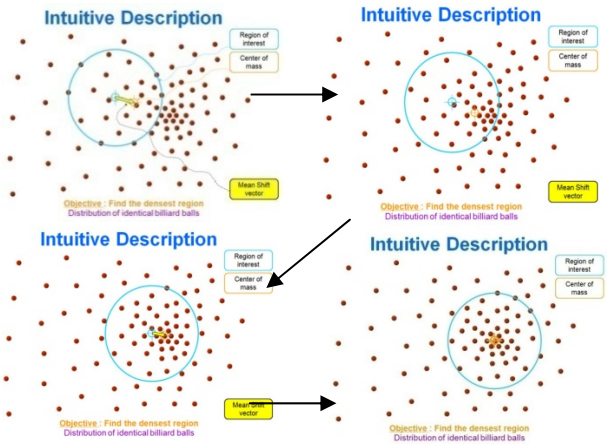


Figure 3. Principle of Mean Shift

3.2 Color Space Transfer Principle

An Euclidean metric, however, is not guaranteed for a color space. The spaces L*u*v* and L*a*b* were especially designed to best approximate perceptually uniform color spaces. In both cases, L*,the lightness(relative brightness) coordinate, is defined the same way, the two spaces differ only through the chromaticity coordinates. In practice, there is no clear advantage between using L*u*v* or L*a*b*; in the proposed algorithms, we employed L*a*b* motivated by a linear mapping property.

A readily accessible conversion formula is as follows: firstly, RGB coordinates should be transferred to CIE-XYZ coordinates using following metric.

$$\begin{bmatrix}(X)\\(Y)\\(Z)\end{bmatrix}=\begin{bmatrix}0.412453 & 0.35758 & 0.180423\\ 0.212671 & 0.71516 & 0.072169\\ 0.019334 & 0.11919 & 0.950227\end{bmatrix}\begin{bmatrix}R\\G\\B\end{bmatrix} \tag{3}$$

The relationship between L*a*b* and CIE-XYZ is illustrated in the following. In the formula, *Xn, Yn, Zn* represent the value of lights stimulus under the CIE standard Lighting conditions. In general, *Xn=95.05, Yn=100, Zn=108.89,* when *X/Xn>0.008856,*

Y/Yn>0.008856, Z/Zn>0.008856.

$$L^* = 116\,(Y/Y_n)^{1/3} - 16$$
$$a^* = 500[(X/X_n)^{1/3} - (Y - Y_n)^{1/3}] \quad (4)$$
$$b^* = 200[(Y/Y_n)^{1/3} - (Z - Z_n)^{1/3}]$$

Otherwise, when *X/Xn<0.008856 Y/Yn<0.008856 or Z/Zn<0.008856*, another expressions represent the relationship.

$$L^* = 903.3[7.787(Y/Y_n)^{1/3} + 16/116]$$
$$a^* = 500[(X/X_n)^{1/3} - (Y - Y_n)^{1/3}] \quad (5)$$
$$b^* = 200[(Y/Y_n)^{1/3} - (Z - Z_n)^{1/3}]$$

3.3 Kernel Function Setting

The kernal function is a weighting function used in nonparametric function estimation. It gives the weights of the nearby data points in making an estimate. In practice kernel functions are piecewise continuous, bounded, symmetric around zero, concave at zero, real valued, and for convenience often integrate to one. They can be probability density functions. Often they have a bounded domain like [-1,1].
Usually, there are some kernel function used widely, as table 1 illustrated.

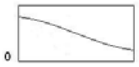
Kernel function	Definition	Profile
Uniform	$K(x) = \begin{cases} 1 & if \ \|\| x \|\| < 1 \\ 0 & if \ \|\| x \|\| \geq 1 \end{cases}$	
Epanechnikov	$K(x) = \begin{cases} 1 - \|\| x \|\|^2 & if \ \|\| x \|\| < 1 \\ 0 & if \ \|\| x \|\| \geq 1 \end{cases}$	
Gaussian	$K(x) = e^{-\|x\|^2}$	

Table 1. Widely used kernel function

In this paper, considering the rate of convergence and the sensitive of feature, Epanechnikov is chosen to be the kernel. Meanwhile, in the fused data*(x, y, z, l, a, b, i), space coordinate (x, y, z), chromaticity coordinate (l,a,b)* and *reflectivity intensity (i)* belong to separately independent Euclidean spaces, therefore, the multi-variate kernel is defined as the product of three radial symmetric kernels and the Euclidean metric allows a single bandwidth parameter for each domain.

$$K(x) = \frac{C}{h_s^3 h_c^3 h_r} k\left(\left\|\frac{x_s}{h_s}\right\|^2\right) k\left(\left\|\frac{x_c}{h_c}\right\|^2\right) k\left(\left|\frac{x_r}{h_r}\right|\right) \quad (6)$$

Xs means *space coordinate (x, y, z), Xc means chromaticity coordinate (l,a,b), Xr means reflectivity intensity (i);* hs, hc, hr means corresponding bandwidth for each domain. C means constant.

4 CASE STUDY

4.1 Pre-processing

4.1.1 Height smooth: Through establishing search window for the point, whether a point is gross error is determined by the comparing the height difference with other points in the search area. The height difference threshold and the search window radius needs to be set, as figure 4 illustrated.

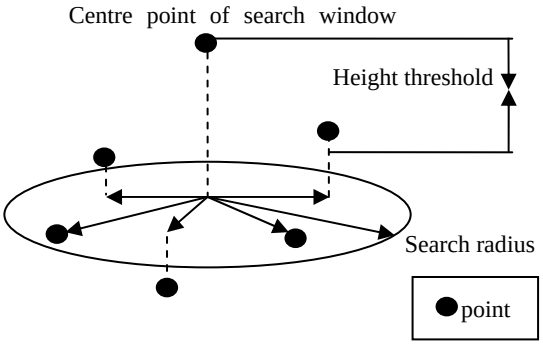


Figure 4. Principle of Height smooth

4.1.2 Registration: Point clouds have three-dimensional position information of high accuracy, while aerial images do not contain coordinates. Before the two types of data can be joined for feature extraction, registering to a single coordinate system should be completed in advance. Fourteen homologous points were selected in the aerial image and the point clouds. The affine transformation based on six parameters of the Plane Coordinate Transformation Model was applied to complete registration and adjustment. The model is given in equation(7).

$$x' = Ax + By + C$$
$$y' = Dx + Ey + F \quad (7)$$

The Least squares solution of six parameters are obtained in table2, through the fourteen homologous points. (x', y') represents the pixel position of the homologous point in the aerial image, while (x, y) represents the point position from lidar data.

Parameter	Value
A	0.6676065
B	0.000883
C	-0.000711
D	-0.669068
E	620313.6
F	4166486.8

Table 2. Coordinate transfer parameter

4.1.3 Data fusion: The spectral information are extracted from the aerial images to the lidar data through overlay operation based on the results of registration. As shown in figure 5, the density of the points is much lower than that of pixels, therefore, the mistake that one point obtains more than one pixel value can not be produced.

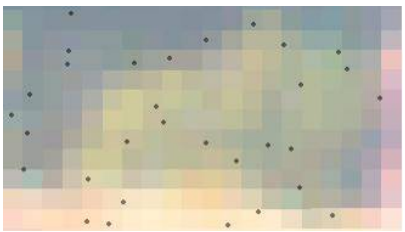
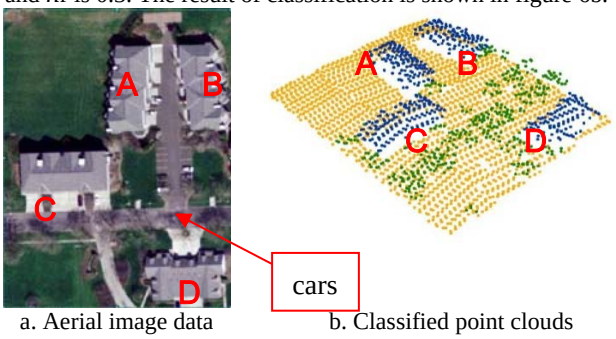


Figure 5. Overlay Details

4.2 Classification

When the location, intensity and spectral vectors are concatenated in the joint domain of dimension, their different nature has to be compensated by proper normalization. Then, by several trials of bandwidth selection, the buildings, the trees

and the land points are well separated, when hs is 0.4, hc is 0.3, and hr is 0.3. The result of classification is shown in figure 6b.



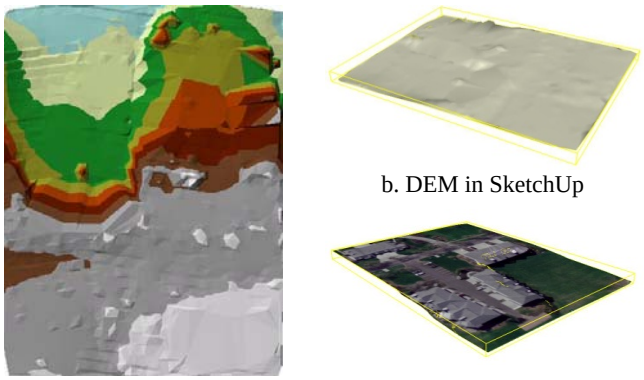
a. Aerial image data b. Classified point clouds

Figure 6. Result of classification

4.3 3D Scene Modelling

With a small amount of manual check of the classification, the misclassification mainly caused by the cars (figure 6a) can be corrected (figure 6b). Then, the accurate classified points representing the ground feature can be the foundation of 3D modelling.

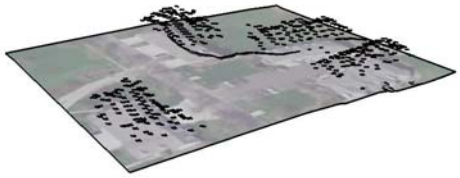
Firstly, the DEM of study area can be extracted from the land point cluster (see figure 7a) in ArcGIS. Then with the ESRI plug-in of SketchUp, the DEM can be imported to the 3D scene of SketchUp (see figure 7b). The aerial images can be the "texture" of the DEM surface (see figure 7c), for the next step of modelling 3D buildings and trees.



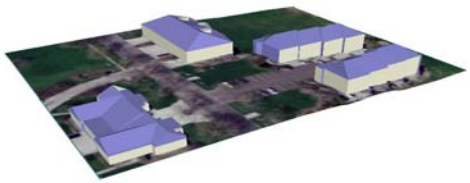
a. DEM in ArcGIS b. DEM in SketchUp c. DEM surface

Figure 7. Procedure of DEM modelling

Secondly, the building point cluster is imported to the SketchUp as a layer(see figure 8a), therefore, the 3D buildings can be modelling, with help of the points and the aerial images information.



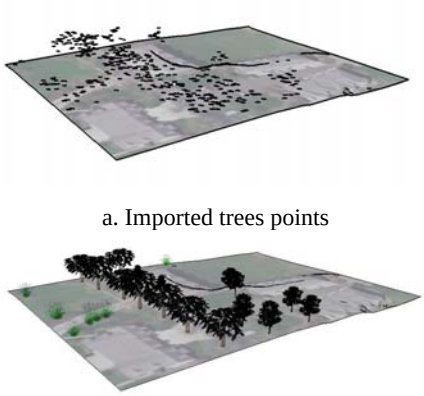
a. Imported building points



b.3D buildings modelling

Figure 8. Procedure of buildings modelling

Similarly, the trees point cluster is used to establish the 3D trees, as figure 9 shown. The single tree model can be found in SketchUp Material Database from Internet.



a. Imported trees points

b. 3D trees modelling

Figure 9. Procedure of trees modelling

5 RESULT AND CONCLUSION

In the end, all the layers are displayed at the same time, and then the ground feature can be reconstructed in figure 10.

Figure 10. Final result of ground feature extraction

On the other hand, the assessment of classification accuracy can be finished by a confusion metric, shown in table 3. The general accuracy of classification can achieve 85.64%, which is a good result of classification.

Number of points		Reference cluster			Total	User accuracy
		building	trees	land		
Practical cluster	building	5582	28	36	5646	98.87%
	trees	288	21524	862	22674	95.06%
	land	1988	5710	26833	34531	77.63%
Total		7858	27262	27731	62708	
Mapping accuracy		71.04%	78.95%	96.86%		
Omission errors		28.96%	21.05%	3.14%		
Commission errors		0.81%	4.22%	28.31%		
General accuracy					85.64%	

Table 3. Confusion metric of classification

This paper proposes a novel method of 3D feature reconstruction, which has been proved feasible in the experiment. Therefore, another thought of efficient feature extraction different from tradition is provided.
The key issues in this paper are: 1) LIDAR data registration with aerial images; 2) classification using mean shift algorithm; 3) reconstruction of ground feature.

ACKNOWLEDGEMENT

This paper is fully supported by the National Natural Science Foundation of China (Project No.40501061), National 863 High-tech Project “Complicated Features’ Extraction and Analysis from Central Districts in Metropolis”(Project ID: 2007AA12Z178) and Key Laboratory of Advanced Engineering Surveying of SBSM (Project ID: ES-SBSM(07)-07)

REFERENCES

Comaniciu, D. and P. Meer (2002). "Mean shift: A robust approach toward feature space analysis." *IEEE Transactions on pattern analysis and machine intelligence* 24(5): 603-619.

Dalponte, M., L. Bruzzone, et al. (2008). "Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas." *IEEE Transactions on Geoscience and Remote Sensing* 46(5): 1416-27.

Debevec, P., C. Taylor, et al. (1996). Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach, ACM New York, NY, USA.

DeCandido, R. (2004). "Recent changes in plant species diversity in urban Pelham Bay Park, 1947–1998." *Biological Conservation* 120(1): 129-136.

Haala, N. and C. Brenner (1997). *Generation of 3D city models from airborne laser scanning data*, Citeseer.

Herold, M., J. Scepan, et al. (2002). Object-oriented mapping and analysis of urban land use/cover using IKONOS data.

Irish, J. L. and W. J. Lillycrop (1999). "Scanning laser mapping of the coastal zone: the SHOALS system." *ISPRS Journal of Photogrammetry and Remote Sensing* 54(2-3): 123-129.

Lefsky, M. A., W. B. Cohen, et al. (1999). "Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests." *Remote Sensing of Environment* 70(3): 339-361.

Maas, H. and G. Vosselman (1999). "Two algorithms for extracting building models from raw laser altimetry data." *ISPRS Journal of Photogrammetry and Remote Sensing* 54(2-3): 153-163.

Mei, Z., X. Bing, et al. (2009). A Classification Method for Building Detection Based on LiDAR Point Clouds. 2009 Urban Remote Sensing Joint Event, Shang Hai.

Nilsson, M. (1996). "Estimation of tree weights and stand volume using an airborne lidar system." *Remote Sensing of Environment* 56(1): 1-7.

Noronha, S. and R. Nevatia (2001). "Detection and modeling of buildings from multiple aerial images." *IEEE Transactions on pattern analysis and machine intelligence* 23(5): 501-518.

Rottensteiner, F. and C. Briese (2002). "A new method for building extraction in urban areas from high-resolution LIDAR data." *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences* 34(3/A): 295-301.

Rottensteiner, F., J. Trinder, et al. (2005). "Using the Dempster-Shafer mehtod for the fusion of LIDAR data and Multi-spectral images for building detection." *Information Fusion* 6(4): 283-300.

Rutzinger, M., B. Hofle, et al. (2006). Object-based building detection based on airborne laser scanning data within GRASS GIS environment.

Sohn, G. and I. Dowman (2007). "Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction." *ISPRS Journal of Photogrammetry and Remote Sensing* 62(1): 43-63.

Tao, G. and Y. Yasuoka (2002). "Combining high resolution satellite imagery and airborne laser scanning data for generating bareland DEM in urban areas." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 30.

Wulder, M., J. White, et al. (2009). "Characterizing boreal forest wildfire with multi-temporal Landsat and LIDAR data." *Remote Sensing of Environment* 113(7): 1540-1555.