Research on Building Extraction Based on High-Resolution DSM Images from Oblique Photography

Shuai Xianghua[1], Li Weiwei[2], Feng Wei[3], Yang Haifang[3], Zhen Meng[1]

1 China Earthquake Networks Center，Beijing 100045，China

2Henan Province Non-ferrous Metals Geological Mineral Resources Bureau，Henan，450016，China

3 Institute of Earthquake Forecasting, China Earthquake Administration，Beijing 100036，China

shuaixhua@sina.com

keywords: oblique photography, DSM, depth image, gray threshold, area threshold, building extraction

Abstract. Digital Surface Model (DSM) can fully exhibit the undulate features of terrain and landmark, which has obvious advantages in buildings extraction rapidly. The experiments about oblique photography measurements during the M6.5 Ludian Earthquake occurred in Longquan village Longtoushan town on August 3, 2014, in this paper, have been researched. The data are acquired by using the 5-lens oblique photography collection techniques carried by electric Six-rotor Unmanned Aerial Vehicle (UAV), and generating the highly resolution DSM images according to the system model with precision of 4.9cm and elevation of 1494.27-1638.69m. We try to convert the DSM images into the DSM depth images that contain Buildings, ground and trees information. Because in undulate terrain of the mountainous research area, Buildings are constructed according to terrain and all in a muddle. The buildings manifest different floors and coexistence between low buildings and multi-storey buildings, observing from different angles, a serial of buildings form larger architecture covered with dense woods. The threshold of gray values of Buildings and background information displays diversity in the study area, therefore, the method of combining multi-level gray threshold and area threshold is used to extract Buildings for researching in this paper. To obtain the gray values for study area, we transform DSM depth image into binary image by gray histogram method of threshold segmentation, minimum of result is ground, maximum is tree, and Buildings show lower gray values. The gray segmentation threshold of Buildings and background information concentrates on 30 and 45, in this paper, so we select the images of 26, 30, 36, 45 thresholds to binary. The images are combined by threshold of 26 and 36, 26 and 45, 30 and 45, 36 and 45, and we confirm 26 and 45 combination is final result according to actual surface contrast. However, the extracted results also include ground and trees. Buildings, trees and ground show regular otherness in size, so we set up different area thresholds to analysis and extraction for the binary images. At last, we evaluate accuracy by completeness and correctness of extracted Buildings, the result of the assessment is satisfactory. It follows that rapid extracting Buildings by oblique photography DSM images is applicable for not only urban regions but also mountainous areas, which is of great significance in quick acquisition for Building measurements pre- and post-earthquake.

Introduction

DSM is surface elevation model that contains buildings, bridges, trees and so on. Based on DEM, DSM can ample express the undulate features of terrain and landmark because it re-covers other surface elevation except ground. At present, DSM data are generated mainly by means of regular gridding of Light Detection And Ranging (LIDAR) data and photogrammetric stereo matching, both of the DSM measurements are gradual and deep studied in detecting buildings. Utilizing the LIDAR data, evaluation grading of buildings was realized according to setting elevation and area threshold (Huang et al., 2015). Using the DSM measurements of airborne three- dimensional imager to divide threshold, the precision of buildings extraction reached 90% (Long et al., 2001). Based on the

classified extraction method of ground object spectra and texture property, DSM and DEM extracted from aerial images had been used successfully for obtaining the building profile and 3D reconstruction (Kan and Cao, 2014). 3D variation monitoring of buildings was achieved from combining the LIDAR point cloud data and aerial images (Peng et al., 2015). The DSM data had been analyzed and compared their advantages and disadvantages, which were acquired from aerial photogrammetry and 3D laser scanning (Mu and Li, 2008). LIDAR technology can obtain high precision DSM measurements efficiently, but the restriction of point interval, scanning angle, especially lack of spectral information, objective color and texture information will be ignored. Therefore, the extracted 3D model is imprecise from LIDAR data, it is always used to detect buildings outline combined with aerial images. Traditional photogrammetry utilizes digital software, based on the DSM data obtained by stereo matching, relies on image processing, feature matching and especially side texture of buildings that needs manual collection for the complex procedure, which is difficult to suit the current demands because of the poor efficiency and precision (Wu et al., 2015).

With the development of photogrammetry and remote sensing technology during the 1990s, airborne multi-angle oblique photography technology emerges as the times require as combining of traditional aerial photogrammetry and digital surface acquisition technology. Oblique photogrammetry technology loads multiple or various sensors in the same flight platform, which collects ground images from multiple angles to acquire high resolution texture information of abundant buildings top and side (Wang et al., 2011). Oblique photography system has the advantage of low cost and extracting detailed building textures from multiple angles, that solves the deficiencies of collecting building lateral images in traditional aerial photography. So oblique photography system plays an increasingly important role in aerial remote sensing, because it is able to obtain the sub-meter precision DSM measurements by aerotriangulation and dense matching (Sun, 2014). DSM data contain non-ground objects such as buildings and trees, but the elevation information is significant to extract buildings. Based on analyzing for DSM data obtained from photography, which is of great importance for buildings formulation and change detection (Wang et al., 2014; Li et al., 2012).

The experiments about oblique photography measurements during the M6.5 Ludian Earthquake occurred in Longquan village Longtoushan town on August 3, 2014, in this paper, had been researched. Because in undulate terrain of the study area, buildings are constructed according to terrain and all in a muddle. The buildings manifest different floors and coexistence between low buildings and multi-storey buildings, observing from different angles, a serial of buildings form larger architecture covered with dense woods that display area difference in the images. We extract DSM measurements from oblique photography data, therefore, the method of combining multi-level gray thresholds and area thresholds is used to extract buildings for research in this paper accordingly.

Method

We convert the DSM images into the DSM depth images, and then digitalize them to detect buildings. As depth images contain buildings, ground, and trees, in order to acquire interested building area, on the one hand we intuitive analyze images by gray histogram, on the other hand we set up gray thresholds to divide buildings and background area. Gray thresholds divide into binary images that contain buildings, trees and mountains, buildings have an area regularity, so we can extract buildings by setting area thresholds.

DSM depth image

DSM depth images are obtained by gray-scale quantifying on elevation from DEM data, which are similar to gray images. DSM depth image is a special gray image, so traditional digital image processing technology is also applied to it. Knowing the elevation range of DSM image, we normalize the elevation values of DSM image into depth image with gray range from 0 to 255 according to Eq. (2-1) (Zhang et al., 2014),

$$\text{Pixel}(i,j) = \frac{Z(i,j) - Z_{min}}{Z_{max} - Z_{min}} \times 255$$

2-1

where Pixel(i,j) is gray value of row of i and column j in DSM image, Z(i,j) is elevation value of row i and column j, Zmin is minimum elevation value, while Zmax is maximum elevation value. According to Eq. (2-1), the object (like ground) with lower elevation displays lower gray value in depth image, while the object (like building, bridge, tree etc) with higher elevation displays higher gray value.

Threshold segmentation method

Image segmentation is very important in image processing and computer vision field, which is basic premise for visual analysis and pattern recognition. Research on image segmentation algorithm can be traced back to 1960s, generally sorted into three kinds, that is segmentation by threshold, edge and area correspondingly. Threshold segmentation is widely used in digital image processing, because it has many advantages such as easy realization, less calculation and steady performance etc. Segmentation method based on threshold has many kinds, for example histogram method, global threshold method (like Otsu algorithm), local threshold method (like Bersen algorithm) and so on (Pan, 2012). In our research, we utilize gray histogram to divide threshold. Gray histogram is gray level function for gray image, it represents the number of pixels for each gray level, which expresses occurrence frequency of each gray in the image. We obtain the binary image by histogram statistics of gray image, and dividing buildings and background by setting up threshold. Binary equation can be written as Eq. (2-2),

$$g(i, j) = \begin{cases} 255 & f(i, j) \geq T \\ 0 & f(i, j) < T \end{cases}$$ 2-2

with g(i, j) the pixel value of the binary image at row i and column j, with f(i, j) the pixel value of the input image at row i and column j, and with T threshold. Image binarization aims to generate two kinds of gray level by appropriate thresholds from gray image with a range of 0 to 255, they are black-and-white images with 0 value and 255 value, 0 represents background image and 255 represents target image.

Study and analysis

Experimental data

The data are oblique photographic images acquired by electric multiple-rotor UAV system that contain one vertical camera and four oblique cameras, and they highly integrated with GPS receiver and high-precision IMU. The five angle cameras provide oblique and vertical images, GPS obtains real-time location, and IMU offers state information. Based on internal parameters of system camera and Position and Orientation System (POS) data, firstly we calculate precise exterior elements of image by aerial triangulation technology of oblique photography. Secondly selecting suitable image matching units for feature and dense matching, we utilize the exterior orientation elements to obtain DSM images in different directions by dense matching of multi-angle images. In this paper, DSM data generated by oblique photograph technology have the precision of 4.9cm and elevation range of 1494.27m to 1638.69m. Oblique image of study area is displayed in Fig. 1.

Fig. 1. 3D image for oblique angle in Longquan village

Transform into deep DSM image and gray threshold to extract buildings

In Eq. (2-1), DSM image is transformed into depth image with the gray range of 0 to 255, that is showed in Fig. 2. There is a significant difference between the gray value of the target building and the gray value of ground and tree in DSM depth image, maximum gray values of target object represent trees and minimum gray values of target object are ground. Threshold choice is most important in segmentation of DSM depth image. Gray statistics of DSM depth image is obtained by gray histogram in study area, shown in Fig. 3, abscissa shows gray values of DSM image, and ordinate represents the number of pixels for these gray values. Through the histogram analysis of DSM and gray images, we can see that buildings have lower gray values. We select the 26, 30, 36, 45 thresholds to binary according to gray histogram, the result is shown in Fig. 4.
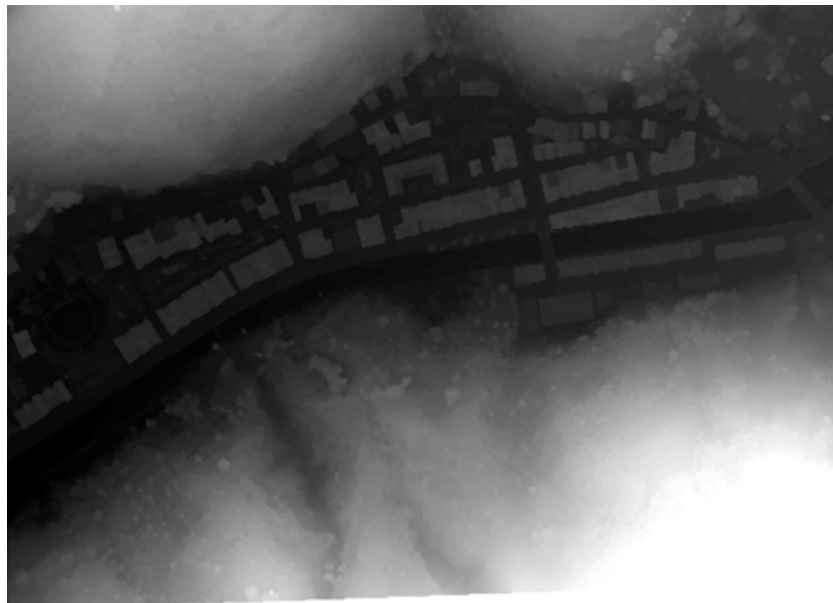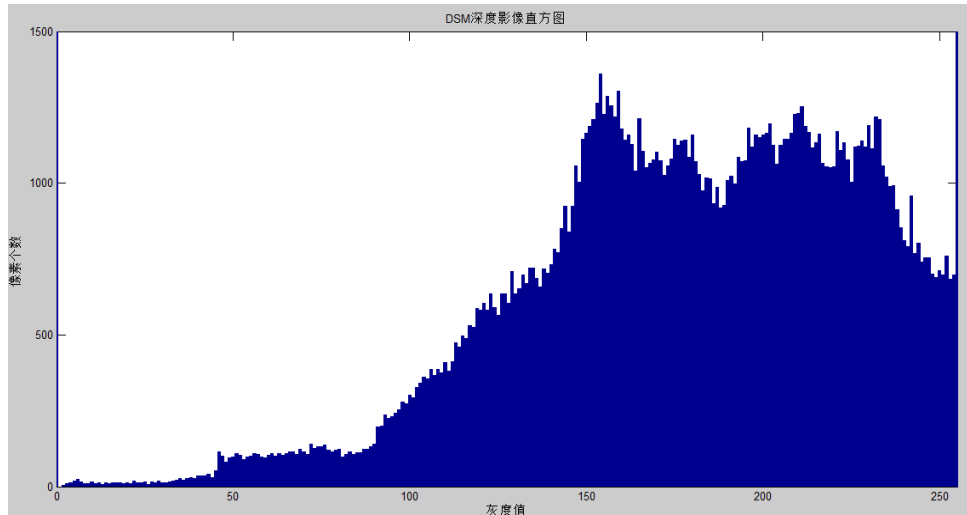

Fig. 2 DSM depth image

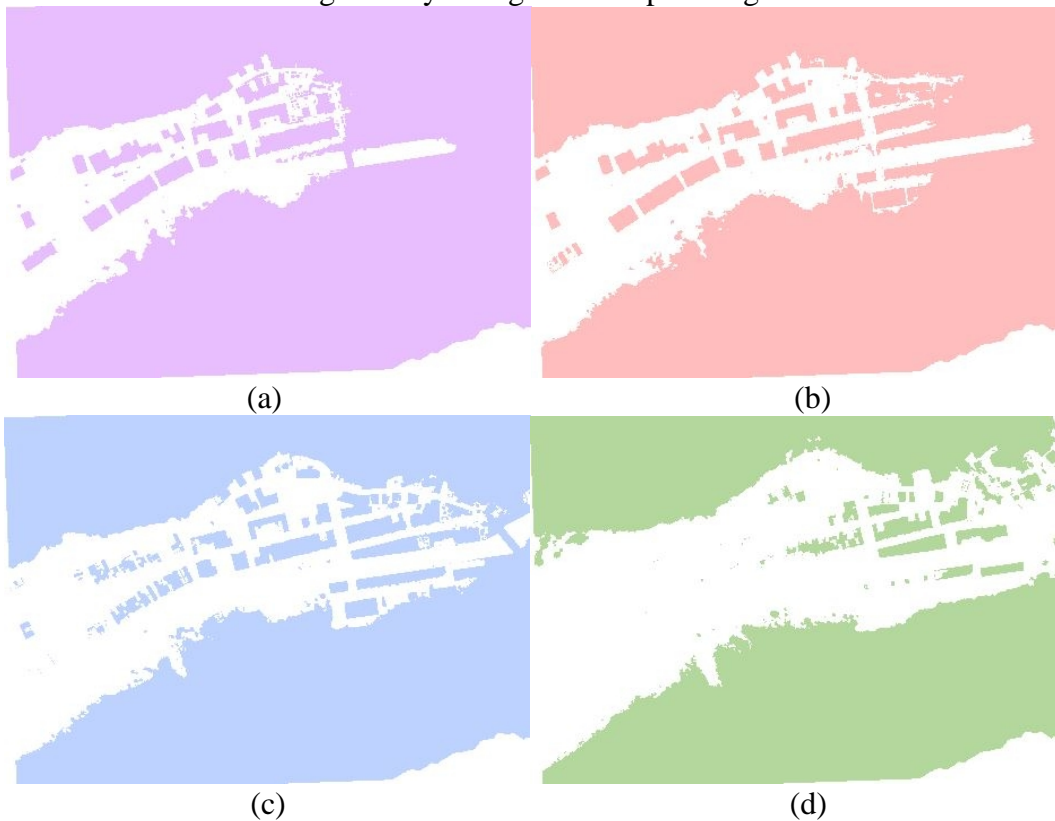Fig. 3 Gray histogram of depth image

(a)

(b)

(c)

(d)

Fig. 4 (a) gray threshold of 26;(b) gray threshold of 30;(c) gray threshold of 36;(d) gray threshold of 45

The result of threshold segmentation indicates that the gray values of the background and the segmentation threshold of the buildings of one part are about 30, and the other parts of the values are about 45. Extracting buildings from background needs to set up multiple thresholds for study area. According to threshold segmentation, the thresholds are combined by 26 and 36, 26 and 45, 30 and 45, 36 and 45, the results are displayed in Fig. 5. By contrast to the actual image, we choose 26 and 45 combination is final result for segmentation of gray threshold.

<center>(a)   (b)</center>
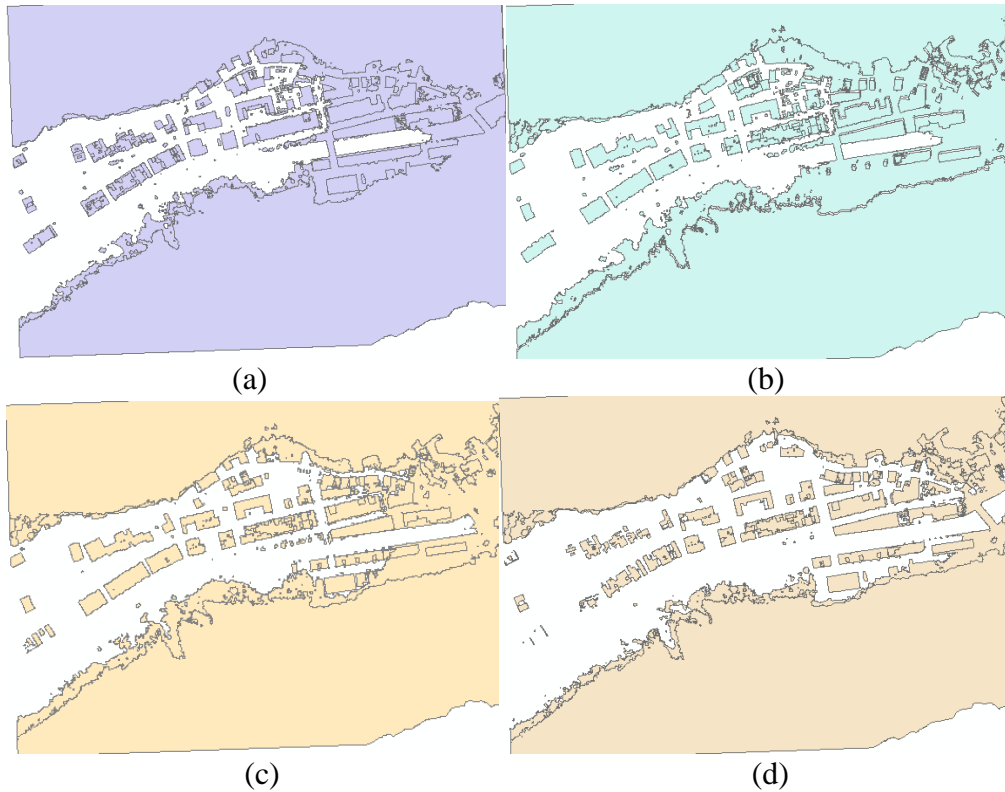
<center>(c)   (d)</center>

Fig. 5 gray threshold of 26 and 36; (b) gray threshold of 26 and 45; (c) gray threshold of 30 and 45; (d) gray threshold of 36 and 45

Extract buildings by area threshold

The images contain not only buildings, but also elevation of trees and mountains after gray threshold segmentation, and buildings have certain size compared with trees and mountains, so we can extract buildings by area thresholds. Firstly we obtain the profile data of buildings based on vectorization of DSM images, in Fig. 6, the pink regions are vectorized buildings. Secondly the area rule of buildings is analyzed after gray threshold segmentation, the result is plotted when area thresholds are more than 100m2 and less than 3000m2 in Fig. 8. Fig. 7 and 8 show that the pink surfaces are buildings data based on Vectorization of DSM image, and the black edges are extracted result that contains building profiles based on area threshold segmentation.
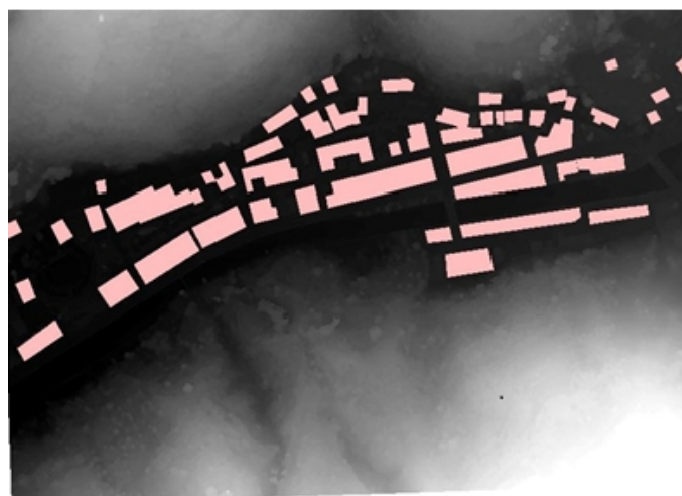

<center>Fig. 6 Vectorized surface data of buildings based on DSM image superposition</center>

Fig. 7 Comparison of extracting result with visual interpretation and area threshold greater than 70

Fig. 8 Comparison of extracted result with visual interpretation and area threshold greater than 100
Accuracy assessment

In our research, we use both of completeness and correctness to assess extracted precision for buildings. Completeness means the proportion of the total number of correct buildings extracted by threshold segmentation and the total number of buildings extracted by vectorization of DSM images, which expresses the integrity of the buildings extraction. Correctness represents the proportion of the total number of correct buildings and the whole buildings extracted by threshold segmentation, which reflects the accuracy of the buildings extraction. The formulas of completeness and correctness are written as Eq. (3-1) and (3-2),

$$Completeness = \frac{TP}{TP+FN} \qquad \text{3-1}$$

$$Correctness = \frac{TP}{TP+FP} \qquad \text{3-2}$$

where TP (True Position) represents the total number of correct extracted buildings, FP (False Positive) represents the number of wrong extractions, FN (False Negative) means the total number of failed extractions that is equivalent to the subtraction of the number of vectorization buildings and the total number of correct buildings.

In the experimental area, assessment results of completeness and correctness of buildings data are displayed in Table 1. There are 52 buildings in the experimental area, the completeness of buildings is 73% if we set up area thresholds larger than 70 and less than 3000. The failed extractions are mainly buildings that are low and built close to mountains, these buildings are not well separated from the

background values, which leads to large error. The correctness of buildings is 88% accordingly, trees with coincidental area threshold are misclassified that causes error. The completeness of buildings is 88%, when we set up area thresholds larger than 100 and less than 3000, the correctness of buildings is 97%. Correctness has been manifestly improved according to increase area threshold, but the completeness has ever reduced.

Table 1 Assessment results of extracting precision of buildings

	70＜area threshold＜3000	100＜area threshold＜3000
TP (True Position)	38	37
FP (False Positive)	5	1
completeness	73.08%	71.15%
correctness	88.37%	97.36%

Conclusions

Compared with the regular gridding of LIDAR data and DSM data generated by traditional photogrammetric stereo matching, DSM data obtained by oblique photograph technology have the advantage of rapid imagery and high precision. In this paper, we transform DSM images obtained by oblique photograph technology into depth images, which are studied by segmentation of gray and area thresholds. According to assessment of completeness and correctness for buildings, we build a extracting method based on DSM data of oblique photography finally. This method is used in undulate mountainous regions where buildings are constructed close to mountains, during experimental analysis, low and built close to mountains buildings that are similar to gray values of background, but we can separate them with a better completeness by setting up thresholds. According to set up two area thresholds for extracting buildings, the completeness is lower and correctness is higher when we set up area threshold larger than 100, while the completeness increases and correctness decreases correspondingly when we set up area threshold larger than 70. We need to perfect the segmentation and extraction method for background and buildings that are low and built close to mountains in the future, in order to further improve the completeness and correctness of buildings.

Acknowledgements

Reference

[1] M. Huang, J. Zhang and B. Zeng: The classification of Building Height in a City Based on Airborne LiDAR Data, Bulletin of Surveying and Mapping, 4, 82-85 (2015)

[2] H. You, T. Liu, L. Su, S. Liu, G. Guo and S. Li.: Fast Acquiring Urban DSM Image and Displaying 3D Image, Journal of Remote Sensing, 5(1), 8-12 (2001)

[3] P. Kan and L. Cao: Building Extraction Using High Resolution Imagery, Remote Sensing Information, 29(1), 3-6 (2014)

[4] D. Peng, Y. Zhang and X. Xiong: 3D Building Change Detection by Combining LIDAR Point Clouds and Aerial Imagery, Geomatics and Information Science of Wuahn University, 40(4), 462-468 (2015)

[5] L. Mu and F. Li: On the generation technology of city digital surface model and application, Journal of Shaanxi Normal University (Natural Science Edition), 36, 117-131 (2008)

[6] J. Wu, M. Cheng, Z. Yao, Z. Peng, J. Li and J. Ma: Automatic generation of high-quality urban DSM with airborne oblique images, Journal of Image and Graphics, 20(6), 0845-0856 (2015)

[7] W. Wang, W. Huang and J. Zhen,: Pictometry Oblique Photography Technique and its Application in 3D City Modeling, Geomatics and Spatial Information Technology, 34(3), 181-183 (2011)

[8] H. Sun: Three Dimensions Modeling of Digital City Based on Oblique Photogrammetry Method, Modern Surveying and Mapping, 37(1), 18-21 (2014)

[9] Q. Wang, H. Ai and L. Zhang: Rapid city modeling based on oblique photography and 3ds Max technique, Science of Surveying and Mapping, 39(6), 74-78 (2014)

[10] S. Li, Y. Song and C. Li: Quick Extraction of Building Based on DSM and RGB-Aerial Image, Bulletin of Surveying and Mapping, 1, 33-36 (2012)

[11] W. Zhang, L. Yang, L. Cao and Y. Jia: Method Research for Constructing Urban DSM Based on Aerial Image Matching Point Cloud, Remote Sensing Information, 29(6), 31-35 (2014)

[12] Z. Pan: Building Extraction Based on the LIDAR Point Cloud Data and Image Fusion[D], China University of Geosciences, Beijing (2012)

[13] L. Chen: Research on Building Detection in Urban Area Based on Aerial LiDAR Data and High Resolution Remotely Sensed Imagery, Nanjing University, Jiang Su (2011)