BUILDING EXTRACTION USING SURFACE MODEL CLASSIFICATION

Hossein, AREFI[1], Amin, ALIZADEH[2*], Ali, GHAFOURI[2*]

[1]Remote Sensing Technology Institute, German Aerospace Center (DLR), Wessling, Germany

hossein.arefi@dlr.de

[2] Dept. of Geomatics, Tehran University, Tehran, Iran

a.alizadeh@ut.ac.ir, ali.ghafouri@ut.ac.ir

Abstract

In many applications like urban planning and environmental simulation, the major solution is building extraction, which can be performed using different airborne or space-borne data or an appropriate fusion of them. This paper presents an automatic building recognition technique using fusion of LIDAR data and multispectral imagery. To this end, a rule-based classification method is considered in order to extract buildings from input data which are DSM, DTM extracted from DSM and an optical Image. To achieve better accuracy classification is performed in both pixel and object level. Accordingly, a user-friendly MATLAB toolbox is provided for both classification and evaluation procedures. It is experimentally shown that the proposed algorithm can successfully detect urban residential buildings, when assessed in terms of different quantitative criteria and visual inspection.

Keywords: Building Extraction, Digital Surface Model, Rule-based Classification

INTRODUCTION

Building detection from remotely sensed data is important to the real estate industry, city planning, home-land security, disaster (flood or bush fire) management and many other applications. The Automated extraction of building boundaries towards generating city models is also an essential step (Cheng et al., 2008). As it can be observed, over the last few decades, a large number of building detection techniques have been reported (Awrangjeb et al., 2010). However, a fully successful automatic building detection is still an ambitious goal. There are several reasons to explain the obstacles in this way including (Awrangjeb et al., 2010, Sohn and Dowman, 2007):

- *Sensor dependency:* the primary data to support the building detection is available from a variety of sources with different resolutions, each source having its own bright and dark points for building detection.
- *Scene complexity:* most of the scenes usually contain very rich information which provide a large number of prompts with geometric or chromatic co-similarity to buildings, but belong to non-building objects.
- *Incomplete cue extraction:* due to occlusions, poor contrast, shadows and disadvantageous image perspective, there is always a significant loss of relevant building cues.

According to literatures, three main categories can be considered for building detection approaches (Lee et al., 2008). First of all, there are many algorithms that employ 2D or 3D information from photogrammetric imagery (Mayer, 1999). The complexity of separating buildings from other objects Increases with increasing image resolution. This is because high-resolution images contain more detailed information (Cheng et al., 2008) and doubtlessly occlusions and shadows (Li and Wu, 2008). Furthermore by using stereo images or multiple images captured from the same scene it is possible to derive 3d information. It should be noted that 3D information is one of the most important feature in building detection. (Sun et al., 2005). nonetheless, nearby trees of similar height also make the use of such derived range data difficult (Lee et al., 2008). Certainly, the problem meaning trees next to building can be resolved by using some texture analyses or

---

* Corresponding authors

NDVI index (Meyong et al., 2001). Nevertheless, height information extracted from a stereo pair of images is much more weak particularly in urban areas because of mismatching in these areas (Grigillo et al., 2012).

As the second group, there have been several attempts to detect building regions from LIDAR (LIght Detection And Ranging) data. This task has been largely solved by classifying the LIDAR points according to whether they belong to bare-earth, buildings, or other object classes (Lee et al., 2008). In fact, the introduction of LIDAR has offered a favourable option for improving the level of automation in building detection process when compared to image-based detection (Vu et al., 2009). A number of problems with building detection have been discussed in the literature (elberink et al., 2011), and they have been shown that the use of raw or interpolated data can influence the detection performance (Awrangjeb et al., 2010). Moreover, building detection with very good horizontal accuracy and appropriate discrimination between trees and buildings is almost impossible (Haala and Brenner, 1999).

Each of LIDAR and photogrammetric imagery has particular advantages and disadvantages in horizontal and vertical positioning accuracy. in comparison with photogrammetric imagery, LIDAR generally provides more accurate height information but less accurate boundary lines. On the other side Photogrammetric imagery can provide extensive 2D information such as high-resolution texture, and different indices like NDVI index.

The third category of methods utilizes both LIDAR data and photogrammetric imagery. More explicitly, intensity and height information in LIDAR data can be used with texture and region boundary information in aerial imagery to improve accuracy and correctness (Lee et al., 2008).

It is necessary to mention that LIDAR data and aerial photography have the same quality as high-resolution DSM and satellite images such as orthorectified worldview respectively. Consequently, all algorithms aforementioned for the former group can be used for the latter group. This point has been mentioned because our data in this study are high-resolution DSM and orthorectified worldview.

Although there are a lot of works having been done about building detection, a few works have been done by integration of LIDAR and photogrammetric imagery.

Building detection techniques integrating LIDAR data and imagery can be separated into two groups. Firstly, there are techniques which use the LIDAR data as the primary cue for building detection and employ the imagery only to eliminate vegetation from the scene (Vu et al. 2009 and Rottensteiner, 2005). As a result, they suffer from poor horizontal accuracy for the detected buildings. Since problem of horizontal accuracy can be somewhat resolved by different outline approximation (Arefi et al., 2008), it can be ignored in building detection part. In another work, Dempster-Shafer theory as a data fusion framework was used to classify points as building, tree, grassland or bare soil. A method based on morphological scale space is proposed for extracting building foot prints from the elevation data and then removing vegetation pixels using the spectral data (Vu et al., 2009). The proposed building detection technique falls into this group.

Secondly, there are integration techniques including (Awrangjeb et al., 2010, Sohn and Dowman, 2007, Chen et al., 2004) which use both the LIDAR data and the imagery as the primary cues to delineate building outlines. Consequently, they offer better horizontal accuracy for the detected buildings. Haala and Brenner (1999) applied a pixel-based classification where the normalized DSM (nDSM) was used as additional channel to the three spectral bands of the aerial imagery (Haala and Brenner, 1999). Chen et. al. followed a region based segmentation of nDSM and ortho-images and then used a knowledge-based classification to detect building (2004). Sohn and Dowman proposed a data-driven approach on the optical imagery and a model-driven approach on the point cloud to extract rectilinear lines around buildings (Sohn and Dowman, 2007). In another work, a similar technique was proposed with precise geometric position. Lee et. al. extracted the initial building boundaries from the LIDAR data and then enhanced the initial boundaries using colour information ,after which edge matching and perceptual grouping techniques were applied to yield the final building boundaries (Lee et al., 2008). Demir et. al. applied four different methods to achieve an improvement by combining the advantages and disadvantages of these approaches and used the edge information from images for quality improvement of the detected buildings (Demir et al., 2009).

This paper aims at following two goals: a successful integration of the LIDAR data and photogrammetric imagery for building detection so that LIDAR performance in tree removal is improved ,and development of a rule-based classification performed in both pixel and object level. In other words, the goal of this study is building extraction based on a rule-based classification in pixel and object level using fusion of multi-source data.

RULE BASED CLASSIFICATION

As mentioned above, a rule-based classification method is used for building extraction. The first stage of this classification is feature selection. After feature extraction, classification is done based on a set of rules written by selected features. An outlook of rule-based classification is depicted in Fig. 1.

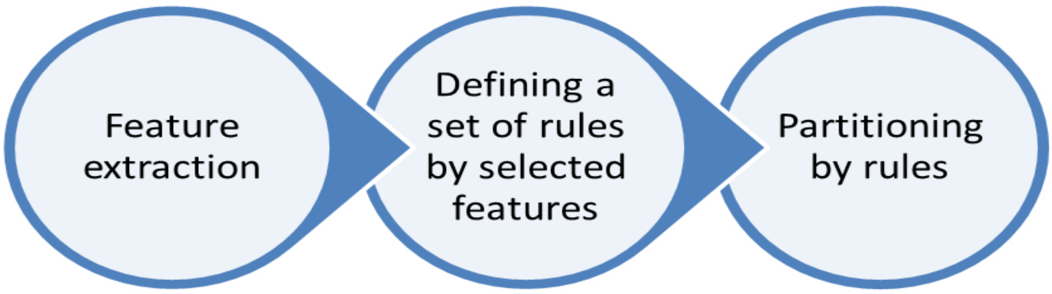

Fig. 1. General procedure of rule-based classification

In the following, at first different features used in this study is explained. After that rule-based classification process is discussed.

FEATURES

In this study, two kinds of features named pixel-based and feature-based are used. In the following each of them has been explained in more detail.

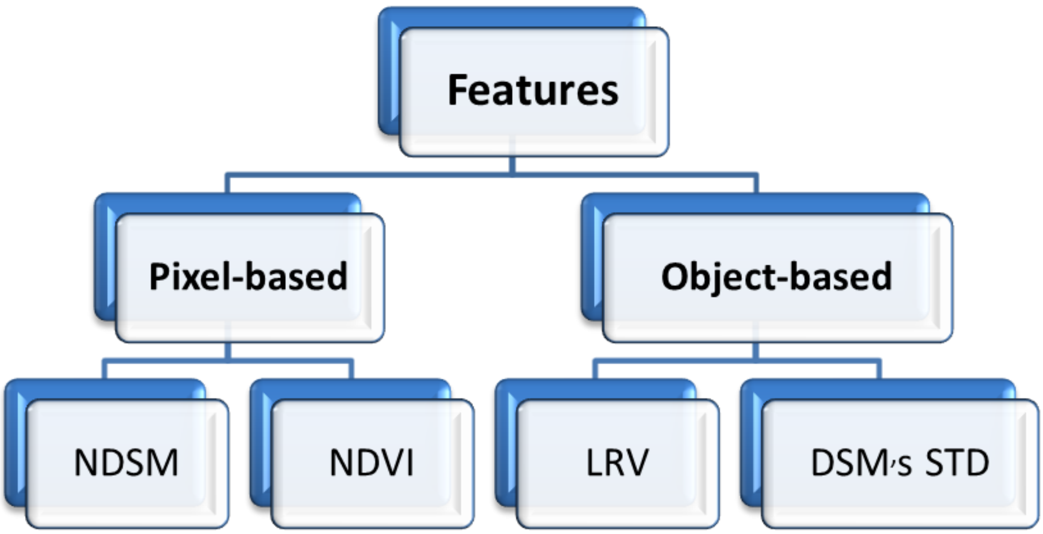

Fig. 2. an outlook of used features

Pixel-based features

NDSM. This feature, so-called normalized digital surface model (nDSM), is computed by measuring the difference between the DSM and the digital terrain model (DTM). Since the nDSM excludes the influence of topography, it represents the height of all overlying objects, such as buildings and trees on the terrain. Therefore, expert knowledge about the appearance of certain objects can be directly used. In many studies, the nDSM was used as an important clue for building detection. That is noticeable that DTM is extracted from DSM by morphological reconstruction algorithm (Arefi and Hahn, 2005).

NDVI. As it is indicated in equation 1, this feature can be computed by near-infrared and red bands. Since in this project the world-view 2 images with eight spectral bands are experimented, this feature can be easily extracted (Krauss et al., 2008).

$$NDVI = (NIR – Red) / (NIR + Red), \qquad (1)$$

Using this feature, vegetation and non-vegetation can be separated from each other.

Object-based features

For these kinds of features, first of all regions or objects have to be extracted. In this regard, there are two different approaches. One of them is segmentation and the other is using regions extracted from a processing. In this project, we used the latter.

Local Range Variation (LRV): As early mentioned, for this feature, regions have to be extracted. For each object on the DSM, borders are extracted. Every border consists of a number of points. Around each point on the border, a 3×3 window has been considered. Then for every window difference between max and min value are computed. This work is done for every point on the border. Consequently, for each border a set of LRV values are computed. In this study the mean of these LR values belonging to each region has been considered as a feature value. The objects such as small hills and buildings can be distinguished from each other using this feature descriptor (Hahn et al., 2007).

STD: As early mentioned, for this feature, regions have to be extracted. This feature is the standard deviation of height values for each region. By this feature, leaf-off trees can be distinguished.

CLASSIFICATION PROCESS

After selecting appropriate features, classification is triggered. Usually in rule-based classification at first based on the most important feature (completely depends on application) an initial classification is done. After that, using other features this classification is improved. In this case since the aim is building extraction, nDSM is the most important feature. Consequently, initial classification is done in terms of nDSM (level0). Then, this initialization will be improved by the other features. This initial classification is includes in leaf-on and off trees, mounts and buildings. Since our goal is building extraction, non-building features should be eliminated in the improvement levels (we considered 3 levels in this study). At first, leaf-on trees and height vegetation are excluded from nDSM (level 1). By this work leaf-on trees are removed. After that, mount and other objects like that are excluded from improved nDSM in the first level. By this work mounts and hills are removed (level 2). Finally leaf-off trees and other objects like that are excluded from the improved nDSM in the second level (Nozaki et al., 1996).

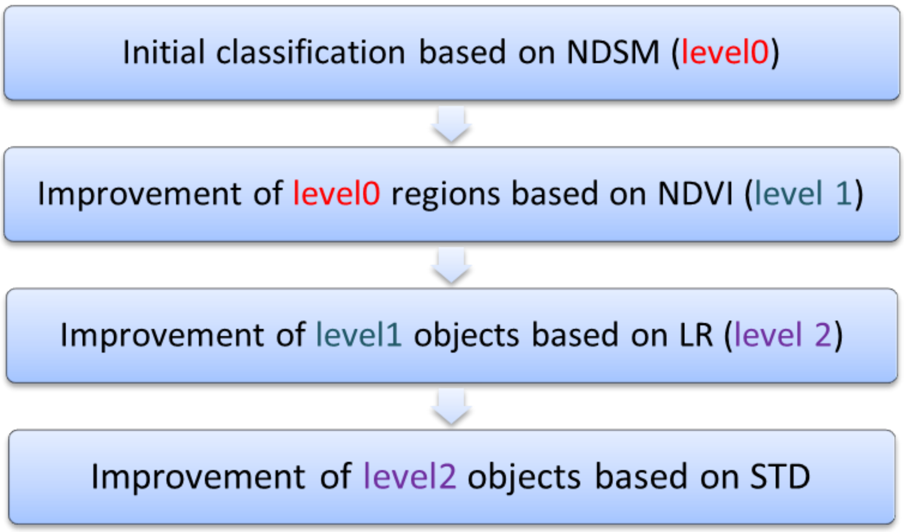

Fig.3. an outlook of classification process

IMPLEMENTATION

In order to building detection a program is developed in MATLAB which can provide rule based classification of DSM considering DTM and Satellite image as well. Generally nDSM is produced by subtraction of DTM from DSM in order to remove natural topographic.

PROGRAMMING AND DEVELOPMENT

Fig. 4 depicts the interface of the developed program. After selection of the type of input data which could be MAT or Image file, by clicking on "Open Image", the input Imagery data would be selectable. In the next step "Digital Terrain Model" is added like an input to the program (Cheng et al., 2008). "DSM" as the second input is considered as the third step. After this data import, normalized DSM can be calculated by subtracting DSM from DTM.

By clicking on the forth button of data input process, Ground Truth map which has been prepared by digitizing the Imagery Data is imported and would be ready to use. Ground Truth preparation is fulfilled in the ArcGIS software as a vectorized output. The vector ground truth had been converted into raster with the same resolution as the optical image and Digital Models (DSM/DTM and nDSM).

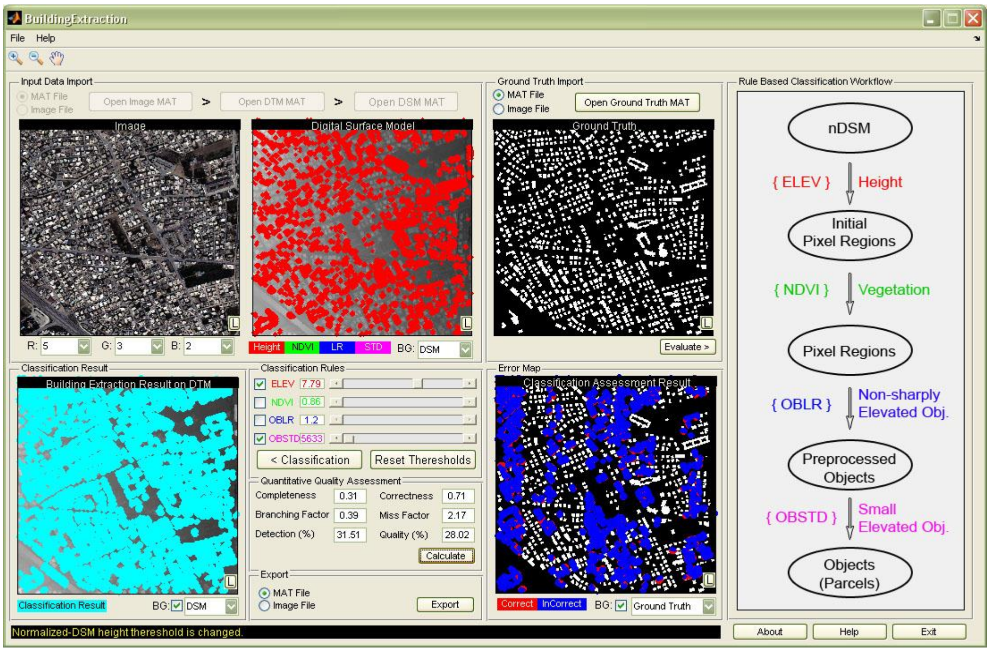

Fig. 4. Program Interface

In the program, three rules are selected to affect the process of building height classification. At the beginning based on heights of the nDSM, it is possible to classify it by moving a slider. Moving the slide changes the height threshold, and finally a logical and acceptable threshold would be considered. Two other rules are also imposed to filter and have a better classification. The "Standard Deviation" and also the "NDVI" are two limits for this goal. Standard Deviation is calculated via the obvious formula as expected value of each pixel value and its average difference. NDVI (Normalized Difference Vegetation Index) roughly is the index of vegetation and is calculated by division of difference and summation of NIR and Red bands of the satellite imagery.

SIMPLE PROCESS OF RULE-BASED CLASSIFICATION

After the data input, some default values for Elevation, STD and NDVI thresholds are considered which could be changed gradually to achieve an efficient threshold. Finally, the user must click on the button of "Classification" to see the result of imposing the rules as described above.

As discussed above, Fig. 4 obviously demonstrates how aforementioned data are loaded and after affecting the rules of Rule-based Classification, the results, assessment and evaluation parameters are calculated and presentable.

For “visual inspection”, an error map is shown as well in the right down of the window and the four categories discussed in future sections are shown. The correctly classified results are shown in red and the incorrect ones are shown in blue. The best inspection can be done by viewing this error map with “Ground Truth” as background; but other backgrounds as DTM, DSM and also nDSM are available to view as base.

Six criteria as will be discussed in the aforesaid section are revealed in the special frame for it. This frame is placed below the threshold sliders in the program interface (Fig. 4).

Users can have a sort of export from the program output in the special frame in the middle bottom of the program. It is made possible to export in form of MAT file or an image.

OUTPUT EXPORT

Result of the classification could be saved for further post-processing via the module of export close to Exit button of the program.

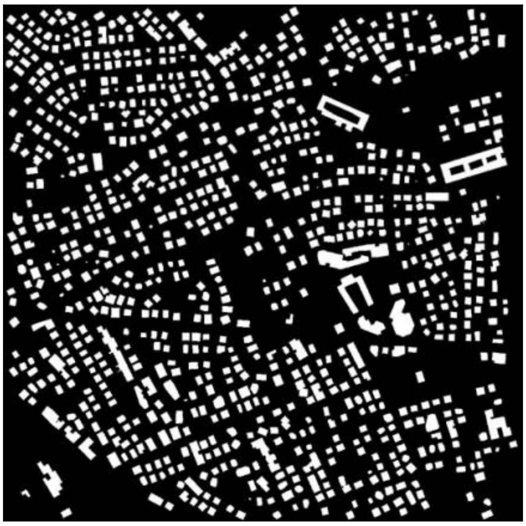

Fig. 5. Ground Truth Map

Fig. 6. Classification results

PROGRAM OUTLINE

Fig. 5 and Fig. 6 obviously demonstrate the ground truth map and the classification result. The ground truth can be used for quality assessment of results. Assessment report will be available after determination of assessment method and clicking the related button.

Fig. 4 (program interface) shows the effect of a fair threshold result for Elevation and Vegetation Index (NDVI) as pixel-based thresholding and also Object-based Local Ranging and Object-based Standard Deviation. After excluding the low height areas from the nDSM and vegetated heights like trees by the index of these features as NDVI from Elevation based selected areas, the process of object analysis is started and in two procedure of object classification Local Range of the pixel regions (which are considered as objects in this situation) are undertaken to process. By a 7x7 window the minimum and maximum values of nDSM on the border of each object is analyzed and can be thresholded. In the process of Object-based Standard Deviation calculation it must be described that the objects (parcels) with STD<Th4 which Th4 is an acceptable threshold, final results are achievable.

RESULTS ASSESSMENT

In this study, two kinds of assessment methods have been considered. The former is visual inspection, and the latter is the quantitative criteria.

QUALITATIVE CRITERIA

In visual-inspection state blue objects indicate regions which have wrongly been detected as building by our method. On the other side red objects indicate regions which have correctly been detected as building (Fig.7).

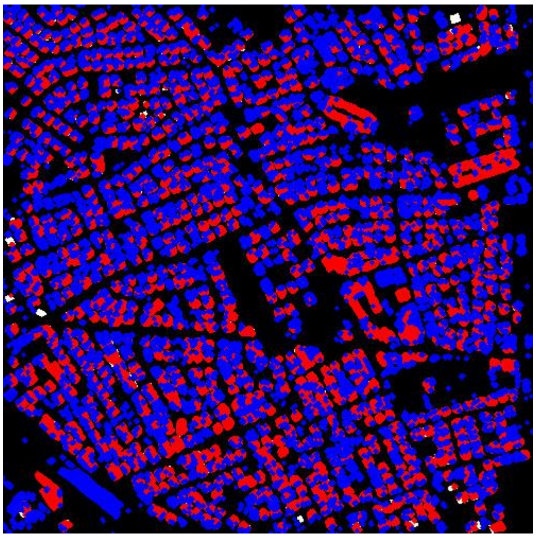

Fig. 7. Classification Result Evaluation based on Qualitative Criteria. In this figure Red pixels demonstrate the true classification and blue pixels depict the incorrect classification based on the ground truth data.

QUANTITATIVE CRITERIA

For evaluation of the raster building extraction quality, the results of the automatic procedure and also the reference building data base were rasterized. The extracted buildings were compared pixel by pixel to the buildings in reference data (ground truth). The standard statistical parameters are defined and measured as follows (Haala and Brenner, 1999):

1- **True positive (TP)** – both the automated method and the rasterized reference building database label a pixel as a building; in other words, The pixels which are building on the ground truth and classification classified them as building.

2- **True negative (TN)** –both the automated method and the rasterized reference building database label a pixel as background (non-building); in other words, The pixels which are not building on the ground truth also, they are classified as non-building.

3- **False positive (FP)** –only the automated method labels a pixel as a building; in other words, The pixels which are not building on the ground truth but they are incorrectly assumed as building.

4- **False negative (FN)** –only the rasterized reference building database labels a pixel as a building; in other words, The pixels which are building on the ground truth and classification classified them as non-building.

Using these four categories, the following statistical measures were computed to evaluate the performance of the automated building extraction process.

$$Branching\ Factor = \frac{FP}{TP} \qquad Detection\ Ration = \frac{TP}{TP + FN}$$

$$Correctness = \frac{TP}{TP + FP} \qquad Completeness = \frac{TP}{TP + FN}$$

$$Miss\ Factor = \frac{FN}{TP} \qquad Quality\ Ration = \frac{TP}{TP + FP + FN}$$

Interpretation of the above calculation is as follows. The **branching factor** indicates the rate of incorrectly labeled building pixels. The **miss factor** gives the rate of missed building pixels (the automated method incorrectly labels pixels as background). The **detection percentage** denotes the percentage of building pixels correctly labeled by the automated process. The **quality percentage** measures the absolute quality of the extraction and is the most stringent measure. It describes how likely a building pixel produced by an automatic approach is correct. The **completeness** is the percentage of the reference data which is explained by the extracted data, i.e., the percentage of the reference network which lies within the buffer around the extracted data. The optimum value for the completeness is 1. The **correctness** represents the percentage of correctly extracted building data, i.e., the percentage of the extracted data which lie within the buffer around the reference networks (Khoshelham et al., 2010).

ASSESSMENT RESULTS

Based on the qualitative evaluation, Fig. 7 illustrates the graphical output of the program. As the quantitative assessment results based on the evaluation parameters given in previous section, in MATLAB are:

$$Branching\ Factor = 0.4 \qquad Detection\ Ration = 84.44\%$$

$$Correctness = 0.71 \qquad Completeness = 0.84$$

$$Miss\ Factor = 0.18 \qquad Quality\ Ration = 63.06$$

DISCUSSION

An algorithm is presented to hierarchically refine the classification results of high resolution worldview DSM and orthorectified images using different feature criteria. A MATLAb tools is produced for this implementation which helps to interactively monitor the effects of including to excluding each feature descriptor to the classification procedure. The refinement of segments is employed in both pixel- and region-based classification steps. The final result is compared statistically to the ground truth which is manually digitized. Final assessment results which are described in the last section obviously figure the effectiveness of the method.

REFERENCES

Arefi, H., Engels, J., Hahn, M., Mayer, H. LEVELS OF DETAIL IN 3D BUILDING RECONSTRUCTION FROM LIDAR DATA. International Archives of Photogrammetry and Remote Sensing, Vol. 37, Com III, Part B3b, pp. 485 - 490.

Arefi, H., Hahn, M. (2005) A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 36, p. 3.

Awrangjeb, M., Ravanbakhsh. M., Fraser, C. (2010) Automatic detection of residential buildings using LIDAR data and multispectral imagery, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 65, pp. 457-467.

Cheng, L., Gong, J., Chen, X., Han, P. (2008) Building boundary extraction from high resolution imagery and lidar data, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 37, pp. 693-698.

Chen, L. C., Tao, T., Shao, Y., Lai, Y., Rau, J. (2004) Fusion of LIDAR data and optical imagery for building modeling, International Archives of Photogrammetry and Remote Sensing, vol. 35, pp. 732-737.

Demir, N., Baltsavias, E. (2009) Extraction of buildings using images & LIDAR data and a combination of various methods, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 38, p. W4.

Elberink, S. O. (2008) Problems in automated building reconstruction based on dense airborne laser scanning data, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 37, pp. 93-98.

Grigillo, D. et al. (2012) Automated building extraction from IKONOS images in suburban areas, International Journal of Remote Sensing, vol. 33.

Haala, N., Brenner, C. (1999) Extraction of buildings and trees in urban environments, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 54, pp. 130-137.

Hahn, M., Arefi, H., Engels, J. (2007) Automatic building outlines detection and approximation from airborne LIDAR data, Urban and Regional Data Management UDMS 2007 Annual, p. 15.

Khoshelham, K., Nardinocchi, C., Frontoni, E., Mancini, A., Zigaretti, P. (2010) Performance evaluation of automated approaches to building detection in multi-source aerial data, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 65, pp. 123-133.

Krauss, T., Lehner, M., Reinartz, P. (2008) Generation of coarse 3D models of urban areas from high resolution stereo satellite images, International Archives of Photogrammetry and Remote Sensing, 37 (B1), pp. 1091-1098.

Lee, D. H., Lee, K. M., Lee, S. U. (2008) Fusion of lidar and imagery for reliable building extraction, Photogrammetric Engineering and Remote Sensing, vol. 74, p. 215.

Li, Y., Wu, H. (2008) Adaptive building edge detection by combining lidar data and aerial images, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 37, pp. 197-202.

Mayer, H. (1999) Automatic object extraction from aerial imagery—a survey focusing on buildings, Computer vision and image understanding, vol. 74, pp. 138-149.

Myeong, S., Nowak, D., Hopkins, P., Brock, R. H. (2001) Urban cover mapping using digital, high-spatial resolution aerial imagery, Urban Ecosystems, vol. 5, pp. 243-256.

Rottensteiner, F., Clode, S., Kubik, K. (2005) Using the Dempster–Shafer method for the fusion of LIDAR data and multi-spectral images for building detection, Information Fusion, vol. 6, pp. 283-300.

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

Sun, J., Li, Y., Kang, S. B., Shum, H. Y. (2005) Symmetric stereo matching for occlusion handling, pp. 399-406 vol. 2.

Vu, T. T., Yamazaki, F., Matsuoka, M. (2009) Multi-scale solution for building extraction from LiDAR and image data, International Journal of Applied Earth Observation and Geoinformation, vol. 11, pp. 281-289.