International Journal of Engineering &Technology, 7 (4.24) (2018) 95-99

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

A Multi Stage Approach for Urban Building Extraction from Remote Sensing Satellite Images

VSSN Gopala Krishna Pendyala[1]*, Hemantha Kumar Kalluri[2], C.V. Rao[3]

[1]National Remote Sensing Centre, Hyderabad, Telangana, India
[2] Department of Computer Science & Engineering, Vignan's Foundation for Science Technology and Research, Vadlamudi, Andhra Pradesh, India
[3]National Remote Sensing Centre, Hyderabad, Telangana, India
**Corresponding author E-mail: gopalakrishna_pvssn@nrsc.gov.in*

Abstract

The most important parameter for urban information system is the building information which is represented by the geographic location of the buildings as well as the area, perimeter, density, inter building distances. This data is integrated with demographic data for various applications. High resolution Remote sensing images are widely used as primary data for automatic extraction of building information. Many researchers have developed different methods for maximizing the detection percentage with minimum errors. This paper analyzes the primary data available for researchers, deriving the secondary information and utilizing it effectively. Case studies by various researchers were analyzed and a methodology has been outlined using their experiences, which is expected to be more efficient and reduced errors.

Keywords: *Automatic Extraction; Digital Elevation Model; High Resolution Satellite Images; LiDAR, Digital Surface Model; Urban Buildings; Urban Information System.*

1. Introduction

The introduction of Geographic Information System (GIS) enabled efficient planning, monitoring and regulating the utilization of natural resources to the optimum. GIS has the ability to correlate, foresee, manage and disseminate geographical information [1]. In particular to the urban areas, this is more effective as the population density is very high and there is a high demand for water, electricity, essential commodities, sewerage management etc. Many of urban administration aspects such as planning, development, infrastructures, transport, telecommunications, electricity, security, health, disasters management, environment protection, education, culture, entertainment involve geo-spatial data for which information systems provide means of management [1, 2].

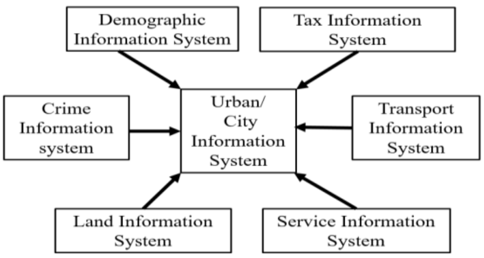


Fig. 1: Urban Information System

Future planning also is very much essential as urban growth is exponential and one has to look for next 10 to 20 years ahead in terms of road network, new settlements, water and other essential provisions. Moreover, revenue collection, monitoring of growth and assessment mechanisms also need to be addressed [3]. To address the multifold requirements to plan, monitor and regulate the resources, administrators have been using various information systems. These systems are either independent or interdependent in nature. The typical urban/ city information system may consist of all or some the sub information systems as shown in figure 1.
The three key aspects of any Information system is (i) accurate (at least in relative terms) depiction of the locations of all physical parameters, which makeup the network. (ii) the ability to define a model of the utility as near to the status on the ground by using the physical parameters and (iii) the ability to integrate a spatial representation of all the individual models into a single system.
The most important input for these models is the urban building information which is represented by the geographic location of the buildings in relation to the ground. The building information such as area, perimeter, density, inter building distance are collected and integrated with the demographic data for each building. This is a very handy tool for many city planners and civic amenity providers in deriving various GIS based applications which include GIS based Digital Taxation, monitor plan deviations, utility service planning, Transportation Planning, Land Information System, Assets Management and Maintenance, Traffic Density Studies, Environmental Impact Analysis, Disaster Management and Mitigation, Community Development, Hot Spot Analysis, Crime Analysis etc.

Table. 1: Comparison of Methodologies for Generating Urban Building data.

Method	Type Data set	Location	Time*/ sq.km.	Cost*/ sq. km.	Remarks
Field method	None / Base map	Field	Very high ~250 hr	Very high Rs. 50000 to 60000	Highly accurate and true shapes are obtained
Manual Mapping	HR satellite data	Manual in Lab	High ~100 hr	High Rs. 3000 to 4000	Accurate and true shapes are obtained.
Automatic / semi-automatic	Fusion of image and height data with another primary / derived data set	Automatic / semi-automatic in lab	Less than ~1 hr	Less than Rs. 500	Base Data set to be good. Accurate corners of the buildings are not preserved

* Assumed 3000 buildings for sq. km. area

2. Building footprint data collection

There are 3 types of approaches in deriving the basic building footprint information for usage in the Urban Information system. The first approach deals with complete manual and field methods in which all the buildings are physically visited and all corners of the buildings are measured using a positioning device such as Global Positioning System (GPS). This data is transferred to the computer and the buildings are drawn digitally by using Commercial-off-the-shelf *(*COTS*)* software.

The second approach is the manual mapping. Having the high resolution aerial / satellite images as base data, the building footprints are manually captured using the mapping software on the computer after preprocessing of images for accurate geo-location. The operator performs the vectorization process on screen and the limiting factors are the experience and performance of the operator and used time [4]. This method requires some amount of ground data for the preprocessing

The third method relies on semi-automatic/ fully automatic methods to generate the foot prints of buildings by using hybrid datasets of HR satellite remote sensing images. The more popular method under this category include using images and by overlaying the derived height information (using photogrammetric technique/ LiDAR).

While selecting an approach, the time required for generating a data set as well as the cost requirement are to be considered. They may be grouped into low, medium, and high average costing, depending on the methodology used and the requirement of accuracy and precision with which it is required. The cost and time requirement are directly proportional and higher average costs are incurred on the data acquisition on which more detailed models are generated. By automatic methods, it is possible to reduce the time and cost requirements to a reasonable level with some limitations.

As seen from the table 1, the accuracy and preserving true shape and size of building are the two considerations for adopting any approach. Utilities such as tax assessment, checking the building plan deviations deals with details of the individual buildings, hence, does not accept any tolerances. However, most of the other utilities mentioned previously, require the outline of the building or center location of the building. Other utility applications deal with groups/clusters of the buildings and the association and dependency properties with respect to one another. These can use the dataset from the automatic/ semi-automatic method of deriving buildings. This offers a significant cost and time reduction in generating the base datasets for that particular urban information system. Hence various researchers have concentrated on the development of methodology either automatic / semi-automatic process of extracting building bases from remote sensing images.

3. Remote Sensing data types:

Remote sensing data is available in various forms and consist of imaging and non-imaging forms. It can be either primary data or secondary data which is derived from primary data. The details of the frequently used data sets are presented in detail.

3.1. Image data

It was analysed by Pendyala et.al. [5] that for automatic building detection in urban areas, better detection results are achieved using high resolution image data supplemented with height information either by DEM generated from optical data / DEM derived from LiDAR data.

3.1.1. High resolution RS satellite images

Optical Remote Sensing satellites of new generation are being launched in order to acquire two types of images: panchromatic and multispectral. The panchromatic (PAN) has better spatial resolution and the multispectral comes with a better spectral resolution in 3 to 4 bands. The panchromatic images have a higher geometric resolution/ spatial resolution varying from Very High Resolution (VHR) of 0.4 to 1 m and multispectral from 1.65 to 4 m [6]. The 4 band multi spectral images consist of Red (R), Blue (B), Green (G) and Near Infra-Red (IR) bands. Table 2 lists the capabilities of the currently available remote sensing high resolution image data [7 - 9]. As the building sizes are going to be as small as 3m X 3m, the data sets that are less than 1m are only useful for this purpose.

Table 2: Various High Resolution Satellites and Capabilities

ID	Satellite / Sensor	Company	Spectral Bands	Resolution in meters
1	IKONOS 2	GeoEye	PAN & 4 band multi Spectral	0.8 3.2
2	QuickBird -2	DigitalGlobe	PAN & 4 band multi Spectral	0.72 2.6
3	GeoEye-1	GeoEye	PAN & 4 band multi Spectral	0.41 1.64
4	WorldView-1	DigitalGlobe	PAN	0.5
5	WorldView-2,3&4	DigitalGlobe	PAN & 8 band Multi Spectral	0.46 1.8
6	Cartosat - 2	ISRO	PAN	<1.0

3.1.2. Aerial images

Aerial Remote Sensing provides high resolution images and elevation data by using the aircraft as a platform. The typical data acquisition sensors include Large Format Digital Camera (LFDC) and Airborne LiDAR. The advantages of this method of data acquisition are that very high resolution (typically 0.05 m). The cost and time factor are very large for acquiring this data.

3.2. Non image data

3.2.1. Digital Surface Model (DSM)

DSM consists of the information about the height of ground as well as vegetation and manmade objects. DSM can be obtained by 2 methods:

From stereo pair of images: A stereo pair is the one which has an overlapped /common area coverage in 2 images acquired from different angles [10]. The geo-location accuracy of the same point on both the images will not be registered properly due to the geometric errors in satellite image orientation. The satellite geometry is reconstructed and satellite images are reoriented using the image geometric model provided by the Rational Polynomial Coefficient (RPC) [11] RPCs and the images and elevation data is generated from the stereo images.

From Airborne LiDAR: Airborne LiDAR (Light Detection and Ranging) is an instrument works on the principle of laser ranging. As the name implies, it is an airborne instrument that measures the distance from ground to receiver by emitting a laser purse from the aircraft. The round trip travel time between aircraft and ground is converted to distance [12]. Airborne LiDAR system provides direct measurement of heights, hence it is more accurate. LiDAR covers large areas with comparatively less turnaround times as the pre-processing takes less time compared to photogrammetric methods [12, 13].

3.2.2. Digital Terrain Model (DTM)

DTM also often called as Digital Elevation Model (DEM), and essentially it contains only ground information. This is the derived product of DSM, generated by removing points belonging to trees and other manmade objects such as buildings, power lines etc.

3.2.3. Normalized DSM (nDSM)

Obtained from DSM and DTM height differences as shown in (1), consisting of information about vegetation, buildings and other manmade objects [14].

$$\text{nDSM} = \text{DSM} - \text{DTM} \quad (1)$$

3.2.4. Normalized Differential Vegetation Index (NDVI)

It is an index value having the relation between red (R) and near-infrared (IR) bands of multi spectral image as shown in (2), widely used for the detection of the vegetated areas [15].

$$\text{NDVI} = \frac{\text{IR} - \text{R}}{\text{IR} + \text{R}} \quad (2)$$

4. Discussion on various approaches

In the effort to analyse the approaches adopted for automatic/ semi-automatic urban building extraction, literature study is taken up. The main objective of this study is to understand the type of data sets used by various researchers, the algorithms applied, and the analysis of the results obtained and to propose an improved approach in terms of datasets and outcome. The literature published by various researchers was analyzed and presented below:

4.1. Data sets used

The table 3 lists the datasets used by various researchers:

Table 3: Datasets Used for Urban Building Extraction

Data set	Resolution	Author Reference
Mono/ Fused satellite images	0.5m - 1.0m	[14, 16, 17]
Aerial images	0.09 m - 0.20 m	[15, 18]
LiDAR DSM	0.15m height accuracy with 1 to 2 m posting	[15, 17]

Various data sets, such as aerial photographs of high resolution, satellite data from mono to multispectral ranging from 0.1 to 1.0 m resolutions, in fusion with Airborne LiDAR generated DSM of different resolutions and point densities are used for automatic building extraction in urban areas.

4.2. Methods

The general approach consists of applying sequence of steps systematically such as data preparation, segmentation, classification, feature extraction and selection. The data preparation uses shadow or NDVI information to mask out the unwanted information. Segmentation is carried out by applying region growing method on images and elevation based threshold [15, 17, 19] on height information. Edge detection [18] is used for building delineation.

The work carried out by various researchers is psented below:

1. Shaker et al. [10] oriented the IKONOS stereo images by using Rational Polynomial Coefficients (RPC) in association with ground control points (GCP) to generate buildings in dense residential areas. Buildings were identified by combining image spectral properties and the extracted DSM from this model.
2. Demir et al. [15] used 4 different approaches on aerial images and Lidar data together. The methods are based on (a) DSM/DTM comparison and then analysis by combining NDVI (b) Image classification of multispectral data combined with Lidar data (c) NDVI classification associated with Lidar DTM and (d) Analysis using Raw Lidar DSM profile data.
3. Benarchid et al. [16] used the method object-based classification and shadow information on GEOEYE-1 very high resolution multispectral images. In the object-based image classification, Mean Shift Clustering algorithm was used for segmentation. Subsequently, classification is carried out by using Support Vector Machines (SVM). Shadow information was related to their buildings to improve efficiency.
4. Hermosilla et al. [17] used high spatial resolution imagery of aerial data and Quickbird data with LiDAR DEM data. 2 different approaches were used on the data. In the first method, height threshold was applied on nDSM (LiDAR data) and spectral response threshold on NDVI (using multispectral image data). The other approach followed the object-based image classification. Segmentation by region growing method applied on nDSM, feature extraction and classification using decision trees. Features are searched having mutually exclusive characteristics and homogeneous subgroups are generated till the stopping condition is satisfied.
5. Jiang et al. [18] a semi-automatic method used on aerial images that applies mean shift segmentation on the image and then the region extraction is implemented. Once satisfied with the results, the boundaries of buildings are delineated by edge detection.
6. Ni et al. [19] used step-wise point cloud segmentation on LiDAR point clouds by binning data according to surface roughness. Then, classification, feature extraction, Random Forests (RF) based feature selection and was applied.
7. Bittner et al. [20] used Fully Convolution Network (FCN) architecture on the Digital Surface Model (DSM) generated

from WORLDVIEW-2 images. This is a semi-automatic method by training the FCN by taking inputs from nDSM and output as ground building mask as to create a final binary building.

Table 4: Details and Outcome of various methods discussed

S. No	Author	Data set	Method	Detection Percentage (3)	Quality percentage (4)
1	Ibrahim F. Shaker et al. [10]	IKONOS stereo images	Image data with DSM extracted from images	77.8 - 87.3 (3 areas) Average: 82.55	70 - 84 (3 areas) Average: 77
2a	N. Demir et al. [15]	Aerial images with LiDAR DEM	DSM/DTM comparison + NDVI analysis	90.0*	68.4*
2b		Aerial images with LiDAR DEM	Classification over multispectral images the applying LiDAR height information.	87*	74.82*
2c		Aerial images with LiDAR DEM	NDVI classification + voids LiDAR DTM	87*	75.69*
2d		Aerial images with LiDAR DEM	Raw LiDAR DSM vertical profile analysis	83*	76.36*
3	O. Benarchid et al. [16]	Geoeye-1	Object based classification	74.60	69.00
4a	Txomin Hermosilla et al. [17]	Quickbird with LiDAR DEM	Thresholding	86.3-89.1 (3 areas) Average: 87.7	81.1 to 66.1 (3 areas) Average: 73.6
4b			Object based classification	85.5 to 86.3 (3 areas) Average: 85.9	53.6 to 73.9 (3 areas) Average: 63.75
5	N. Jiang et al. [18]	Aerial images	Segmentation, region selection and edge detection	Not provided	Not Provided
6	Huan Ni et al. [19]	LiDAR DEM	Segmentation feature selection and classification	85.23 - 92.2* (2 areas) Average: 88.72	82.26 to 91.24* (2 areas) Average: 86.75
7	K. Bittner et al [20]	WorldView -2	Fully Convolution Network (FCN)	86	78

* After Normalization

4.3. Parameters for analysis

It is necessary to compare the methods on a common scale. The Detection Percentage (DP) and Quality Percentage (QP) are the quantitative parameters used by various researchers [14, 16, 17, 20].

Where, True Positive (TP) indicate the number of correctly extracted buildings, False Positive (FP) indicate the number of buildings detected by the automatic approach that are not truly present on the ground and False Negative (FN) represents undetected buildings actually present on the ground, then

Detection percentage: This is the percentage of buildings which were actually detected by the automated system to that are available on the ground.

$$\text{Detection Percentage (DP)} = \frac{TP}{TP+FN} \quad (3)$$

Quality percentage: This indicates the overall algorithm performance indicating the measure of wrong detection or misbehavior of algorithm. True negatives and False positives will lower the quality percentage.

$$\text{Quality Percentage (QP)} = \frac{TP}{TP+FN+FP} \quad (4)$$

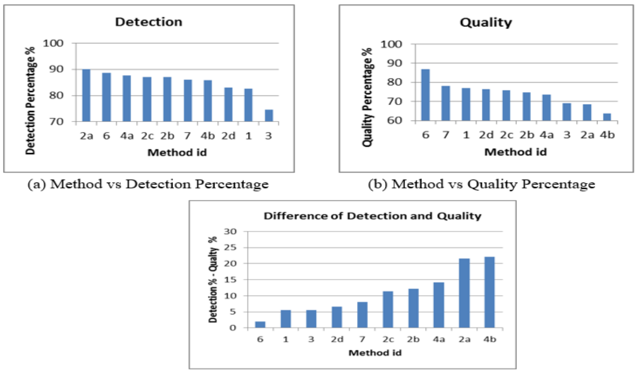
4.4. Achieved Results

The quantitative results obtained are tabulated below in table 4.

5. Analysis

The analysis of various methods by comparing the Detection %, Quality % and the difference between Detection percentages to Quality Percentage is carried out. The Charts outlining the same derived from Table 4 is shown in figure 2. Quality Percentage is carried out. The Charts outlining the same derived from Table 4 is shown in figure 2 below:

After analysis, it can be seen that, from figure 2(a), highest detection percentage is achieved by combining image and LiDAR data and incorporating NDVI information for analysis [15] or using only LiDAR data without any image data [19]. The quality factor improves by utilizing the LiDAR DEM [19] or FCN method [20] as can be viewed in figure 2(b). The most robust method which offers least deviation between the detection to quality percentage is found to be LiDAR data (1.97%) [19] and then IKONOS image data with the extracted DSM from the same image data (5.55%) [10]. This is shown in figure 2(c).



(a) Method vs Detection Percentage (b) Method vs Quality Percentage

(c) Method vs Detection – Quality Percentage

Fig. 2: Analysis of various methods

As the LiDAR data acquisition is costly and time consuming, and the FCN method needs the extensive training data sets, these methods are not very practicable for fully automated approaches in production environment for large areas. Hence, a new method is designed to combine the advantages of (i) HR satellite stereo images (ii) derive the DEM from photogrammetry to improve robustness and (iii) Associate NDVI to improve detection percentage.

6. Proposed Methodology

After the literature study and analyzing, an efficient multistage approach is proposed to be implemented for research to extract the automatic extraction of buildings from HR satellite data.

6.1. Methodology flow chart

The flow chart is provided in Fig.3.

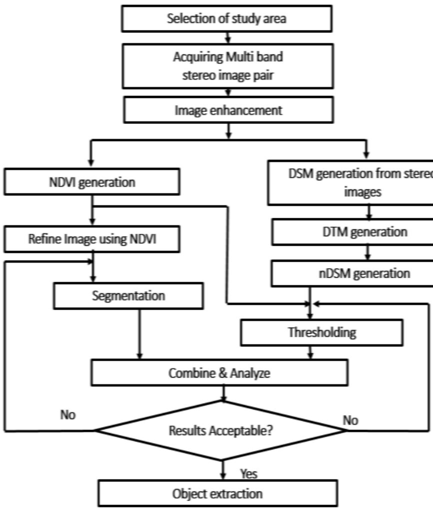


Fig. 3: Flow chart of proposed methodology

6.2. Description of stages

The various stages and their functionalities are explained below:

Image enhancement: Image enhancement is to be carried out to highlight the details of buildings in the image. This algorithm can be a min-max or histogram equalization using thresholding.
Generation of DSM: DSM is generated using stereo images by applying photogrammetry technique and edit to remove noise.
Refine DSM and generate DTM: Remove the vegetation and manmade feature height points from DSM to generate DTM.
NDVI generation and applying on the image data: Generate the NDVI using multiple channels of Multispectral image and apply this mask on the image to remove the vegetation areas.
Generate nDSM: The differential surface creation is done as per the equation (3).
nDSM Thresholding: Thresholding operation on NDSM is carried out by to extract possible building points.
Segmentation: Use the segmentation method on the refined image (after applying NDVI) to get the regions of possible buildings by either Thresholding / region growing/ clustering approach and identify the most possible building areas.
Combine and Analyze: Select the best combination method (union/ intersection) and analyze the results with ground truth collected. Refinement may be required on any of the above operations till satisfactory results are obtained.
Object extraction: Combine the results obtained from DSM Thresholding and image segmentation and identify the region boundaries of most probable buildings.

7. Conclusions

It is essential to have the height information with High Resolution satellite images to generate urban building footprint information effectively. Since obtaining height data through LiDAR is costly and time consuming, a new methodology is proposed to generate data having better detection and quality percentage simultaneously. The proposed method uses same image data set to derive the DSM which offers the advantages such as automatic registration of elevation data and image data. This method offers robust building footprint information, with very less number of false detections as nDSM and NDVI are simultaneously being used. The present method may have limitations as NDVI method may remove some green buildings.

References

[1] Carmen Radut, Ramona Chitu (2009), Geographical Information Systems and Urban Management. Theoretical and Empirical Researches in Urban Management.
[2] Itzhak Benenson, Shai Aronovich, Saar Noam (2005), Let's talk objects: generic methodology for urban high-resolution simulation. Computers, Environment and Urban Systems 29 425–453.
[3] Mahak, Khurana., Vaishali, Wadhwa.(2015), Automatic Building Detection Using Modified Grab Cut Algorithm from High Resolution Satellite Image. International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 8.
[4] Topan, H., Oruc, M., Jacobsen, K. (2009), Potential of manual and automatic feature extraction from high resolution space images in mountainous urban areas. ISPRS Hannover Workshop, www.isprs.org/proceedings /XXXVIII/ 1_4_7-W5.
[5] VSSN Gopala Krishna Pendyala, Hemantha Kumar Kalluri, V Raghu Venkataraman and C V Rao (2018), Comparative study of Automatic Urban Building Extraction Methods from Remote Sensing Data. Advs in Intelligent Syst., Computing, Vol. 815, in press.
[6] Małgorzata, Verőné Wojtaszek.(2015), Opportunities of object-based image analysis for detecting urban environment. presentation at IGIT, Internet resource, www.geo.info.hu.
[7] Huan, Ni., Xiangguo, Lin., Jixian, Zhang.(2017), Classification of ALS Point Cloud with ImprovedPoint Cloud Segmentation and Random Forests. Remote Sensing
[8] "Satellite Mission Database" Internet Resource https://earth.esa.int/web/eoportal/satellite-missions 2018.
[9] Internet Resource "https://www.satimagingcorp.com/satellite-sensors/geoeye-2", 2018
[10] Ibrahim, F. Shaker., Amr, Abd-Elrahman., Ahmed, K. Abdel-Gawad., Mohamed, A. Sherief.(2011), Building Extraction from High Resolution Space Images in High Density Residential Areas in the Great Cairo Region. Remote Sesiing.
[11] K Jacobsen (2008), Satellite Image Orientation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B1. Beijing.
[12] Shan, J. and A. Sampath (2005), Urban DEM Generation from Raw Lidar Data: A Labeling Algorithm and its Performance. International Journal of Remote Sensing 71, pp. 217-222.
[13] Meng, X., L. Wang, J. L. Silván-Cárdenas and N. Currit (2009), A multi-directional ground filtering algorithm for airborne LIDAR. ISPRS Journal of Photogrammetry and Remote Sensing 64(1), pp. 117-124.
[14] Sefercik, U G., Karakis, S., Atalay, C., Yigit, I., Gokmen, U. (2017), Novel fusionapproach on automatic object extraction from spatial data: case study Worldview-2 and TOPO5000. Geocarto International.
[15] Demir, N, Poli, D, Baltsavias, E.(2008), Extraction of Buildings and Trees Using Image and LiDAR data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII. Part B4. Beijing.
[16] Benarchid, O., Raissouni, N., El Adib, S., Abbous, A., Azyat, A., Ben Achhab, N., Lahraoua, M., Chahboun, A. (2013), Building Extraction using Object-Based Classification and Shadow Information in Very High Resolution Multispectral Images, a Case Study: Tetuan, Morocco. Canadian Journal on Image Processing and Computer Vision Vol. 4 No. 1.
[17] Txomin, Hermosilla., Luis, A Ruiz., Jorge A Recio., Javier, Estornell.(2011), Evaluation of Automatic Building Detection Approaches Combining High Resolution Images and LiDAR Data. Remote Sensing.
[18] Jiang, N., Zhang, J X., Li, H T., Lin, X G. (2008),Semi-Automatic Building Extraction from High Resolution Imagery Based on Segmentation. International Workshop on Earth Observation and Remote Sensing Applications, Beijing.
[19] Huan, Ni., Xiangguo, Lin., Jixian, Zhang. (2017), Classification of ALS Point Cloud with ImprovedPoint Cloud Segmentation and Random Forests. Remote Sensing.
[20] Bittnera, K., Cuia, S., Reinartz, P. (2017), Building Extraction from Remote Sensing Data Using Fully Convolution Networks. ISPRS Workshop, Hannover.