**ISSN 1608-7844**

DOI: 10.54470/planplus.v11i1.2

Plan Plus, Volume-11, 2021 (17-28)

# AUTOMATIC BUILDING FOOTPRINT EXTRACTION FROM HIGH RESOLUTION STEREO SATELLITE IMAGE

**Md Rimu Mia*[1], Kazi Saiful Islam[2], Md Didarul Islam[3]**
*1.*Corresponding Author: Urban and Rural Planning Discipline, Khulna University, Khulna,*
*E-mail: rimumiandc86@gmail.com*
*2. Professor of Urban and Rural Planning Discipline, Khulna University Khulna 9208, Bangladesh*
*3. PhD student, Department of Geography and Geoinformation Science, George Mason University, VA, USA*

## ABSTRACT

Automatic buildings footprint extraction is of great importance to city planning, urban growth management, and landscape visualization. Although traditional building footprint extraction from two-dimensional images is relatively easy, but are often both time-consuming and costly. Automated building footprint extraction from imagery has been studied extensively based on image segmentation using the pixel value, while the other dimension of segmentation, such as height value, have not been fully explored to extract the building footprints that have been used in this paper. This approach uses the Digital Terrain model (DTM) and Digital Surface Model (DSM) generated from the stereo imagery using the panchromatic bands with high spatial resolution to calculate the Normalized Digital Surface Model (nDSM) to separate the features which are above the ground surface. After extraction of Elevation (Height) of each building feature, the image segmentation has been performed to separate the building features using the threshold value. When the building features are separated, the Canny Edge Detection algorithm is used to delineate the actual building boundary. After the extraction of building boundary, it is vectorized. After simplification of the vector data, the building footprints are extracted. The vector data is compared to the digitized data sets, which show that the approach can be consistent and precise as the building segmentation approach has achieved greater accuracy because of incorporation of height value. There is no human error involved in the whole process.

**KEYWORDS:** Stereo image, DTM, DSM, nDSM, Edge detection, Segmentation, Vectorization.

## 1. INTRODUCTION

With the increase of urban population, planned development, resource management and land distribution became the main concern for the planners and decision makers. And all these applications require related database for monitoring and ensuring efficient living environment (Suliman, 2017). Since buildings are the most prominent objects in urban areas, continuously updated building information, in geographic information systems (GIS), is important for city planning and management applications (Hussain & Shan, 2016). The collection of building footprints often needs a lot of manual work and is both time-consuming and costly. It is challenging to extract building information from remotely sensed data because of the sophisticated nature of urban environments (Davydova et al., 2016). For decades, this extraction of geospatial data has been performed manually because of the inhomogeneities of the building structures. However, manual extraction is a very slow and the invested time and labour increase the cost of operation (Song & Shan, 2008). Building height is an important parameter for defining buildings. Unlike Three-Dimensional (3D) data, Two-dimensional (2D) data does not provide elevation related information (Zeng, 2014).

Using the proposed novel method, it is possible to increase the robustness of the complex process of shape extraction. Over the past decades, several automated methods have been developed using different sources of data. Many methods used spectral reflectance values (Ghandour & Jezzini, 2018; Lee & Shan, 2002) to detect buildings based on aerial and high-resolution satellite imagery. However, these methods often face problems caused by imaging radiometry of optical sensors when similar spectral reflectance of the ground occurs (Davydova et al., 2016). Because of recent advances in the availability of high resolution remotely sensed imagery such as IKONOS, automatic extraction of man-made structures such as buildings and roads from aerial and satellite images has gained increased attention (Aytekin et al., 2009). Stereo pair images from very high-resolution sensors of satellites such as IKONOS, Quick-Bird, World-View, and GeoEye opened opportunities of generating urban DSMs (Alobeid et al., 2009). Using stereo pair images help generation of the DSM and height of the buildings which creates enormous opportunities for automation of the digitization process (Kux et al., 2006), map updating (Caetano & Santos, 2001) and monitoring urban growth which occurs quick in third world cities. In this study, the high-resolution stereo satellite image is used to incorporate elevation data in the urban building footprint extraction.

## 2. STUDY AREA AND DATA DESCRIPTION

The proposed approach has been implemented on the stereo image of Santiago, California, USA. This city is heavily populated and has many high-rise buildings. Because of Bangladesh's dense vegetation and inaccessibility in online data sources, the stereo image is both costly and inappropriate for building extraction; however, it is freely available and suitable for building extraction in the study region chosen. The IKONOS stereo pair image data sets with spatial resolution of 1m (panchromatic band) has used in this study. The images have cross scan: 0.95 meters along scan: 0.94 meters with geotiff format.

## 3. METHODOLOGY

The building extraction approach proposed in this paper is an automatic method, the purpose of which is to delineate the building polygon using the IKONOS stereo pair image. Main steps followed in this research are given in the Figure 1. The approach consists mainly of two parts for extracting the buildings from high resolution stereo satellite images; (i) Data Preparation and (ii) Building Footprint Extraction.

### 3.1 Data Preparation

This stereo pair is composed of two images over the same area at same captured time. Then the two images are inserted to orient the exterior and interior. In the image orientation stage, the RPC coefficient value for each image has been integrated. After doing that, the tie points of the image pair have been generated. Afterwards following tasks are performed-

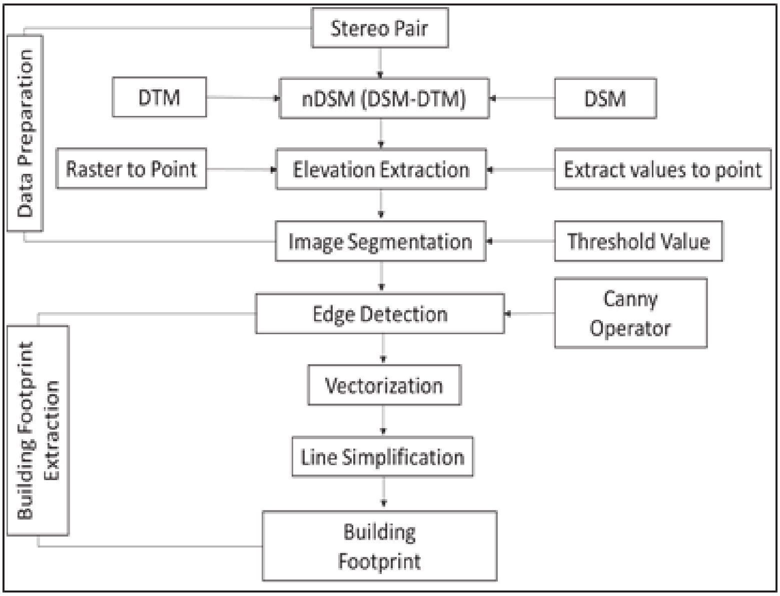


Figure 1: The main steps of the proposed building extraction procedure from high resolution stereo satellite images.

***3.1.1 DSM Generation:***

The DSM represents the features which are above the ground surface. That means the feature that is visible in the generated DSM includes the surface and the features that are above the surface, such as tree, buildings, and roads. Previously, different investigations have been carried out to examine the production of accurate DSMs from satellite images, which have faced many difficulties in producing reliable algorithms that recognize the geometric components of planes, curves as well as suffered from the accuracy. However, some of them revealed expected accuracy, but there was ample scope for improvements (Zhang & Gruen, 2006). The most crucial part of DSM generation is the image matching, which if achieved a high standard leads to a higher quality of DSM. But it is a challenging task (Bertin et al., 2015). And because of some difficulty, in this research, the automatic terrain extraction (ATE) methods have been used for the DSM generation from stereo satellite image (Xu et al., 2008). This method can deliver higher and expected accuracy for stereo satellite image. This method at first extracted the interest points (Förstner & Gülch, 1987) in one of the two images and located these points within another image using the pyramid matching algorithm (Lehner & Gill, 1992). Afterwards, a huge amount of correlation points between two images are generated using the least square matching procedure (Alobeid, 2011). And then the camera parameters have been provided based on the provided RPC files for image orientation (Interior and Exterior). Automatic terrain extraction operation is performed based on the point clouds were. Eventually, this process generates the DSM (Figure 2).

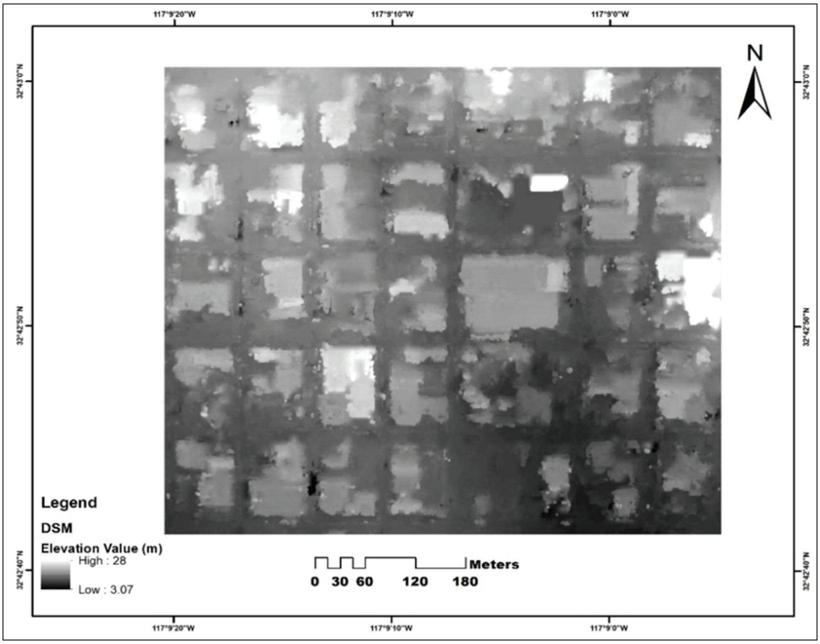


Figure 2: DSM generated from Ikonos stereo pair image
*Source : Author*

### *3.1.2 DTM Generation*

The DTM represents only the ground surface of the earth, which does not contain the feature that is above the earth's surface. This generated DTM provides information about the surface elevation from the Mean Sea Level (MSL). In the need of high quality DTM, several methods have been proposed by different researchers within which the Photogrammetry and Airborne Laser Scanning (ALS) are still two of the most popular approaches (Baltsavias, 1999). ALS data are more dominant in DTM generation, especially in urban region provide more accuracy and building boundaries appears better. In the forest region, the trees are well separated and the more accurate DTM are generated by filtering the trees (Tian et al., 2014). Compared to ALS data, DTM generation from stereo satellite image are more economical as well as helpful for application in larger region (Zhang, 2005). Because of the helpful of high-resolution stereo satellite image to generate the DTM in higher accuracy, in the last decade it became available in photogrammetry and the automatic DTM generation algorithm have replaced the human interaction in the post-processing step of the satellite image (Briese, 2010). In this research, at first the interior and exterior orientations have been performed for the two images using the RPC file, which transforms the original sensor coordinate systems into image coordinate systems (Perko et al., 2014). Then the tie points were automatically computed through multi-image area-based correlation. Then the SGM method (Hirschmuller, 2007) applies to pixel wise image matching with known interior and exterior orientation and epipolar geometry. Finally, the DTM has been generated automatically (Figure 3).

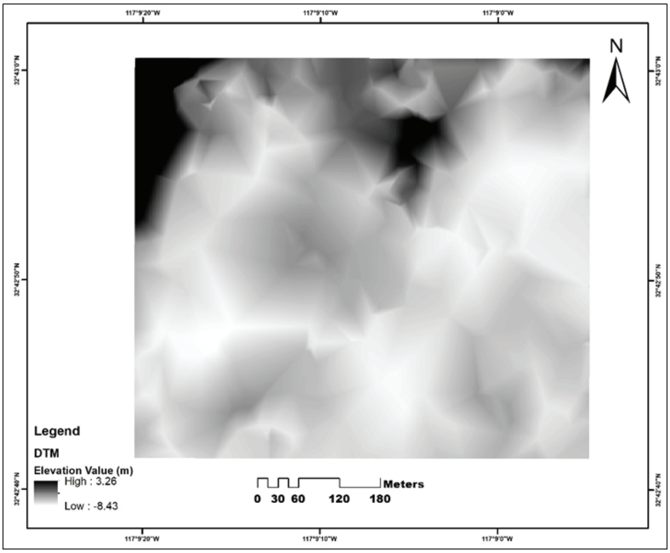


Figure 3: DTM generated from Ikonos stereo pair image
*Source : Author*

### *3.1.3 Normalized DSM*

When the DTM and DSM are generated where DTM is the elevation model of the earth surface that does not include the objects above the ground surface (Podobnikar, 2005) and DSM includes the objects with their heights above the ground surface and topography (Li et al., 2004).

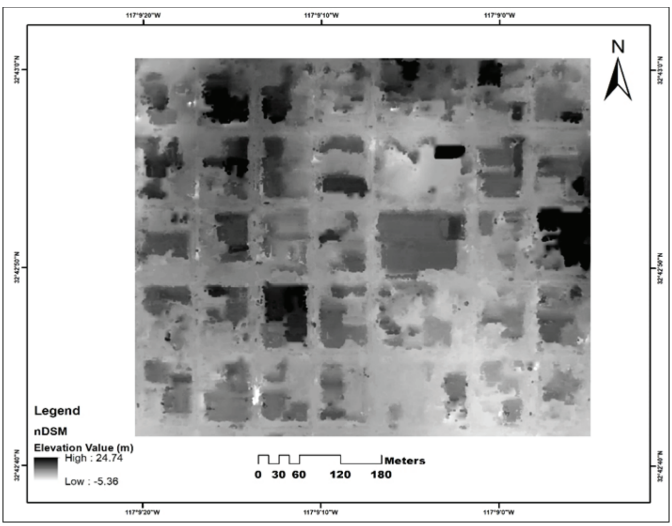


Figure 4: nDSM generated by subtracting DTM from DSM
*Source : Author*

The nDSM is generated by subtracting the DTM from DSM to separate the man-made objects with different heights over the terrain surface (Weidner & Förstner, 1995) (Figure 4).

### *3.1.4 Extraction of Height Value*

When the nDSM has been generated, it is needed to extract the height value of each point of the generated nDSM. In order to do that, the dense point clouds are generated by converting the raster image into point feature. Then elevation values are extracted for each of the points. Then, the point features are cropped for better visualization (Figure 5).

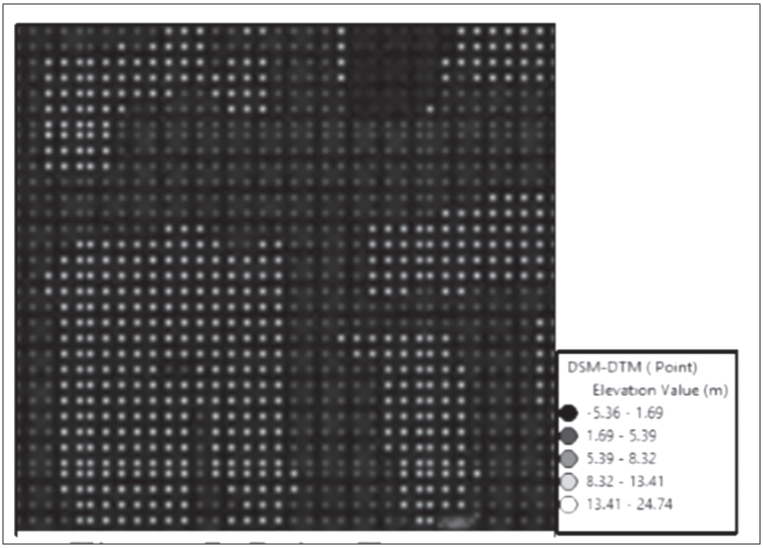


Figure 5: Dense point cloud
*Source : Author*

### 3.2 Building Footprint Extraction

When the data has been prepared to perform the further process the 1. Image Segmentation, 2. Edge Detection, 3. Vectorization and 4. Line Simplification are performed to extract the building footprint.

### *3.2.1 Image Segmentation*

In order to separate the buildings from the other features, the image segmentation has been performed. In this method, a threshold value is used to delete pixels with a lower value than the specified threshold, allowing the pixel containing buildings to be distinguished. Where the 8m elevation has used as the threshold value and the threshold value has selected based on the trial-and-error process. Different threshold values are used until the expected accuracy is achieved. Then the feature which has the elevation of less than 8 m is replaced with 0 m and the value greater than 8 m is replaced with 1 m because of a better representation of the segmented image (Figure 6).

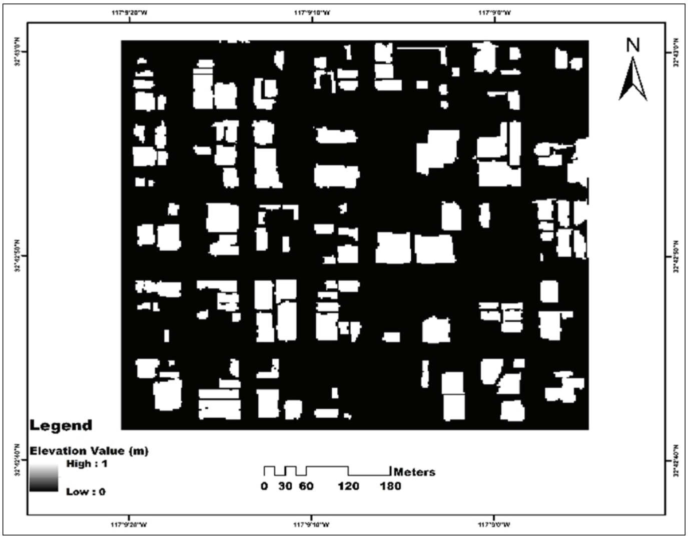


Figure 6: Segmented image
*Source : Author*

### *3.2.2 Edge Detection*

When the building features are separated, the edge detector operator is used to delineate the boundary of the building, which is the most crucial part of this research. For this purpose, Canny edge detection operator is used where the image is processed in the anaconda platform using the python programming language. Where at first the image noises are reduced by smoothing the image and then secondly determined the gradient at each pixel of smoothed image and then found the non-edge pixel which has a certain distance from the actual image and replaced with zero as the pixel value to eliminate them. After that, it was assumed that the output of non-maxima suppression still contains the local maxima created by noise in the image. In order to eliminate them threshold value used to eliminate the pixel which have lower value than define threshold (Maini and Aggarwal, 2009) (Figure 7).

### *3.2.3 Vectorization*

Building edges are delineated with an intention to convert into vector. To convert the edge line into a vector at first, the images are converted to a binary image where nonzero pixels represented an object while the pixels with zero values corresponded to the background. Then the boundary pixels of the buildings are analyzed and the vector line segments are generated from the input binary image. The vector lines are then generated by connecting the coordinates of the detected object pixels (San & Turker, 2007). After doing that, the polygon of the actual buildings are extracted into vectorized format (Figure 8).

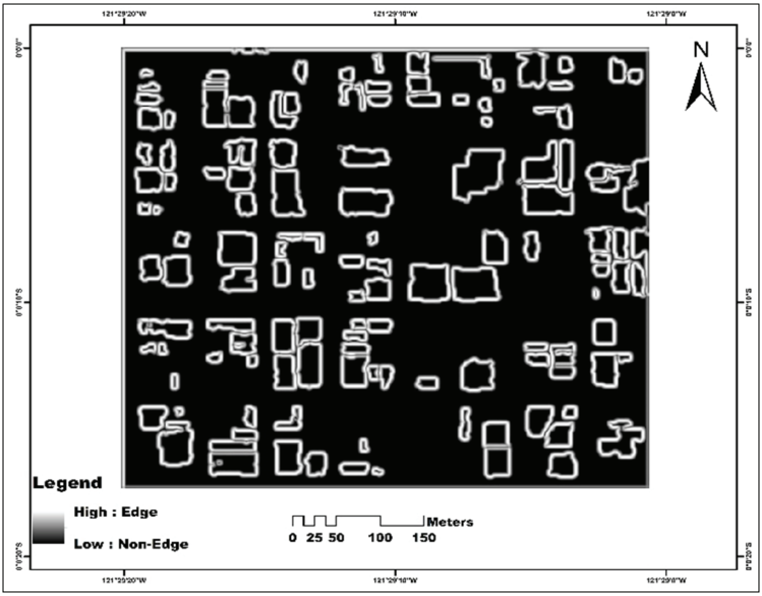


Figure 7: Detected edge
*Source : Author*

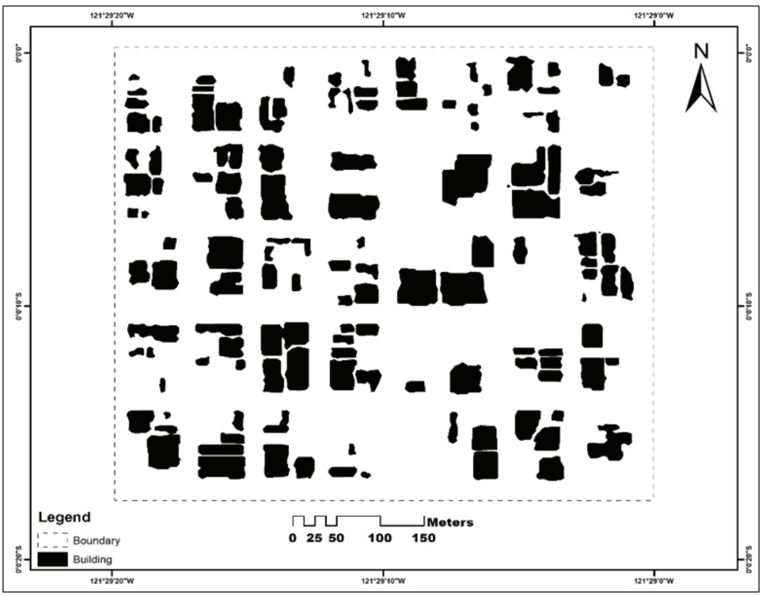


Figure 8: Extracted buildings
*Source : Author*

### *3.2.4 Line simplification*

Vectorized building edges have some issues with the curvature. To get rid of these issues, the line features are simplified to smooth the edge line by adding the vertex point with the edge line which is under 12 m distance (Hershberger & Snoeyink, 1992) (Figure 9).

By doing all the processes, the final goal of my research to automatic building footprint extraction from high resolution stereo satellite image has been achieved.


Figure 9: Building polygons
*Source : Author*

## 4. CONCLUSION

The findings of the automated building footprints derived from the stereo satellite image compared to the digitized data sets show that the approach can be consistent and precise as the building segmentation approach has achieved greater accuracy as the height value has been implemented, and there is no human interaction involved. This automatic method saved a considerable amount of improvement in time and cost to digitize the satellite image for mapping purpose.

Though the research findings have achieved better accuracy, of course, there are obviously a lot of drawbacks. Many buildings are not fully extracted in our experimental results, some extracted buildings are not real buildings and some buildings are not extracted. In fact, the building's real form has been distorted and the edge of the extracted building is not smooth enough. This approach also has limited capacity to detect the building, particularly in the area where trees cover the buildings.

## REFERENCES

Alobeid, A, Jacobsen, K., & Heipke, C. (2009). Building height estimation in urban areas from very high resolution satellite stereo images. ISPRS Hannover Workshop, 5, 2–5.

Alobeid, A. (2011). Assessment of matching algorithms for urban DSM generation from very high resolution satellite stereo images. Fachrichtung Geodäsie und Geoinformatik der Leibniz-Univ.

Aytekin, O., Ulusoy, I., Erener, A., & Duzgun, H. S. B. (2009). Automatic and unsupervised building extraction in complex urban environments from multi spectral satellite imagery. 2009 4th International Conference on Recent Advances in Space Technologies, 287–291. https://doi.org/10.1109/RAST.2009.5158214

Baltsavias, E. P. (1999). A comparison between photogrammetry and laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2–3), 83–94. https://doi.org/10.1016/S0924-2716(99)00014-3

Bertin, S., Friedrich, H., Delmas, P., Chan, E., & Gimel'farb, G. (2015). Digital stereo photogrammetry for grain-scale monitoring of fluvial surfaces: Error evaluation and workflow optimisation. ISPRS Journal of Photogrammetry and Remote Sensing, 101, 193–208. https://doi.org/10.1016/j.isprsjprs.2014.12.019

Briese, C. (2010). Extraction of digital terrain models. Airborne and Terrestrial Laser Scanning, 135–167.

Caetano, M., & Santos, T. (2001). Updating land cover maps with satellite images. IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217), 3, 979–981. https://doi.org/10.1109/IGARSS.2001.976720

Davydova, K., Cui, S., & Reinartz, P. (2016). Building footprint extraction from digital surface models using neural networks. Image and Signal Processing for Remote Sensing XXII, 10004, 100040J. https://doi.org/10.1117/12.2240727

Förstner, W., & Gülch, E. (1987). A fast operator for detection and precise location of distinct points, corners and centres of circular features. Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, 281–305.

Ghandour, A. J., & Jezzini, A. A. (2018). Autonomous building detection using edge properties and image color invariants. Buildings, 8(5), 65. https://doi.org/10.3390/buildings8050065

Hershberger, J. E., & Snoeyink, J. (1992). Speeding up the Douglas-Peucker line-simplification algorithm. University of British Columbia, Department of Computer Science Vancouver, BC.

Hirschmuller, H. (2007). Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2), 328–341. https://doi.org/10.1109/TPAMI.2007.1166

Hussain, E., & Shan, J. (2016). Object-based urban land cover classification using rule inheritance over very high-resolution multisensor and multitemporal data. GIScience \& Remote Sensing, 53(2), 164–182. https://doi.org/10.1080/15481603.2015.1122923

Kux, H. J. H., de Pinho, C. M. D., & e Souza, I. de M. (2006). High-resolution satellite images for urban planning, studies in progress at inpe (national institute for space research), Brazil.

Lee, D. S., & Shan, J. (2002). Generalization of building polygons extracted from IKONOS imagery. ISPRS Commission IV.

Lehner, M., & Gill, R. S. (1992). Semi-automatic derivation of digital elevation models from stereoscopic 3-line scanner data. https://elib.dlr.de/34586/

Li, Z., Zhu, C., & Gold, C. (2004). Digital terrain modeling: principles and methodology. CRC press.

Perko, R., Raggam, H., Gutjahr, K., & Schardt, M. (2014). Assessment of the mapping potential of Pléiades stereo and triplet data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2(3), 103. https://doi.org/10.5194/isprsannals-II-3-103-2014

Podobnikar, T. (2005). Production of integrated digital terrain model from multiple datasets of different quality. International Journal of Geographical Information Science, 19(1), 69–89. https://doi.org/10.1080/13658810412331280130

San, D. K., & Turker, M. (2007). Automatic building extraction from high resolution satellite images for map updating: A model based approach. Urban Data Management: Urban Data Management Society Symposium 2007, Stuttgart, Germany, 10-12 October 2007, 3.

Song, Y., & Shan, J. (2008). Building extraction from high resolution color imagery based on edge flow driven active contour and JSEG. IAPRSIS, 37, 185–190.

Suliman, A. M. H. (2017). Building detection in off-nidar very high resolution satellite images based on stereo 3D information. University of New Brunswick.

Tian, J., Krauß, T., & Reinartz, P. (2014). DTM generation in forest regions from satellite stereo imagery. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(1), 401. https://dx.doi.org/10.5194/isprsarchives-XL-1-401-2014

Weidner, U., & Förstner, W. (1995). Towards automatic building extraction from high-resolution digital elevation models. ISPRS Journal of Photogrammetry and Remote Sensing, 50(4), 38–49. https://doi.org/10.1016/0924-2716(95)98236-S

Xu, F., Woodhouse, N., Xu, Z., Marr, D., Yang, X., & Wang, Y. (2008). Blunder elimination techniques in adaptive automatic terrain extraction. ISPRS J, 29(3), 21.

Zeng, C. (2014). Automated Building Information Extraction and Evaluation from High-resolution Remotely Sensed Data

Zhang, L. (2005). Automatic digital surface model (DSM) generation from linear array images. ETH Zurich.

Zhang, L., & Gruen, A. (2006). Multi-image matching for DSM generation from IKONOS imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 60(3), 195–211. https://doi.org/10.1016/j.isprsjprs.2006.01.001