Automatic Extraction of Arbitrarily Shaped Buildings from VHR Satellite Imagery

Evans Belly, Imdad Rizvi, M. M. Kadam

Abstract—Satellite imagery is one of the emerging technologies which are extensively utilized in various applications such as detection/extraction of man-made structures, monitoring of sensitive areas, creating graphic maps etc. The main approach here is the automated detection of buildings from very high resolution (VHR) optical satellite images. Initially, the shadow, the building and the non-building regions (roads, vegetation etc.) are investigated wherein building extraction is mainly focused. Once all the landscape is collected a trimming process is done so as to eliminate the landscapes that may occur due to non-building objects. Finally the label method is used to extract the building regions. The label method may be altered for efficient building extraction. The images used for the analysis are the ones which are extracted from the sensors having resolution less than 1 meter (VHR). This method provides an efficient way to produce good results. The additional overhead of mid processing is eliminated without compromising the quality of the output to ease the processing steps required and time consumed.

Keywords—Building detection, shadow detection, landscape generation, label, partitioning, very high resolution satellite imagery.

I. INTRODUCTION

IN today's scenario almost more than 50% of the population resides in urban and sub-urban environment [2]. The manual monitoring of the land coverage area is difficult and not feasible as it would provide inaccurate values leading to inconvenient informatory data. So to obtain an acceptable database, the satellite imagery technologies come in picture. As the main concern for humans is the building areas in an environment hence the reliable and accurate extraction of buildings from satellite images becomes an important task. Taking into account the application perspective of satellite imagery, this field emerges as an active research field. To justify the problem the basic concentration is on the automatic building extraction technique from the VHR images.

Automatic detection of buildings in very high spatial resolution remotely sensed imagery has been an important and critical problem because the detection/extraction results can be used in various applications viz. structure change detection, urbanization monitoring, and digital map production. This task also offers an excellent domain for studying the general problems of scene segmentation, 3D inference, and shape description under highly challenging conditions [3]. VHR satellite images provide valuable information to researchers. Among these, urban-area boundaries and building locations play crucial roles. For a human expert, manually extracting this valuable information is tedious and time consuming. One

possible solution to extract this information or data is using automated techniques. The most important data input source to be utilized for the purpose of object extraction is the VHR satellite images [1]. Since more than 50% of the world population lives in urban and sub-urban environments [2], reliable and accurate detection of building objects from satellite images is an essential task and is a very active research field. QuickBird, GeoEye I, GeoEye II, Worldview I, Worldview II are the sensors which provide VHR satellite images whose resolution is 1 meter or less. Human settlement analysis for slum and unorganized settlement monitoring can be assisted by automatically extracted building information because slum areas can generally be characterized by a high density of short and small buildings in irregular spatial arrangements [3].

The recent work performed by Akçay and Aksoy investigates the shadow evidence to focus on building regions [3]. With accordance to this concept, the directional spatial relationship between buildings and their shadows with the prior knowledge of illumination direction is modelled. For the same, a fuzzy landscape generation approach is proposed which is especially designed for modelling the directional relationship between buildings and their shadows. Once all landscapes are collected, a trimming process is applied for the elimination of the landscapes that may occur due to non-building objects viz. roads, sewages, garden wall, bridges etc.

II. LITERATURE SURVEY

Many studies and research have been carried out in the context of building detection, extraction and reconstruction. Simultaneously other man-made structures have also been considered for maintaining and updating geographic information system (GIS) databases. A number of surveys and methodologies were considered in the past to do the same. A state-of-the-art automatic object extraction technique from aerial imagery [4] was surveyed in 1999. This survey included approaches for object extraction from satellite images, which influenced the extraction from aerial imagery. It only covered models and strategies using well defined criteria. Algorithms and underlying technologies were not reviewed. Assessment, Complexity Criteria for the Assessment of Images/Models/ Strategies, Characterization of Models, Characterization of Strategies, Classification of Models and Strategies were the approaches carried out in the survey. Since this was a survey, it rendered as an information source for further analysis. With the existing geo-data a building and road detection technique [6]

Evans Belly is with the Mumbai University, India (e-mail: evns23@gmail.com).

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:12, No:10, 2018

was also developed which focused on the analysis and aspects of knowledge that could be used for extraction such as types of knowledge, problems in using existing knowledge, knowledge representation and management, current and possible use of knowledge, upgrading and augmenting of knowledge [6]. Approaches were also developed for building extraction and updating from high resolution satellite imagery [7].

The developed approaches include two main stages:

(i) Detecting the building patches and
(ii) Delineating the building boundaries.

The building patches were detected from high resolution satellite imagery using the Support Vector Machines (SVM) classification, which is utilized for both the building extraction and updating approaches. In the building extraction part, the previously detected building patches were delineated using the Hough transform and boundary tracing-based techniques.
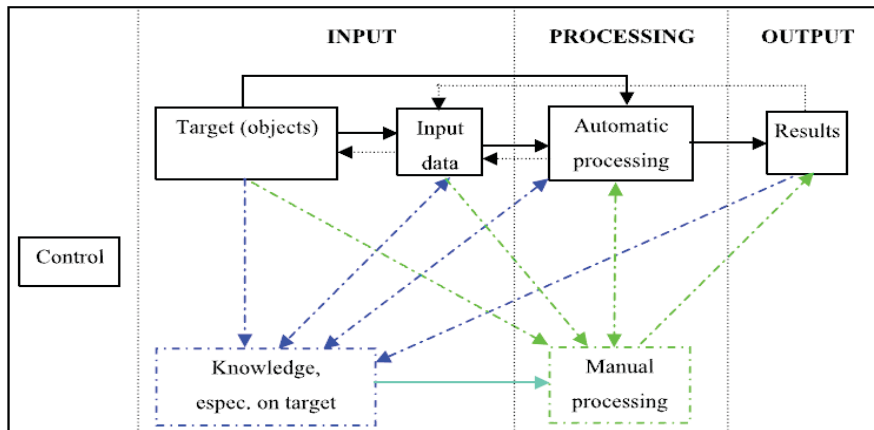

Fig. 1 Architectures and components of image analysis systems for object extraction [7]

Extraction and description of cultural (man-made features and objects, such as buildings and transportation networks were also a research topic in the past. The textural features such as densities, shape of the structures, image quality were analyzed. The methodology consisted of the following procedure: Detecting lines and corners, label corners based on shadows, trace object boundaries and verify hypotheses.

Since the shadow of an object also plays a vital role during the building extraction process, studying it was also expected. A computational technique for utilizing the relationship between shadows and man-made structures to aid in the automatic extraction of man-made structures from aerial imagery was studied. Four methods were described that performed the prediction of structure shape, grouping of related structures, verification of individual structures, and structure height estimation. In each method the relationship between structure and cast shadows was exploited in a unique fashion [8].

After the twentieth century the emergence of fuzzy logic was a widely accepted area of interest. An attempt was made to present an Object-based approach for urban land cover classification from high- resolution multispectral image data that builds upon a pixel-based fuzzy classification [9] approach. This combined pixel/object approach was demonstrated using pan-sharpened multispectral IKONOS imagery from dense urban areas.

The fuzzy pixel-based classifier utilized both spectral and spatial information to discriminate between spectrally similar Road and Building urban land cover classes.

Further the images were segmented and accordingly the non-building, non-road surface were eliminated. Using these techniques, the object-based classifier was able to identify Buildings, Impervious Surface, and Roads in dense urban areas with 76%, 81%, and 99% classification accuracies [5].

As the resolution quality of the satellite sensors upgraded, there arises a need for better quality performance tool for computer aided interpretation. Hence a system was designed for the detection and recognition of man-made objects in high resolution optical remote sensing images. Detection was done by finding a small rectangular area in the image containing an object. Recognition was the attribution of a class label [10]. Supervised learning approach based on support vector machines was used. The system would learn a generic model for each class of objects by using a geometric characterization of the examples in the database (SPOT 5 THR images, 2.5 m resolution). High number of geometric image features were utilized which allowed to characterize several classes of objects with different geometric properties using a supervised learning approach [3]. The results showed the possibility of discrimination of several classes of objects with classification rates higher than 80%.

III. DESIGN THEORY

The system consists of multiple stages which are initially being segregated and then these individual stages are designed, which are further pooled to obtain the final required output.

A. Image Preparation and Pre-Processing

The image used or selected is pan-sharped which undergoes pre-processing. The pre-processing includes grey scale conversion and enhancing the image to fulfil the characteristics of the input image. The pre-processing also includes thresholding at various levels.
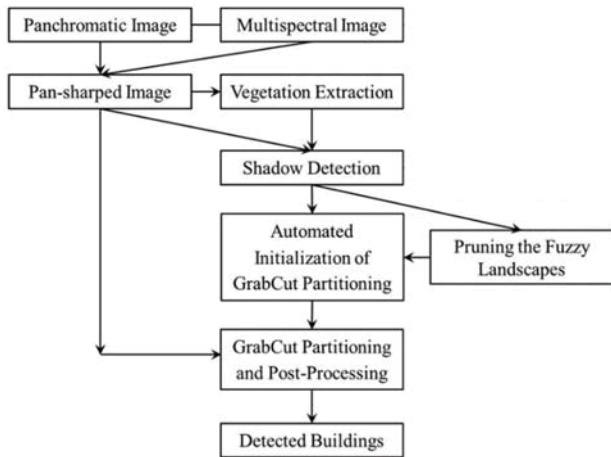
World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:12, No:10, 2018

Fig. 2 Building detection approach

B. Vegetation Extraction and Shadow Detection

For vegetation extraction NDVI (Normalized Difference Vegetation Index) is the widely accepted metric. By applying an appropriate threshold, we compute a binary vegetation mask.

$$NDV\,I = \frac{(NIR - R)}{(NIR + R)} \tag{1}$$

NIR and R represent the normalized near-infrared and red image bands.

For automatic shadow detection the multispectral false color shadow detection [11] is the convenient technique due to two reasons i) it utilizes advantage of near-infrared (NIR) image ii) it is fully independent of user and data-dependent thresholds.

Fig. 3 Color to HSI normalization

$$(Ratio\ Map)\ Rs = (S - I) / (S + I) \tag{2}$$

where, (S) - normalized saturation and (I) - normalized intensity.

C. Shadow Detection and Removal

A model for the spatial arrangement between shadow and building is designed using a morphological fuzzy relation. With reference to the object and a specified direction, the landscape around the reference object along the given direction can be defined as a fuzzy set of membership values in image space. The landscape membership values are defined in the range of 0 and 1. Fuzzy relation approach is used to determine the spatial arrangement between buildings and their shadows. Morphological characteristics information are utilized to find the exact relationship. With a reference (shadow) object and a direction specified by an angle α, the landscape around the reference object along the given direction is defined as a fuzzy set of membership values in image space.

In an urban area, it is essential for a building detection task to eliminate the landscapes that may occur due to shadows cast by non-building objects. To separate the landscapes of building and other non-building objects, the height difference of the objects compared to the terrain height is assessed. A minimum shadow length is computed, which is then compared with the perimeter pixels of a shadow object. If the length is found to be satisfying the length Lmin, an assumption is made that the shadow is cast from a non-building object, and thus, the generated fuzzy landscape is rejected [12].

D. Building Detection

Finally, the building and the non-building region need to be extracted. The Classical image segmentation tools use either texture (colour) information, e.g. Magic Wand, or edge (contrast) information, e.g. Intelligent Scissors.

The filtered image is then passed through GrabCut [13] label methodology, bw and rgb to extract the building structure. The resultant blobs are numbered which is then outlined to reconstruct the building structures.

IV. RESULTS AND DISCUSSION

Step 1: Pre-processing.

Fig. 4 Input Image

Fig. 5 Input Grey Image

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
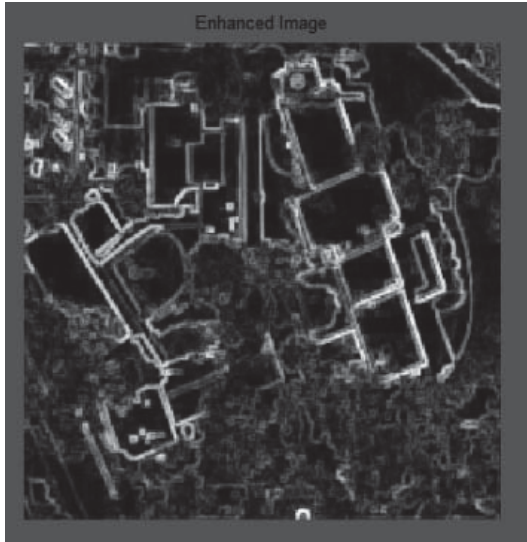Vol:12, No:10, 2018

Fig. 6 Enhanced Image

Fig. 7 Threshold, n=2

Fig. 8 Threshold, n=3

Fig. 9 Threshold, n=4

Above are the pre-processing stages. The process includes conversion of Input Image (the image that contains building structure) Fig. 4 to grayscale Fig. 5 which then acts as the input to the enhancement stage. Enhancement is done using the morphological dilation and erosion methodology, the Fig. 6 shows the enhanced image. The enhanced image then is made to undergo different levels of thresholding, Figs. 7-9 to provide an appropriate input for the Vegetation Extraction & Shadow Detection stage.

Step 2: Vegetation Extraction.
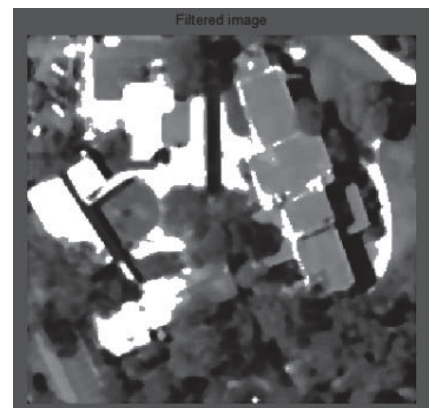
Fig. 10 Filtered Image

Fig. 11 Smoothened Image

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
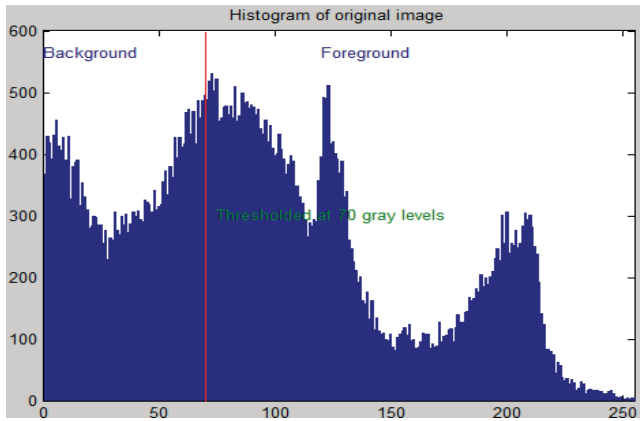Vol:12, No:10, 2018
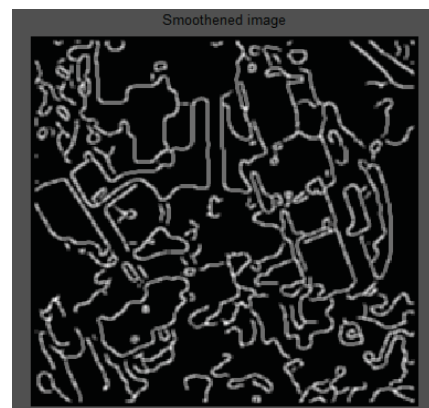
Fig. 12 Histogram: Background and Foreground

For vegetation extraction the NDVI is used. The filtered image Fig. 10 seen is that which is obtained based on the output of the pre-processing and the actual input image which undergoes the NDVI processing. This then is smoothened Fig. 11 to negate out the inducive noise and connected components. The histogram Fig. 12 manifests the background and foreground for the original image, which is thresholded at a gray level of 70.

Step 3: Shadow Detection and Extraction.

Fig. 13 Thinned Morphology Image

Fig. 14 Edged Image

Shadow detection and its extraction play a vital role in obtaining efficient and accurate output. The shadow here is similar to any unwanted noise (which needs to be taken care of) either in speech or image processing. In this stage of execution the image is filtered and thinned Fig. 13 to eliminate the noise followed by edge detection Fig. 14.
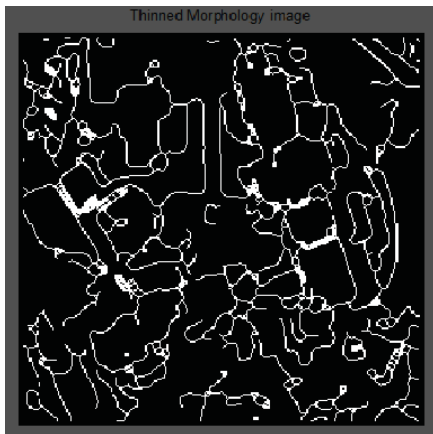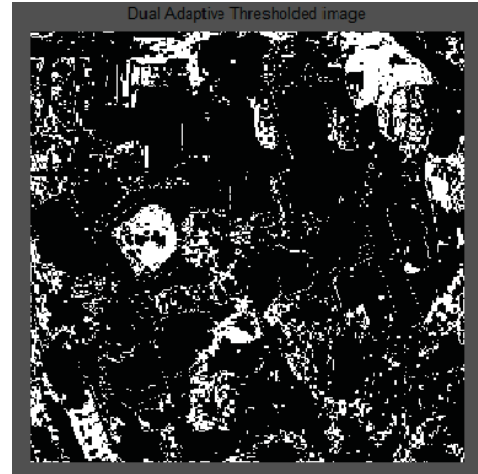

Fig. 15 Dual Adaptive Threshold Image

Fig. 16 Filed Holes

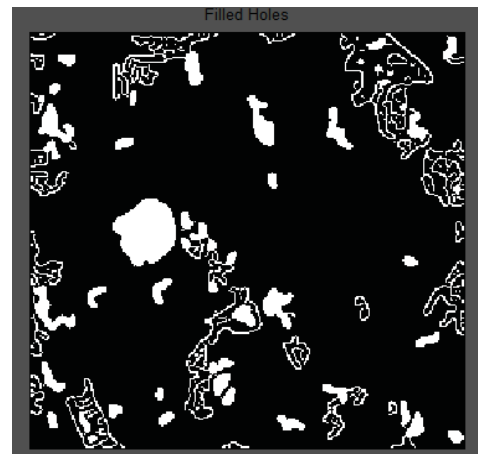Fig. 17 Dilated Image

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:12, No:10, 2018
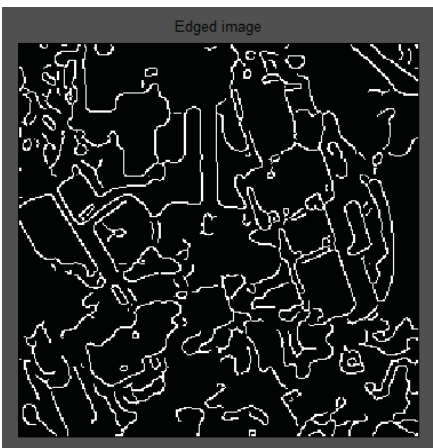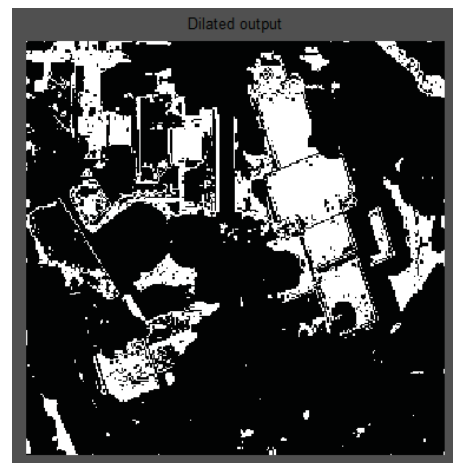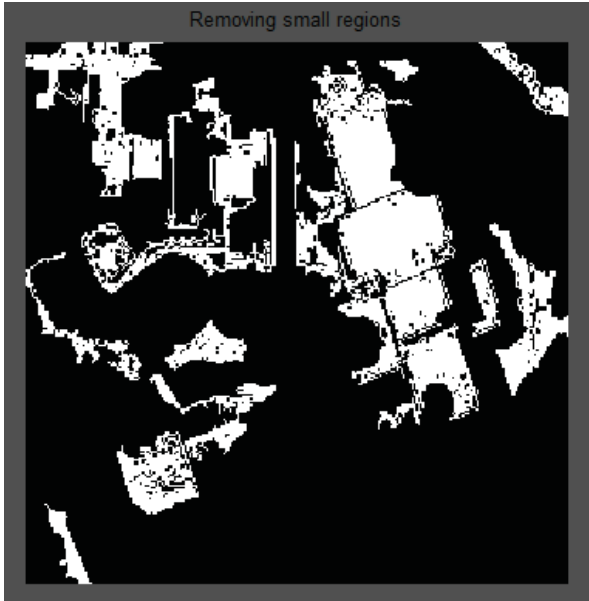
Fig. 18 Removing Small Regions

This edged image is then made to go through a dual adaptive threshold Fig. 15 process to detect the abnormal regions and eliminate them too. Once this is done, "hole fill" is initiated to get rid of any background pixels inside the blobs Fig. 16. The resultant image is then dilated Fig. 17 and the regions which are minuet Fig. 18 and merely useful are discarded so that the image has only those regions which have building structure in it.
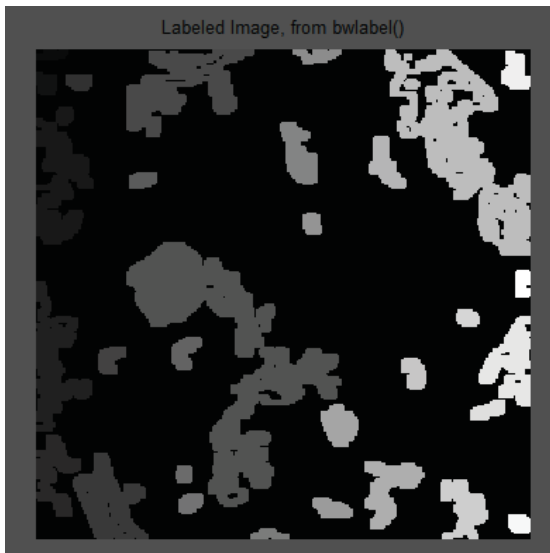
Step 4:   Building Detection.

Fig. 19 BW Labelled Image

The final stage involves partioning and post processing. The grab cut partioning utilizes the foreground and the background Fig. 12 along with the BW labeled Fig. 19 and pseudo colored labeled Fig. 20 to segment the landscape. The resultant image is then outlined Fig. 21 for the structure boundaries which at last is mapped over the original grayscale image to display the building region Fig. 22.
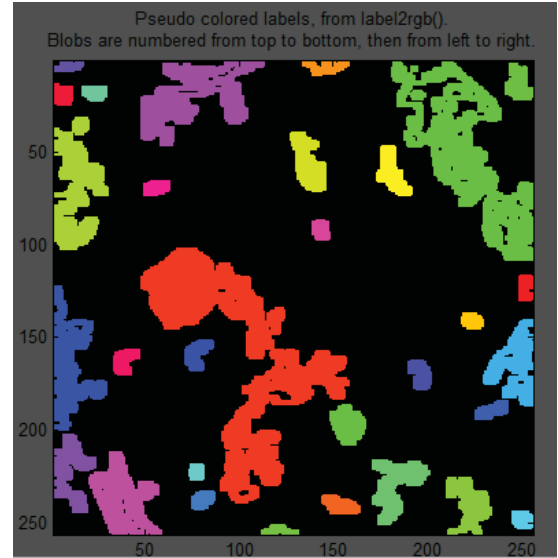

Fig. 20 Pseudo Coloured Labelled Image
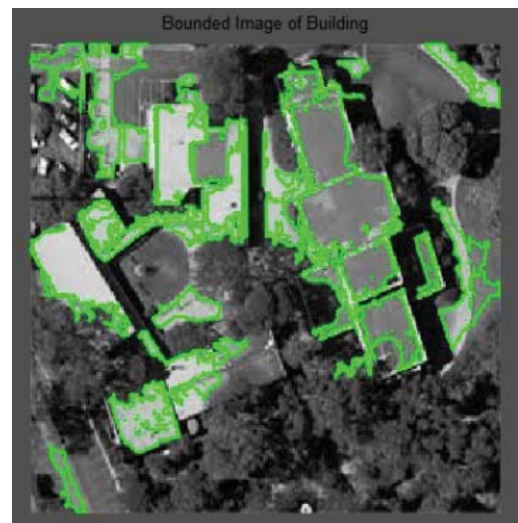

Fig. 21 Outlined Image

Fig. 22 Building Extracted Image

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:12, No:10, 2018

V. CONCLUSION

The automatic extraction of building is possible with,

- Higher accuracy.
- Least Processing time.

The performance of the above approach is affected majority by the shadow generation which seems to be a tentative drawback. Also, the non-building regions come into picture to impact the quality of the output.

In future the proposed method could be used to generate 3-D representation of the detected buildings. The problem of detecting buildings retains many complexities requiring substantial future research. The future scope would be to develop and integrate road and/or bridge detection with the current methodology to eliminate the superfluous land area which do not fall under the building category. By this way, most of the road segments that are erroneously labeled can be identified and eliminated. The other future scope would be to improve the boundary detection to enhance the output quality leading to a higher level of accuracy of the building by means of a generalization process. Additionally, there is also a possibility to reconstruct the detected buildings regions; therefore, as a final future work, there can be a plan to generate a 3-D representation of the detected buildings.

REFERENCES

[1] Ali Ozgun Ok, Caglar Senaras, and Baris Yuksel, "*Automated Detection of Arbitrarily Shaped Buildings in Complex Environments from Monocular VHR Optical Satellite Imagery*" IEEE transactions on Geoscience and Remote Sensing, vol. 51, no. 3, March 2013.

[2] Fischer, T. H. Kolbe, F. Lang, A. B. Cremers, W. Förstner, L. Plümer, and V. Steinhage, "*Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D*," Computer Vision Image Understand., Nov. 1998.

[3] H. G. Akçay and S. Aksoy, "*Building detection using directional spatial constraints*," in *Proc. IEEE IGARSS*, 2010.

[4] H. Mayer, "*Automatic object extraction from aerial imagery -- A survey focusing on buildings*," Computer Vision Image Understand., May 1999.

[5] Imdad Ali Rizvi and B.K.Mohan, "*Object-Based Image Analysis of High-Resolution Satellite Images Using Modified Cloud Basis Function Neural Network and Probabilistic Relaxation Labeling Process*" IEEE Trans. on Geoscience and Remote Sensing, vol. 49(12), Jan. 2011.

[6] E. Baltsavias, "*Object extraction and revision by image analysis using existing Geodata and knowledge: Current status and steps towards operational systems*," ISPRS J. Photogramm. Remote Sens., vol. 58, no. 3/4, pp. 129–151, Jan. 2004.

[7] D. Koc San, "*Approaches for Automatic Urban Building Extraction and Updating From High Resolution Satellite Imagery*," Ph.D. Thesis, Middle East Tech. Univ., Ankara, Turkey, 2009.

[8] R. B. Irvin and D. M. McKeown Jr., "*Methods for exploiting the relationship between buildings and their shadows in aerial imagery*," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, no. 6, pp. 1564–1575, Nov./Dec. 1989.

[9] A. Huertas and R. Nevatia, "*Detecting buildings in aerial images*," Computer Vision, Graph. Image Process, vol. 41, no. 2, pp. 131–152, Feb. 1988.

[10] J. G. Liu, "*Smoothing filter-based intensity modulation: A spectral preserve image fusion Technique for improving spatial details*," Int. J. Remote Sens., Jan. 2000.

[11] M. Teke, E. Ba¸seski, A. Ö. Ok, B. Yüksel, and Ç.¸Senaras, "*Multispectral false color shadow detection,*" in Photogrammetric Image Analysis, vol. 6952, U. Stilla, F. Rottensteiner, H. Mayer, B. Jutzi, and M.Butenuth, Eds. Berlin,Germany: Springer- Verlag, 2011, pp. 109–119.

[12] I. Bloch, "*Fuzzy relative position between objects in image processing: A morphological approach*," IEEE Trans. Pattern Anal. Mach. Intell., Jul. 1999.

[13] C. Rother, V. Kolmogorov, and A. Blake, "*GrabCut: Interactive foreground extraction using iterated graph cuts*," *ACM Trans. Graph. (TOG)*, 2004.