

Automatic Building Detection from Satellite Images using Internal Gray Variance and Digital Surface Model

Amit Raikar
P.G Student
P.D.A College of Engineering
Kalaburagi, Karnataka, India

Geeta Hanji
Associate Professor
P.D.A College of Engineering
Kalaburagi, Karnataka, India

ABSTRACT

Automatic building extraction is considered recently as an active research in remote sensing operation. It has been going on for more than 20 years but the automated extractions still encounter problems due to image resolution, variation and level of details. Because of high-object density and scene complexity this is going to be an even greater challenge especially in urban areas. This paper is going to present an ideal framework for high-resolution panchromatic images which helps in reliable and accurate building extraction operation. Proposed framework along with the consideration of domain knowledge (spatial and spectral characteristics) provides features like the nature of objects in the scene, their optical interactions and their impact on the resulting image. To analyze geometric nature of scene in better way we are using Digital Surface Model (DSM). Proposed algorithm has been evaluated using a variety of images from IKONOS and QuickBird satellites. The results demonstrate that the proposed algorithm is accurate and efficient in comparison with the state of art methods.

General Terms

Building detection, clustering, enhancement, feature extraction, high resolution, morphology, remote sensing, automatic detection, segmentation, and thinning.

Keywords

Digital Surface Model, DSM, Internal Gray Variance, and IGV.

1. INTRODUCTION

Satellite and aerial images are playing major role in acquiring information about objects on the Earth's surface. For many applications the main attention is to identify the objects and targets within the aerial images. Some example are rescue operations and defense applications. Since from the past human used to analyze the aerial image to recognize the building objects, and human understanding of these objects has become expensive and tends to be impractical because of quality of data and increased applications. In the study of aerial images, the identification of buildings and other man-made structures has become a common topic. Other applications such as creating maps or databases for geographic information systems, urban planning are also more demanding. The potential of identifying the building automatically and efficiently helps to understand the scene collected from the image contents and going to be used in database application like content-based retrieval. Further applications to be considered are planning of residential development, evaluation of damage and detection of military target. Because of basic geometrical nature of building it has

attracted most of the applications, which in turn going to decrease the effect of inter-building occlusion in aerial

images. The building detection in aerial image is considered to be a tough task because along with the building there are huge numbers of other objects like vegetation, water bodies, and roads. The potential for similarity of imaged roofs to a background is also another issue to be considered. The main objective here is to identifying the structure of object of interest and segmenting it from the background so that it can be represented for later understanding.

The primary operation to be considered here is identification and isolation which are difficult operations to perform, because of the presence of natural texture of vegetation, area occupied by water and other different kinds of elements which are generally present on or near the object of interest, specifically in the aerial image. Since from late 1980s, the identification of man-made structure and buildings has considered as the active field of interest. One of the standard method considered for solving the problem of identification is assuming that an object (or building) has four edges. Making this assumption, other solutions depend on parameters such as edge, line and corner detection. These parameters are used by grouping and achieving shape detection of objects like rectangles and parallelograms. Later in the second method, DSM module is used. It also helps in obtaining the height of the building from the first ground. DSM helps in extracting buildings, if there is no shadow present in the scene for the relevant building object. The combination of both IGV and DSM leads to detection of buildings.

2. LITERATURE SURVEY

Previously standard approaches considered as the probability model, in this approach spatial context parameters are able to increase the accuracy value of classification process. Additionally other sources of data have also been considered, such as range data or stereo vision by making use of more than one image of the scene for identification, with gray-scale image. Specifically considering the single image data, the information like height parameter is not available; this is the why multiple images were considered to deduce the height information. This height information is combined with spatial information in order to get better performance of detection techniques. In one of the method [17], the parameters like corner and edge values collected from specific families of polygon objects are generally going to be compared with corner and edge values that are specifically collected from images. Others have also achieved to collect and analyze image contents for building extraction by making use of morphological methods; the approaches make use of morphological filter values for identification of building. Few to mention are white top-hat, black top-hat or other geometrical filters. One of the technique considers shadow as case of elevated man-made structures; although this technique needs details like direction of the sun within the specific image, its angular elevation. Large numbers of methods for

identifying the building characteristics are reported in computer vision, remote sensing application and in photogrammetry [3]. Recent studies [4] have highlighted the work on extracting the building object details from standard high-resolution satellite imagery, and also considering digital elevation models (DEM) which are of high-quality and also considering the aerial images. The satellite images to be consider are QuickBird and IKONOS, has been used to obtain plane-metric building details with standard 1-m horizontal accuracy [5], [6]. For the study, we use the parameters like irregular structure of building and closeness of different buildings in urban areas; efforts are needed towards the research in the domain of standard automatic building extraction specifically from high-resolution aerial and extracting the details of the satellite images are implemented with the integration of LIDAR data and images [7]. considering high-quality DEM images such as those specifically available from standard LIDAR has been used to generate three-dimensional (3-D) building models [8], [9]. However the data collected from LIDAR is going to provide precise 3-D position values, the cost of collecting this data is high and it is inadequate to give surface texture value [7],[8].

The satellite images and common aerial images are high as well as low resolution satellite images considered as the basic source for acquisition of geospatial information which is precise and also reliable one. Muller and Zaum [10] have presented an algorithm for identification and performed classification of buildings, considering aerial images with help of region growing technique. Large number of techniques for extraction of building objects automatically from standard digital surface models (DSM) or multispectral imagery is discuss in [16] and [17]. Song and Shan [11] implemented a technique to identify buildings from color imagery which are generally high-resolution images. The authors have discussed on the ways to highlight the boundary of building and performed segmentation of buildings. Active contour models like snakes for identifying the boundary of buildings have been presented in [12]-[14]. Mayunga et al. [13] developed a semi-automatic building identification technique specifically taken from QuickBird images generally by selecting a point on the boundary of each particular building. Then, the initial curves of the specific prototype are provided and precise boundary values of corresponding buildings are identified by making use of an iterative function. For building identification from less or equal to 1-m/ pixel resolution imagery hardly few techniques have been developed, due to the challenges such as low signal to noise ratio and weak object signal.

In the past, to acquire 3D data, the DSM models have been used as input data for automatic and semi-automatic tools. DSM not only contains information like Digital Elevation Models (DEM), but also about the buildings and other objects higher than surrounding topographic surfaces such as trees. There are several methods using DSM to detect the buildings. Most of the studies [16], [17] determine the height so as to remove the non-building objects/structure. They focused on the polygon structure of the buildings and rooftop contours. Tournaire et al. [15] used point processes on digital elevation models. Brunn and Weinder [16] separated buildings and vegetation areas using height data and geometric information on DSM data. Abdullah H, Cem Unsalan, and Peter R. [17] proposed two novel methods to detect buildings by combining panchromatic and DSM data. The major drawback of DSM based scheme is that a group of trees may look like a building and there is no easy way to separate them. The work proposed through this paper attempts to address this limitations using

DSM module. The details of proposed work are presented in section 3.

3. DESCRIPTION OF THE PROPOSED WORK

The focus of this paper is the detection of buildings from high-resolution panchromatic images. An approach presented in this paper is designed to effectively extract the building features as outlined in Figure 1. The details of each step are presented as below.

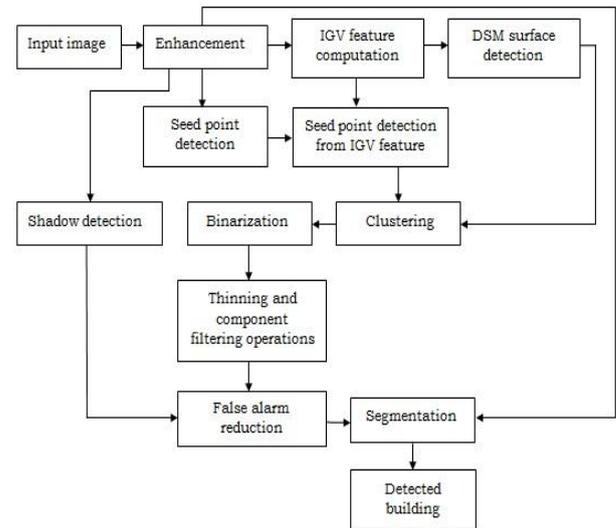


Fig 1: Block diagram of proposed work

3.1 Image Enhancement

The main goal of enhancement is to enhance the visual effects input image to improve the contrast between the target (buildings) and non-target (non-buildings) for high level processing. An enhancement operator will significantly increase (or reduce) the brightness of bright (or dark) building structures in the original image, however the non-target pixels are not affected. The key point is that a manmade object tends to be more homogeneous than that of natural objects. Therefore, morphological operators are useful to eliminate them. Since ‘Opening’ suppresses bright region and ‘Closing’ suppresses dark region, they are used in combination as morphological filters for image smoothing noise removal [2],[18].

3.2 IGV Feature Extraction

In manmade structure, the internal region is more homogeneous than the outer region. The variation in the internal pixels is low; the corresponding IGV is also low. At the same time, the IGV will be high at the border pixels. Whereas, in non-manmade structures the variation of the internal pixels and the external pixel is less, thus the internal pixels and external pixels of the natural structure are almost same. Application of standard edge operators to find the boundaries of buildings is unlikely to be successful, since the manmade objects are not well separated from natural objects. Thus to enhance the difference between them, it is proposed to use of the IGV feature. The enhancement technique tends to increase (or reduce) the brightness of bright (or dark) building structure and blurs the non-manmade objects such as trees. As a result, the variance of each pixel within an object is low, and is high at the boundary of the objects. In this way the IGV simplifies the identification of boundary points of the manmade objects. Since non manmade regions are not strongly different from the background in the enhanced

image, they are blurred. Thus, the enhancement and IGV feature extraction is very important in order to extract the buildings effectively. The IGV feature extraction is detailed as below.

IGV feature extraction includes calculating the average gray intensity within a 5×5 working window and then computing the IGV feature. The following steps explain the IGV feature extraction.

1. Average gray value of the pixels within a working window is computed as

$$\mu(x, y) = \frac{1}{w \times w} \sum_{i=-w/2}^{w/2} \sum_{j=-w/2}^{w/2} h(x+i, y+j) \quad (1)$$

2. IGV feature is computed as

$$IGV(x, y) = \sum_{i=-w/2}^{w/2} \sum_{j=-w/2}^{w/2} [h(x+i, y+j) - \mu(x, yz)]^2 \quad (2)$$

3.3 Digital Surface Model (DSM) Technique

The proposed method is based upon the detection of corner points and building shadows. In order to detect the corner points and building shadows, we have used two different digital surface models. In the DSM proposed in [16], for each of the local feature of the underlying image, two vector are generated; one for the dark shaded buildings and the other is for the lightly shaded buildings. Methods outlined in [16] fail to provide 100% building detection. Hence it is proposed to use another DSM [17] method to achieve maximum detection efficiency along with the method proposed in [16]. Description of DSM method is as below.

In the work presented through this paper, approach used for building detection exploits the fact that, the normalized DSM is the difference between DSM and Digital Elevation Model (DEM) describing the topographic surface. Normalized DSM also provides the information about buildings approximately referenced to a plane. Global threshold for segmentation yields proper output for the small vegetation group. The size criterion selection is not sufficient for larger vegetation area or vegetation area close to buildings [17]. To overcome this problem, information using Bayesian network classification is obtained which is an improvement over binary classification scheme. In the following lines description of these classification schemes is presented.

3.3.1 Binary Classification Scheme

The detection of vegetation areas within the DSM is an extension of the approach to building detection. In a scene the vegetation area can be easily recognized in the data sets of step edges and crease edges. These information's are used for a binary classification procedure. Along with the height information of the normalized DSM, the step and crease edge information is extracted and classified using the expected roughness of vegetation as threshold. The entire procedure determines the vegetation area.

The main drawback of this approach is the use of fixed thresholds. This drawback can be overcome by using the Bayesian networks for the classification which is described below.

3.3.2 Bayesian Networks for Classification

This approach uses the following features namely: the height information from the normalized DSM (that is 'ΔNh'), the step edge magnitudes (that is 'StepE') and the variances of

surface normal's (that is 'N Var')[17]. Pyramids of Random Variables (RV),

$$b_d = b_d(x, y, \mathbf{f}) \quad (3)$$

are generated based on these features. In the above equation x, y denotes the position, \mathbf{f} is determined as the used group of features f . The probability that a particular point is member of a building region in an image is denoted with probability $P(b_d(x, y) = T)$, and the probability $P(b_d(x, y) = F)$. Thus the probability of the complimentary event, is given by

$$P(b_d(x, y) = F) = 1 - P(b_d(x, y) = T) \quad (4)$$

Using height information of Normalized DSM ΔNh , the probability $P(b_d(x, y) = T)$, that is the point (x, y) is member of a building segment and is given as

$$P(b_d(x, y)=T) \propto$$

$$\sum_{(S_1, S_2)} P(b_d(x, y) = T | b_{d+1}(x, y) = S_1, b_d(x, y, \Delta Nh) = S_2)$$

$$P(b_{d+1}(x, y) = S_1) P(b_d(b_d(x, y, \Delta Nh) = S_2) \quad (5)$$

where the tuple (S_1, S_2) has the values $\{(T, T), (F, T), (T, F), (F, F)\}$. The term $P(b_{d+1}(x, y) = T)$ defines the posterior probability obtained from on the $(d+1)^{th}$ level of the pyramid, $P(b_d(x, y, \mathbf{f} = \Delta Nh))$ denotes the probability derived taking the feature on the d^{th} level into account, and $P(b_d(x, y) | b_{d+1}(x, y), b_d(x, y, \Delta Nh))$ is the conditional probability. The probability $P(b_d(x, y, \mathbf{f}) = T | \Theta)$ that a feature vector belongs to the region in the feature space a priori defined by 'Θ', Covariance matrix 'C' is calculated by

$$P(b(x, y)) = \frac{1}{\sqrt{2\pi det C}} \int_{\theta} \exp(-\frac{1}{2}(f - \theta)^T C^{-1}(f - \theta)) \quad (6)$$

3.4 Seed Point Detection Technique

There are different manmade structures like buildings, bridges, roads, and oil tanks present in the scene. The panchromatic remote sensing satellite image with limited spectral resolution has three characteristics, viz; 1) unknown number of material classes (nonparametric); 2) The brightness values of the classes overlap; 3) The variation of a particular material class may not be unimodal (non-Gaussian). Due to these properties it is difficult to correctly segment the original panchromatic image, especially manmade and non-manmade classes in different segments. To overcome this problem, the enhanced image is transformed into IGV feature space. Later is used a clustering-based segmentation technique to isolate manmade structures from the background.

Seed-based clustering methods start with some initial seed points and growing clusters around them. Two fundamental problems with all seed based techniques are the lack of well defined methods to select initial seed points, and inability of these techniques to deal with complex clusters. To overcome this problem multiseed technique-based clustering technique is used [19], [20].

The seed point technique, uses the information both from the enhanced image and IGV feature. Initially, the seed points are detected using multi seed technique [19], [20] of the enhanced image. Then the final seed points are detected using the IGV feature of the enhanced image from the algorithm explained below. These seed are called variance seeds. The multiseed

technique works effectively in the enhanced image rather than the IGV feature space, because the range of the feature space is much larger than that of the enhance image. If the multi seed technique is applied directly to IGV feature space, it is comparatively expensive due its occupancy of large data volume and the seed points are very close to each other due to small variation. As a result clusters are not formed properly. Therefore seed points from enhanced image are selected instead of the IGV feature images.

Let m_1, m_2, \dots, m_k be the 'k' seed points extracted from multiseed technique[19]. Each of these seed points $m_i, i = 1, 2, \dots, k$ is the mode of a particular homogeneous cluster. These seed points are used in the enhanced image to determine the seeds in the IGV feature space explained below.

- 1) Find the set of all pixels of enhanced image for each seed point $m_i, i = 1, 2, \dots, k$: whose gray values are m_i , i.e., $PX_i = \{(x_i, y_i) : g(x_i, y_i) = m_i\}, x = 1, 2, \dots, M, y = 1, 2, \dots, N$ and for all $i = 1, 2, \dots, k$, where M and N are the rows and columns of the image.
- 2) IGV value is formed at those pixels (x_i, y_i) in IGV feature space image, that is IGV $(x_i, y_i), i = 1, 2, \dots, k$ is obtained.
- 3) Calculate gray variance

$$GV[i] = \sum_{g(x_i, y_i) = m_i} (7)$$

$$IGV(x_i, y_i), i = 1, 2, \dots, k \quad (8)$$
- 4) Variance seed $VS[i] = GV [i] / \#PX_i, i = 1, 2, \dots, k$, is obtained where "#PX_i" is the number of points of the set PX_i .

3.5 Clustering

Applying multiseed clustering technique[19], [20] directly to enhanced image, many clustered regions are obtained and it is very challenging to identify building structures from the clustered image in unsupervised method. This problem can be overcome using a technique of nearest neighbor clustering method in the IGV feature space using detected seed points $VS[i], i = 1, 2, \dots, k$ are the k seed points (cluster centers) in the IGV feature space. In the IGV feature space each sample is grouped to the nearest cluster centre. This approach ensures that grouping of the boundary pixels of manmade structures into a single cluster are distinct from the clusters made by non-manmade features. Thus the detection of threshold value in the binarization process is quite easier and yields more accurate results for manmade structures.

3.6 Binarization

It is difficult to distinguish internal region and border region in the image as the manmade objects have natural brightness. In order to overcome this problem, the thresholding-based binarization technique is used [20] with the intension of mapping the points in the feature space back into the image space by binarizing the IGV clustered features. Though it is simple to use threshold-based binarization, detecting the threshold value automatically is difficult. In such a situation, a threshold on the bimodality detection has been proposed [20]. Explain of this scheme is given below.

Let, 'Pl' be the population of points in a cluster in the IGV feature space. 'Pl' is defined as bimodal, it can be divided into mainly two components or subpopulations, say $Pl_{s(u)}, Pl_{g(u)}$ such that,

- ' $Pl_{g(u)}$ ' contains all the pixels with cluster value \geq some 'u' and ' $Pl_{s(u)}$ ' contains all the pixels with cluster value $<$ 'u'.
- The variances of ' $Pl_{g(u)}$ ' and ' $Pl_{s(u)}$ ' are small relative to the variance of 'Pl'.

Let us assume that 'n' and ' σ^2 ' are the total frequency and variance of 'Pl', ' $n_{s(u)}$ ' and ' $\sigma_{s(u)}^2$ ' be the total frequency and variance of ' $Pl_{s(u)}$ ', and similarly ' $n_{g(u)}$ ' and ' $\sigma_{g(u)}^2$ ' be the total frequency and variance of ' $Pl_{g(u)}$ '. Now, the gray level value 'u' is determines such that the function

$$W(u) = \frac{n_{s(u)}\sigma_{s(u)}^2 + n_{g(u)}\sigma_{g(u)}^2}{n\sigma^2}, \quad cl_{min} \leq u \leq cl_{max} \quad (9)$$

is minimized. In the above equation cl_{min} and cl_{max} are the minimum and maximum gray level values of Pl. The gray level value 'u' is called bimodality parameter.

The cluster feature (cl_f) space is binarized and a binary image is generated by using the bimodality parameter 'u'. That is, for each pixel (x, y),

$$V(x, y) = \begin{cases} 0, & \text{if } cl_f(x, y) < u' \\ 1, & \text{if } cl_f(x, y) \geq u' \end{cases} \quad (10)$$

3.7 Thinning and Component Filtering Operation

In the image after the binarization process, many false alarms are present. These can be removed by using some prior knowledge [1]. Initially, for a particular region, if there are no building edges and shadows in the image region is removed. Very small regions are unlikely to be buildings, are also removed. To determine the accurate position of the building and its shadow, it is important to use the thinning process. The important shape factor to represent the object is perimeter and length. However it is difficult to estimate shape factor accurately. The boundaries of the manmade structure are highlighted with image binarization. However the boundaries obtained are thicker. To represent these thick boundaries linear/curvilinear representation is used in thinning process.

To reduce the thickness of the building some authors have used Sobel operator [21]. This operator reduces the threshold output of edge detector to the lines which are of single pixel thickness, but pixels at the extreme ends of lines are not affected. After thinning process, due to noise many small regions remain. These regions are removed using a simple filtering step based on the length. In this way, all thinned regions having lengths less than a predefined threshold (LT) are eliminated from further processing.

3.8 Shadow Detection Technique

The shadows can be clearly distinguished from the background in the scene, in case of satellite images. The image consists of many shadows of different objects such as buildings, towers, bridges. To detect the building the shadow information has been frequently used [1], [22], [23].

An algorithm to automatically detect and remove shadows in an image is proposed by Paul [22]. A method to detect the shadow using select correct features and parameters had been proposed by Liu et al. [23]. Proposed method uses shadow information is used to confirm that a detected object is a building. The important property of the shadow is its lower luminance [1], which determines the presence of shadow in the image. An enhanced image is used for the detection of shadows since they manifest as clear and darker regions in the image. The proposed shadow detection technique for panchromatic satellite images is a modified version of the threshold-based technique proposed by Otsu's method [24].

Otsu's method and the proposed (modified Otsu's) method is outlined as below.

1. Compute histogram and probabilities of each intensity level
2. Set up initial $\omega_i(0)$ and $\mu_i(0)$
3. Step through all possible thresholds $t = 1 \dots$ maximum intensity
 1. Update ω_i and μ_i
 2. Compute $\sigma^2_{b1}(t)$ and $\sigma^2_{b2}(t)$
4. Desired threshold corresponds to

$$\text{Desired threshold} = \frac{\sigma^2_{b1}(t) + \sigma^2_{b2}(t)}{2} \quad (11)$$

where $\omega_i(t)$ is class probability and $\mu_i(t)$ is class mean, $i = 0$ and 1 . $\sigma^2_{b1}(t)$ is the greater max and $\sigma^2_{b2}(t)$ is the greater or equal maximum.

3.9 False Alarm Reduction Technique

This step is used to reduce the false alarm obtained from the output of the thinning technique step. The technique uses the positional information of building edge and shadow. The position of the shadow with respect to the building depends on the position of sun and look angle of the satellite. At any instant of time the shadow of all buildings in a particular image is in the same direction, i.e., the position of shadow either may be right/left/above/below/diagonal. Using this information, the buildings are identified using following steps.

- 1) Remove such shadows that don't have edge in the edge map image in an appropriate position.
- 2) Identify at least one nonzero pixel of edge map image as the "building seed."
- 3) If there exists a building seed then eight-neighbor connected component of the edge map image for the seed is kept as it is; else, the edge component from the edge map image is deleted.

3.10 Segmentation

For effective segmentation, thresholding the target and background in the image is considered. Otsu method is used for segmentation [24]. The final building detection step considers an enhanced image. The segmentation process is based upon the local adaptive thresholding. Any local adaptive thresholding-based operation depends on the region size (that is the area of the image in which segmentation operation is to be performed). The output of false alarm reduction technique module is used. Each connected edge region of the final (high probable) building edge image is considered as a candidate for region thresholding. The adaptive threshold based segmentation procedure using the final building edge and enhanced images is described as below.

1. First, find four points, that is right-most, left-most, top-most, and bottom-most points of a particular connected edge region from the final building edge image. Based upon these four points, the bounding rectangle region is obtained and this rectangle is the region of interest for segmentation.

2. Find the corresponding rectangular region of the enhanced image.
3. Determine the threshold value (OT) by applying the Otsu's algorithm to the rectangular region of the enhanced image.
4. Segment the rectangular region of the enhanced image as

$$g(x,y) = \begin{cases} 0, & \text{if } h(x,y) < OT \\ 1, & \text{if } h(x,y) \geq OT \end{cases} \quad (12)$$

where the pixel (x, y) belongs to the rectangular region.

5. Repeat the above steps for all connected edge regions of the final building edge image.

The result of this step is the set of building structures detected in the image.

4. RESULTS AND DISCUSSIONS

In this paper, we have evaluated the performance of our approach with scenes obtained from IKONOS, QuickBird panchromatic, and Google maps imagery. We have also compared our results with other algorithms in literature that have used IKONOS, QuickBird and Google maps images.

Figure 2(a) shows an image of a satellite of size 240 x 210. We have used a 5 x 5 seed template for enhancement of the image. The enhanced image using the proposed algorithm is shown in Figure 2(b) and it is observed that the building structures are smooth and enhanced as compared to the original image [Figure 2(a)]. The algorithm for shadow detection shows a white color overlaid on the original image as shown in Figure 2(c). Figure 2(d) shows the clustered image by the proposed technique and it is observed that pixel intensity is low at the centre of the buildings and high at the border of buildings. It is also observed that the most of low IGV feature values are formed as a single cluster. Figure 2(e) shows the image using the binarization process. It is observed from Figure 2(e) that, most of the manmade structure boundaries are detected. Figure 2(f) shows the thinned image of the buildings. It shows the presence of thin curvilinear feature in the image. These features are eliminated as shown in the Figure 2(h). The corresponding image with edge details is shown in Figure 2(g) for edge extraction canny edge operator is used. It can be seen from Figure 2(g) that small unwanted edges of non-manmade structures are present and hence it is difficult to isolate the building edges from the non-building edges in many places. Thus, a qualitative analysis of the results shows that the proposed algorithm is successful in extracting the building edge features better than other edge operators. Figure 2(i) shows the final building edge image after false alarm reduction. Finally, the detected building structure (after segmentation) is shown in the Figure 2(j).

For an image of size 600 x 600 shown in Figure 3(a). The final detected building structure by the proposed method is shown in Figure 3(b). From the results it is observed that, though the original image is shadow free, the results obtained are much encouraging. Thus it can be said that, the proposed approach is quite efficient in detecting the maximum number of buildings in a scene without shadows.

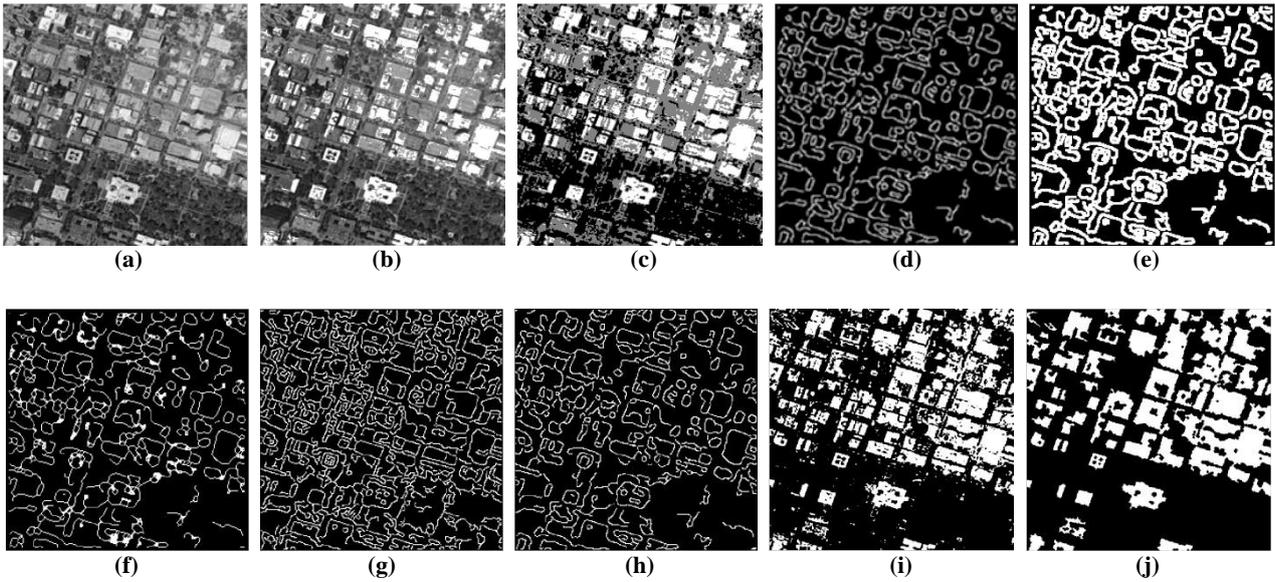


Fig: 2 (a) Original image. (b) Enhanced image. (c) Shadow overlade image. (d) Clustered image of the IGV features. (e) Binary image. (f) Thinned image. (g) Edge image by canny operator. (h) Small component filtered image. (i) Building edge image. (j) Final detected building structures.

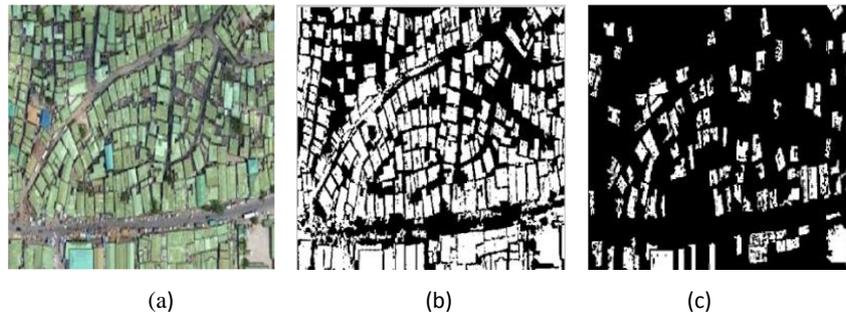


Fig: 3 (a) Original image of size 404 x 402. (b) Building detected by proposed method. (c) Buildings detected by R [1].

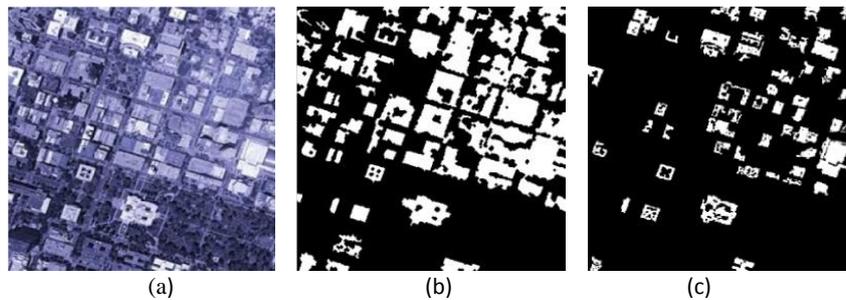


Fig: 4 (a) Original image of size 401 x 400. (b) Building detected by proposed method. (c) Buildings detected by R [1].

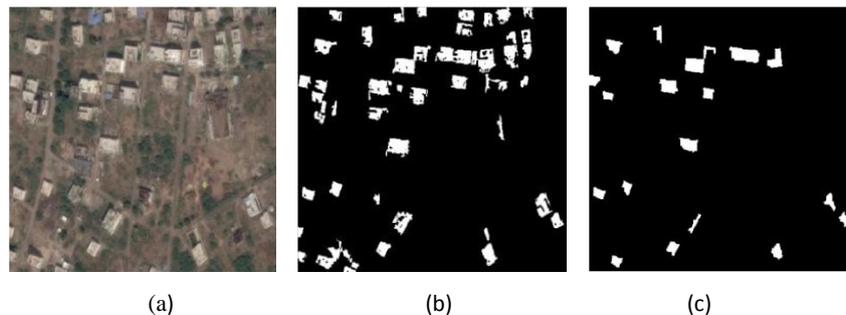


Fig: 5 (a) Original image of size 401 x 401. (b) Building detected by proposed method. (c) Buildings detected by R [1].

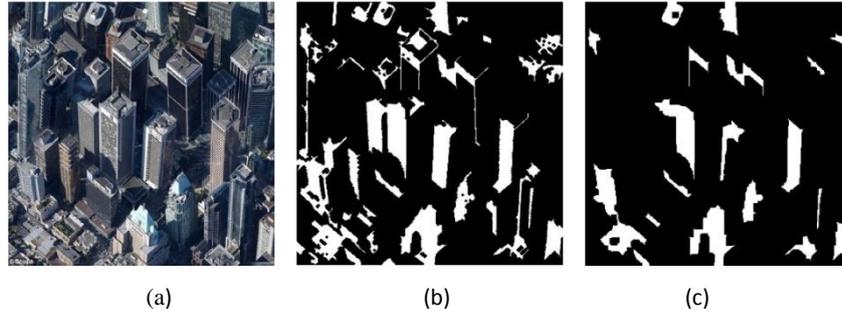


Fig: 6 (a) Original image of size 404 x 403. (b) Building detected by proposed method. (c) Buildings detected by R [1].

Figure 4(a) shows another panchromatic image without shadow. The building detected using the proposed method is shown in Figure 4(b). In this case also the results obtained very encouraging.

Figure 5(a) shows a Kalaburagi City image taken from the Google Maps source. Figure 5(b) shows the corresponding building detected by the proposed method. From the results it is observed proposed method works better even for a low resolution images. Observe that a low resolution image can also be used for our method. All the buildings in the image are detected.

Figure 6(a) shows the original image of the New York City. The image captured from the satellite is in the elevated form. That is the satellite makes an angle with the buildings. In this case also proposed method works very well. As shown in the Figure 6(b).

Finally, we calculate the performance of our method using a metric proposed by Lin and Nevatia [26]. Here, we compare the results with the ground truth derived manually. The metrics used are:

$$\text{Detection Percentage (DP)} = \frac{100 \times TP}{TP+TN} \quad (13)$$

$$\text{Branch Factor (BF)} = \frac{100 \times FP}{TP+FP} \quad (14)$$

where ‘TP (true positive)’ is the building detected manually and using the proposed algorithm, ‘FP (false positive)’ is building detected by proposed algorithm only, and ‘TN (true negative)’ is building detected only by manual method.

The building is said to be detected if a small part of it is detected by the proposed method. Here, we calculate the two metrics by comparing the output image with the ground truth. DP describes how many existing buildings in the image are found by the automatic approach, and BF indicates how many buildings are found erroneously. If DP is 100% which is true in case of ‘class image’; the BF also increases. Thus the goal is to maximize the DP while keeping the BF low. This requirement has been met in the proposed approach as shown in through the results tabulated in Table 1. Table 1 describes the values of TP, TN, FP, DP, and BF for the four different images. Performance results of DP and BF with proposed method and the existing method R [1] is shown in Figure 7. The mean detection percentage by the proposed method is more, while the mean branch factor is comparatively less.

Table 1: Evaluation Results for different satellite images

Image	Proposed Method					Existing method R[1]				
	TP	FP	TN	DP (%)	BF (%)	TP	FP	TN	DP (%)	BF (%)
Fig. 3	245	7	17	93.51	2.77	100	2	162	38.16	1.96
Fig. 4	63	2	9	87.5	3.07	41	2	31	56.94	4.65
Fig. 5	44	3	3	93.61	6.38	19	1	28	40.42	5
Fig. 6	32	2	8	80	5.88	24	0	16	60	0
Mean	DP = 88.65% BF = 4.52% (overall)					DP = 48.88% BF = 2.90%				

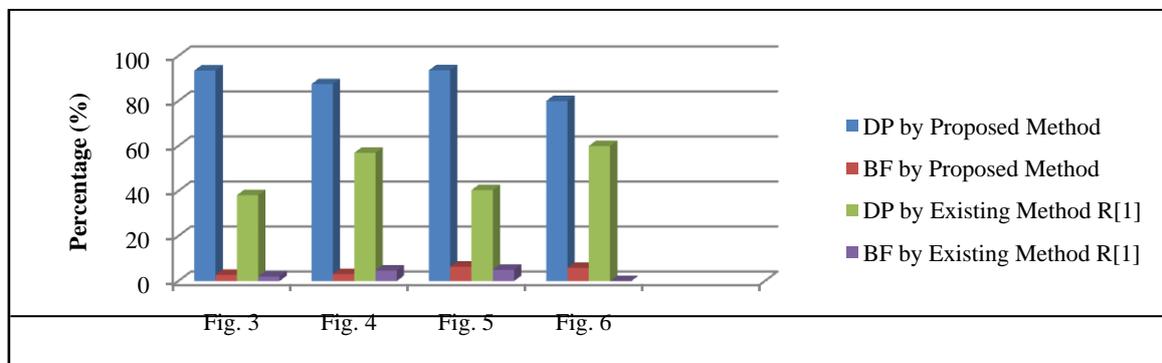


Fig 7: DP and BF Performance results obtained with proposed and the existing R [1] method.

5. CONCLUSION

Satellite imagery contains important information and it is also source of knowledge for many different applications. Hence, extracting the main characteristics from these images is a challenging task. Therefore, there is need of addressing different issues before performing operations like isolating the characteristics and identifying them. Building is considered as an important object in satellite image. Thus the building detection has a large number of applications. Few applications to mention are: civilian, commercial and military. This process is also useful in natural or manmade disaster, where the operations like finding out the existence of building can be performed. During the development of our work we found two issues, low signal-to-noise ratio and weak object signal in the specified images. Most of the previous algorithms discussed in this paper are not fully automated. This paper attempted to present fully automated technique to precisely identify the building from high-resolution panchromatic images which are remotely sensed. The method implemented provides the spectral and spatial characteristics by making use of multi-step mechanism. The results show that the proposed algorithm is very precise and effective in comparison with other algorithms reported in the literature.

6. ACKNOWLEDGE

Both the authors would like to express a deep gratitude to HOD (E&CE Dept.), Principal, and Nodal officer (TEQIP), P.D.A College of Engineering, Kalaburagi for providing platform for research work. Authors are also grateful to anonymous reviewers for providing valuable feedback to improve upon the quality of the research paper.

7. REFERENCES

- [1] D. Chaudhuri, Senior Member, IEEE, N. K. Kushwaha, A. Samal, Senior Member, IEEE, and R. C. Agarwal, "Automatic Building Detection From High-Resolution Satellite Images Based on Morphology and Internal Gray Variance" Manuscript received February 19, 2014; revised March 13, 2015; accepted April 09, 2015.
- [2] D. Chaudhuri and A. Samal, "An automatic bridge detection technique for multi-spectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2720–2727, Sep. 2008.
- [3] K. Khoshelham, C. Nardinocchi, E. Frontoni, A. Mancini, and P. Zingaretti, "Performance evaluation of automated approaches to building detection in multi-source aerial data," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, pp. 123–133, 2010.
- [4] Weidner, U. and Forstner, W., 1995. Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *ISPRS Journal* 50(4), pp. 38–49.
- [5] L. B. Theng, "Semi-automatic building extraction from satellite imagery," *J. Eng. Lett.*, vol. 13, no. 3, pp. EL_13_3_5, 2006.
- [6] D. S. Lee, J. Shan, and J. S. Bethel, "Class guided building extraction from IKONOS imagery," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 2, pp. 143–150, 2003.
- [7] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Using the Dempster-Shater method for the fusion of LIDAR data and multi-spectral images for building detection," *Inform. Fusion*, vol. 6, no. 4, pp. 283–300, 2005.
- [8] R. O. C. Tse, C. M. Gold, and D. Kidner, "A new approach to urban modeling based on LIDAR," in *Proc. Winter Sch. Comput. Sci. 14th Int. Cof. Cent. Eur. Comput. Graphics Visual. Comput. Vision*, Plzen, Czech Republic, Jan./Feb. 2006, pp. 279–286.
- [9] F. Tarsha-Kurdi, T. Landes, P. Grussenmeyer, and E. Smigiel, "New approach for automatic detection of buildings in airborne laser scanner data using first echo only," in *Proc. ISPRS Symp. Photogramm. Comput. Vision*, Bonn, Germany, Sep. 2006, pp. 1–6.
- [10] S. Muller and D. W. Zaum, "Robust building detection in aerial images," in *Proc. Joint Workshop ISPRS Ger. Assoc. Pattern Recog. Object Extr. 3-D City Models Road Databases Traffic Monit. Concepts Algorithms Eval.*, Vienna, Austria, 2005, vol. 36 (3/W 24), pp. 143–148.
- [11] Y. Song and J. Shan, "Building extraction from high-resolution color imagery based on edge flow driven active contour and JSEG," in *Proc. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Beijing, China, 2008, vol. XXXVII, pp. 185–190, Part B 3a.
- [12] J. Peng, D. Zhang, and Y. Liu, "An improved snake model for building detection from urban aerial images," *Pattern Recog. Lett.*, vol. 26, no. 5, pp. 587–595, 2005.
- [13] S. D. Mayunga, Y. Zhang, and D. J. Coleman, "Semi-automatic building extraction utilizing QuickBird imagery," *IAPRS*, vol. XXXVI, pp. 29–30, 2005, Part-3/W24.
- [14] S. Ahmady, H. Ebadi, M. J. V. Zouj, and H. A. Moghaddam, "Automatic building extraction from high-resolution aerial images using active contour model," in *Proc. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Beijing, China, 2008, vol. XXXVII, pp. 453–456, Part-B3b.
- [15] O. Tournaire, M. Bredif, D. Boldo, and M. Durupt, "An efficient stochastic approach for building footprint extraction from digital elevation models," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 4, pp. 317–327, 2010.
- [16] Abdullah H. Ozcan, Cem Unsalan, Peter Reinartz. "Building Detection Using Local Features and DSM Data" 978-1-4673-6396-9/13/\$31.00 ©2013 IEEE.
- [17] Brunn, A., Weidner, U., 1997. Extracting buildings from digital surface models. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences* 32 (Part34W2), 2734.
- [18] D. Chaudhuri, N. K. Kushwaha, and A. Samal, "Semi-automated road detection from high-resolution satellite images by directional morphological enhancement and segmentation techniques," in *Proc. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1538–1544, Oct. 2012.
- [19] D. Chaudhuri and B. B. Chaudhuri, "A novel multi-seed non-hierarchical data clustering technique," *IEEE Trans. Syst. Man Cybern., B*, vol. 27, no. 5, pp. 871–877, Sep. 1997.

- [20] D. Chaudhuri and A. Agrawal, "Split-and-merge procedure for image segmentation using bimodality detection approach," *Defence Sci. J.*, vol. 60, no. 3, pp. 290–301, 2010.
- [21] Irwin Sobel and Gary Feldman "Sobel operator or Sobel-Feldman operator", 1968.
- [22] D. M. Paul, "Shadow analysis in high-resolution satellite imagery of urban area," *Photogramm. Eng. Remote Sens.*, vol. 71, no. 2, pp. 169–177, 2005.
- [23] J. Liu, T. Fang, and D. Li, "Shadow detection in remotely sensed images based on self-adaptive feature selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5092–5103, Dec. 2011.
- [24] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
- [25] J. Canny, "A computational approach to edge detection," *IEEE Trans Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.
- [26] C. Lin and R. Nevatia, "Building detection and description from a single intensity image," *Comput. Vision Image Understanding*, vol. 72, no. 2, pp. 101–121, 1998.