Chapter 6

Building Extraction from Satellite Images

6.1 Introduction

Automatic building extraction is being extensively used for urban planning and management. For decades, the extraction of geospatial data process has been performed manually because of the inhomogeneities of the building structures. However, manual extraction is a very slow operation method and requires well-trained operators. The invested time and labour extremely increase the cost of operation. For this reason, building extraction using automatic techniques are developed. Extracting buildings is one of the most complex and challenging tasks as there exist a lot of inhomogeneity due to varying hierarchy. The variety of the type of buildings and also the shapes of rooftops are very inconstant. Also, in some areas, the buildings are placed irregularly or too close to each other. For these reasons, even by using high resolution satellite imagery, and considering soft computing methods and whatever be the algorithm formulated, the quality percentage of building extraction is very less. In our research work, two different approaches for building extraction are discussed to find out the best method to overcome these difficulties. One method is by considering region of interest segmentation and the other one is by considering region out of interest segmentation. Usually for all feature extraction problems, we will develop algorithms to extract the region of interest, but different methods using this concept is not to be ensured better efficiency for the case of
building extraction alone. Therefore, in our work, algorithms for masking region out of interest is also developed and compared with the other method to find out the best choice.

6.2 Existing Techniques

Building extraction systems can be categorized by the type of sensor data used. Some of the researches concentrate on the fusion of more than one data sources. These methods usually use the advantages of height information in 3D data set. One common approach is to use more than one aerial or satellite images and getting the height information using photogrammetric calculations (Henricsson, O., 1998; Fischer, A., 1998; Moons, T. et al., 1998; Koc San, D. & Turker, M., 2007). Also usage of new technologies such as LiDAR, which provides high vertical accuracy and high point density, becomes popular. Some of these make use of the fusion of LiDAR and satellite image (Sohn, G. & Dowman, J., 2003; Rottensteiner, F. & Briese, C., 2002). Another important category of the building extraction systems extracts objects by using monocular aerial or satellite imagery. Wei & Zhao (2005) introduced an approach, where they first cluster of the satellite image using an unsupervised learning method and used the shadow information to verify the existence of building. Then, for each building boundary, Canny operator is used for extracting edges and finally the system detects lines using hough transform. Mayungaa (2005) works on an active contour model which is commonly known as a snake algorithm for a semiautomatic building extraction method. In this method the user has to click to the approximate centre of each building; then the algorithm generates the border of this building. Jin & Davis (2005) proposed an automated building extraction strategy for high resolution satellite imagery that utilizes structural, contextual, and spectral information. The system runs automatically without pre-classification or any training sets, although some initial algorithm parameters must be set by the user.

Recent researches in this area focus on automatic and unsupervised extraction of buildings. Akçay & Aksoy (2008) proposed a method for unsupervised segmentation and object detection in high-resolution satellite images but the system performance varies depending on different rooftop structures. Aytekin & Erener (2009) proposed an algorithm for automatic and unsupervised building extraction.
from urban environments. Better performance is ensured by the method but the major drawback is over detection, *i.e.* detection of buildings at levels greater than that are actually the case. Use of local feature vectors and a probabilistic framework for the extraction of buildings having diverse characteristics and appearance is also discussed (Katartzis, A., 2007; Sirmacek, B., 2011). Though the method is efficient the algorithm is not strictly unsupervised. Recently, a novel approach for automatic detection of buildings with a gabled roof from very-high-resolution aerial images, covering particularly rural areas is proposed (Hazelhoff, L., 2011). Not many researches are carried out in this area. But the method is only semi-automatic and a number of assumptions are used to set the algorithm parameters which leads to the final result.

Another viable approach to building extraction is the use of artificial neural networks as proposed by Lari, Zahra. & Ebadi, Hamid. (2007). This approach must first be trained on a set of training images before being able to perform the extraction. This approach will initially perform segmentation on the image by using a seeded region growing algorithm. What this algorithm does is evenly distribute seed points across the entire image; it then compares the seed point’s value to neighboring pixels and adds them to the region if the neighboring pixels are below a set threshold. The region growing algorithm continues recursively with the newly added pixels until there are no more neighboring pixels which fall below the set threshold. This approach is able to extract building boundaries from the image, but since the resulting boundaries are a best fitted prior building shape, there might be slight variations from the actual boundary of the building.

Most of the works in the literature are either designed for specific applications or need some prior knowledge, such as human’s interaction for the extraction of buildings. Recent works which are focussed on unsupervised and automatic detection techniques are mostly restricted to specific types of shapes or surface features. A complex urban environment includes various shapes and surface materials which make the detection process complicated and in many cases pixels belonging to roof tops of buildings may wrongly identify as road pixels because both have linear features. Depending on roof top structure the system performance varies drastically and a single algorithm which is fully automatic and yet unsupervised which can be applied for any type of roof top structures with any complexity levels
is difficult. A solution for this problem is proposed in this paper by eliminating inhomogeneities due to varying hierarchy. Two different algorithms are developed focussing the region of interest having building features and region out of interest other than the areas having building features. The two different methods are evaluated with various qualitative and quantitative measures to conclude with the best method.

6.3 Study Region

All the test images used in this work are PS-MS IKONOS level 3A 1m resolution images of Ankara city, Turkey acquired in year 2007. The images are distributed by Space Imaging Middle East (SIME), Dubai, UAE (http://www.spaceimagingme.com) which is a regional affiliate of GeoEye, providing high-resolution satellite imagery of Middle East region collected from various earth observation satellites and aerial sensors, with resolutions varying from 15 centimeters to 20 meters. SIME’s satellite constellation includes Indian Remote Sensing satellite (IRS), IKONOS and CARTOSAT satellites. SIME offers full-resolution GeoEye-1 and IKONOS imagery products to download and view in standard GIS software applications. The test regions contain denser buildings with different kinds of challenges. Some buildings have different shapes of building rooftops. Also, some buildings have similar intensity reflectance to the roads and this causes interference of the roads and the buildings.

6.4 Method I – Region of Interest Segmentation

The method proposed in this work utilize the possibilities of Differential Morphological Profile-DMP (Jin & Davis, 2005) at the initial stage and training using artificial neural networks (Lari, Zahra. & Ebadi, Hamid., 2007) at the final stage. The pre-processing stage of the proposed method is the well known watershed segmentation using wavelet filters which is already explained in sections 5.4.1 & 5.4.2 of chapter 5 for road network extraction. So this part is not described in this chapter. The flow chart of the proposed method is given in fig.6.1. The test image used is a PS-MS IKONOS image converted to gray scale (fig.6.2 (a)) as the spectral information is not required in this method. The output of the preprocessing stage is given in fig. 6.2 (b). The remaining part of the proposed method is explained in the following subsections.
6.4.1 Extraction of Large Buildings

Most of the significant buildings can be detected using morphological opening and closing operations according to the shape characteristics from the reconstructed image. For this DMP with disc shaped Structuring Element (SE) having radius \( r = 3 \text{m} \) to \( 24 \text{m} \) (step size equal to \( 3 \text{m} \)) is used (Jin & Davis, 2005) as the structure of most of
the buildings varies from rectangular to round. DMP is a multi-scale image processing algorithm that employs a combination of morphological operators and derivatives of the resulting morphological profile. A multi-scale approach is required for two reasons: For the general case of object extraction, it is impossible to assume that all objects will always be of uniform size in a given image; thus, structuring elements of multiple scales are necessary. Additionally, even if this assumption is possible, a multi-scale technique also allows for the identification of object substructure. By varying the size of the structuring element used in the morphological operations, different regions corresponding to candidate objects will provide different levels of response. Objects with sizes varying from small to large will show an increased level of response as the size of the structuring element is changed correspondingly.

The DMP is divided into two profiles, opening morphological profile and closing morphological profile. Let $\gamma$ denotes morphological opening operator and $\varphi$ denotes closing operator, using structuring element $c$. The differential morphological profile $\Delta(x)$ can be written as eqn. 6.1.

$$
\Delta(x) = \left\{ \Delta_c : \begin{align*}
\Delta_c &= \Delta \varphi_{n-c+1}, \forall c \in [1, n] \\
\Delta_c &= \Delta \gamma_{c-n}, \forall c \in [n+1, 2n]
\end{align*} \right\}
$$

(6.1)

with $c = 1, \ldots, 2n$ where $n$ is the total number of iterations. Fig.6.3 and 6.4 shows derivative of the opening and closing profiles with different size of SEs and we can see the changes while SE is varied. The same have to be done for varying radius.

Figure 6.3: Derivative of the Opening Profile with (a) r=3 (b) r=6 (c) r=9 (d) r=12 (e) r=15 (f) r=18 (g) r=21 (h) r=24

Figure 6.4: Derivative of the Closing Profile with (a) r=3 (b) r=6 (c) r=9 (d) r=12 (e) r=15 (f) r=18 (g) r=21 (h) r=24
After these operations, there are still a lot of regions which are similar to building areas. The size of these areas is very small compared to buildings and can be eliminated by connected component labelling. To verify the hypothesized connected components as building regions, area analysis is performed. For that, the Minimum Enclosing Rectangle (MER) of each and every connected component suggested by Costa, L. da F. & Cesar, R. M. Jr. (2001) is used. Then rectangular fit is calculated as the area of the connected component divided by the area of its MER. If the rectangular fit is lower than a threshold, the connected component is rejected which reduces the percentage of over detection (6.2 & 6.3).

\[
\text{Rectangular Fit} = \frac{\text{Area of Connected component}}{\text{Area of MER}} \quad (6.2)
\]

\[
\text{IF} \ (\text{Rectangular Fit} \geq 0.5) \ \text{THEN ASSIGN Building};
\]
\[
\text{ELSE ASSIGN Not Building}; \quad (6.3)
\]

Fig. 6.5 & fig. 6.6 shows the result after calculating rectangular fit for the above shown DMP profiles. Thus we have investigated the presence of buildings with all opening profiles and all closing profiles using SE of size ranging from \( r=3 \) to 27. After adding all the extracted components we get the result as shown in fig. 6.7.

Figure 6.5: Result after rectangular fit for opening DMP with (a) \( r=3 \) (b) \( r=6 \) (c) \( r=9 \) (d) \( r=12 \) (e) \( r=15 \) (f) \( r=18 \) (g) \( r=21 \) (h) \( r=24 \)

Figure 6.6: Result after rectangular fit for closing DMP with (a) \( r=3 \) (b) \( r=6 \) (c) \( r=9 \) (d) \( r=12 \) (e) \( r=15 \) (f) \( r=18 \) (g) \( r=21 \) (h) \( r=24 \)
6.4.2 Neural Network based classification

After this initial segmentation, most of the significant buildings are extracted with less error. But some of the small building structures, bright buildings and building rooftops having irregular shapes are not extracted by differential morphological operations. For this a neural network is implemented with two different attributes for the classification process. The first attribute measures the linearity of the region boundaries, while the second attribute measures the roundness of the regions. Using these two attributes all other building shapes which lie in between rectangle and circle are extracted with high degree of accuracy.

Border linearity is measured using a modified version of hough transformation, (Koc San, D. & Turker, M., 2010). Roundness (Lari, Zahra. & Ebadi, Hamid., 2007) is independent of region’s size and calculated as ratio of area to square of perimeter (6.4). It varies from 0 to 1.

\[
\text{Roundness} = \frac{4 \pi \text{Area}}{\text{Perimeter}^2}
\]  

(6.4)

The neural network implemented in this research is a feed-forward back-propagation network (fig.6.8 (a)). The network consists of three layers; an input layer, one hidden layer, and an output layer. The number of neurons in the first layer is two. The optimum number of neurons in the hidden layer is selected to be ten. The number of neurons in the last layer is one. The output of this neuron is either one in case the region is a roof region or zero in case the region is not a roof region. The activation function for all neurons in the first and second layers is the sigmoid functions. For the
output neuron, the step function is chosen as the activation function. To study the performance of the neural network a variety of training datasets randomly selected from different regions of the input images having roof and non-roof samples with different shapes is chosen. As seen from fig.6.8 (b) after a certain number of epochs the performance approaches the goal which validates the network performance.

![Fig.6.8: (a) The Implemented Neural Network (b) Performance Evaluation](image)

To train the neural network used in this system, numeric features of the image calculated in the previous section are entered in network as inputs, network’s output is compared with desired output and network’s weight coefficients are modified using backpropagation algorithm. This process repeats until network’s output could classify roof and non-roof regions with desirable accuracy. Fig.6.9 shows the final building extraction result after neural network based classification.

![Figure 6.9: Final Building Extraction Result](image)

### 6.5 Method II – Region Out of Interest Segmentation

In this method a dedicated algorithm is developed for masking region out of interest so that region of interest i.e., buildings are projected in the final result. The algorithm
first calculates NDVI and chromaticity to intensity ratio for the initial level of segmentation. Next, rooftops and roads are detected and eliminated. Then principal component analysis and area analysis is done to get accurate results. The block diagram of the proposed system is shown in fig. 6.10.

![Block Diagram of the System](image)

**Figure 6.10: Simplified Block Diagram of the System**

### 6.5.1 Vegetation Masking

The NDVI is a simple graphical indicator to assess whether the target being observed contains live green vegetation or not which is well explained in section 2.2.11 of chapter 2. Nearly all satellite Vegetation Indices employ the difference formula given in eqn.2.1 to quantify the density of plant growth on the Earth. Calculations of NDVI for a given pixel always result in a number that ranges from minus one (-1) to plus one (+1); however, no green leaves gives a value close to zero. A zero means no vegetation and close to +1 (0.8 to 0.9) indicates the highest possible density of green leaves and thus the algorithm for vegetation masking is developed as given in eqn.6.6. Here also a suitable threshold (T) is determined by Otsu’s method (refer section 3.5.4).

\[
\text{IF } (\text{NDVI} \geq T) \text{ THEN ASSIGN Vegetation region;}
\]

\[
\text{ELSE ASSIGN Building region;}
\]  

(6.6)
6. Building Extraction from Satellite Images

The NDVI image can be obtained in MATLAB by reading the NIR data of the image from its .lan file generated using the software Multispec32 (refer section 2.2.8). Thus vegetation area can be eliminated by the comparison of obtained NDVI image and the original image. Regions are considered to be vegetation and the building hypotheses are rejected when the NDVI value is higher than the threshold (T). Fig.6.11 shows PS-MS IKONOS test image and corresponding vegetation masked image.

Figure 6.11: (a) IKONOS PS-MS Satellite Image (b) Vegetation Masked Image

6.5.2 Shadow Masking

The procedure for shadow detection is explained in detail on chapter 3. For the algorithm developed, we only require the RGB bands of the PS-MS image. After masking the detected shadow pixels, the figure shown below (fig. 6.12) is obtained.

Figure 6.12: Shadow Masked Image
6.5.3 Rooftop Detection

For the detection of rooftops a segmentation algorithm is required. Most of the segmentation algorithms work only on gray scale images but our image is a 4-band multispectral image. In order to retain the spectral features even after segmentation, a mean shift algorithm (Comaniciu, D. & Meer, P., 2002) is used. Here, the image from RGB space is first converted to LUV space in which \( L \) is the luminance value and \( U & V \) are the chrominance values. LUV color space is designed to be perceptually uniform and a given change in value corresponds roughly to the same perceptual difference over any part of the space. As for image segmentation, the aim is to cluster pixels sharing a similarity in pixel values. For this purpose, the filtering procedure is run and all convergence points are stored. The mean shift vector always points toward the direction of the maximum increase in the density. The mean shift procedure, obtained by successive computation of the mean shift vector \( m_h(x_t) \) and translation of the window \( x_{t+1} = x_t + m_h(x_t) \) is guaranteed to converge to a point where the gradient of density function is zero. Iterative calculation of mean shift vectors converges to a stationary point of the density, which corresponds to the modes of the image, i.e. homogenous structures in general. The pixel points converging to the same mode, are closer to each other in terms of spatial extend and color bandwidth. These pixels are segmented as the same cluster. In fact, mean shift vectors are aligned towards the similarity of colors incorporating spatial information. The algorithm for mean shift segmentation is given in fig.6.13. The mean shift segmented image is shown in fig.6.14.

6.5.4 Masking Long Segments

After segmentation, some noise is seen with the buildings in segmented regions. The edges of these regions are irregular. Hole filling followed by morphological opening operation is used to remove pseudo pixels and smooth building’s edge. A hole may be defined as a background region surrounded by a connected border of foreground pixels. Let \( A \) denote a set of points whose elements are 8-connected boundaries, each boundary encloses a background region (a hole). Given a point in each \( X_k \) defined by eqn.6.9 as a hole, the objective is to fill all the holes with ones.

\[
X_k = (X_{k-1} \oplus B) \cap A^c \quad k=1,2,3,......
\]  

(6.9)

Opening operation is of the form (6.10):
6. Building Extraction from Satellite Images

\[ X \circ B = (X \Theta B) \oplus B \quad (6.10) \]

That is, image \( X \) is eroded by structure element \( B \) (eqn.6.11), then it is dilated by \( B \). We define,

\[ B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (6.11) \]

1. Convert the image from RGB space to LUV space.
2. Select a data point (centre pixel) from the image randomly (not labelled).
3. Set the value of that point as the mean.
4. Find all the data points having the same value (a value within a specified range).
5. Cluster those points.
6. Select a point from the cluster.
7. Check whether the point is within a window or specified bandwidth (radius) from the centre point:
   - Yes-label it
   - No-no change
8. Go to step 5 until all the points in the cluster are evaluated.
9. Calculate the peaks of points having the same label.
10. Check whether the distance between the peaks of two labelled segments is less than or equal to bandwidth/2:
    - Yes-merge them
    - No-no change
11. Compute the mean shift vector of the clustered points \( m_h(x) \).
    \[
    m_h(x) = \frac{\sum_{i=1}^{n} x_i g \left( \frac{\|x-x_i\|^2}{h} \right)}{\sum_{i=1}^{n} g \left( \frac{\|x-x_i\|^2}{h} \right)} - x
    \]
    where \( h \) is the bandwidth parameter and \( g \) is the gradient density function.
12. Translate the window to the next point \( x^{t+1} \).
    \[
    x^{t+1} = x^t + m_h(x^t)
    \]
13. Check whether \( x^{t+1} = x^t \):
    - Yes-go to step 14
    - No-go to step 2
14. Check whether all the points are labelled:
    - Yes-end

Figure 6.13: Algorithm for Mean Shift Segmentation

Fig. 6.15 (a) and 6.15 (b) show the results of hole filling and morphological opening. After morphological treatment, edges have become smooth, but there are still some discrete noises. Many cases pixels belonging to rooftop of buildings may wrongly identify as road/ pavement pixels because both have linear features. To eliminate this, skeletonise the image (fig.6.16 (a)), as now our region of interest is main roads alone.
For skeletonization, label pixel p if and only if the rules 1, 2, 3, 4 are all satisfied.

Rule 1: The pixel under consideration must presently be black. If the pixel is already white, no action needs to be taken.

Rule 2: At least one of the pixel's close neighbours must be white.

Rule 3: The pixel must have more than one black neighbour.

Rule 4: A pixel cannot be removed if it results in its neighbours being disconnected.
After each iteration the labelled pixels are deleted. The algorithm is performed until one pixel wide skeleton is obtained.

For identifying whether a segment is building rooftop or not, find out the length of the segments which is equal to the number of pixels in skeleton. From the distribution of length of segments threshold is automatically estimated using Otsu's method. If length is greater than this threshold then surely it will not be a building rooftop and have to be masked. In the skeletonised image, there may be unwanted branches which are obviously non-building segments. So end points of the skeletons, which have only one neighbor, are removed (fig.6.16 (b)) for better result. Fig. 6.17 is the resultant image after these operations.

![Figure 6.16: (a) Skeletonized Image (b) Skeletonized Image with End Points Removed](image)

![Figure 6.17: Image after Masking Long Segments](image)
6.5.5 Masking Very Thin Segments

Some small objects such as cars, trees on the side of roads can also be viewed as thin stripes after skeletonization. This can be eliminated by applying, Principal Component Analysis (PCA) to each segment. Very thin segments show large variances along the first principal component whereas small variance along the second principle component. Therefore, the ratio of the corresponding eigen values provides the variances along the corresponding eigenvectors and gives an idea of how thin the segment is. Higher ratios represent unreasonably thin segments. Then, this ratio is thresholded in order to detect whether the segment is road or not. Here also, the threshold is automatically determined by Otsu’s method (refer section 3.5.4).

PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables, called principal components. The number of principal components is less than or equal to the number of original variables. PCA is the simplest of the true eigenvector-based multivariate analysis. PCA is done by calculating the mean of each co-ordinate X and Y of the segment and then finding the covariance matrix. Covariance matrix for a set of data with n dimensions is given by eqn.6.12 and eqn. 6.13.

\[ \text{cov}(X,Y) = \frac{\sum_{i=1}^{n} (X_i - \bar{X}) (Y_i - \bar{Y})}{(n-1)} \]  

\[ C_{\text{cov}} = \text{cov}(\text{Dim}_i, \text{Dim}_j) \]  

(6.12)

(6.13)

Then find the eigen values \( \lambda_1, \lambda_2 \) of \( C_x \) using eqn. 6.14.

\[ |C_x - \lambda I| = 0 \]  

(6.14)

The eigen vector with the highest eigen value is the principal component. The eigen vectors are ordered by eigen value from highest to lowest. This gives the components in the order of significance. Form the feature vector as eqn. 6.15.

\[ \text{Feature vector} = (\text{eig}1 \ \text{eig}2 \ \text{eig}3 \ ... \ \text{eigin}) \]  

(6.15)

The final data is derived by the following equation (6.16):

\[ \]
Final data = Row feature vector x Row Data Adjust

Row Feature Vector is the matrix with the eigenvectors in the columns transposed so that eigen vectors are now in the rows, with the most significant eigenvector at the top. The data items are in each column, with each row holding a separate dimension.

6.5.6 Area Analysis

To eliminate very small structures wrongly identified as buildings, area analysis is performed by finding rectangular fit in the same way as discussed in section 6.3.1. The results of PCA and area analysis for the PS-MS image are shown in fig.6.18.

6.6 Results and Discussions

The extracted buildings using both methods are compared with the manually labelled buildings as reference (fig.6.19). Within the test area, 49 buildings were manually delineated. As seen from the output image, using method-I only 38 buildings are correctly detected; 6 buildings are wrongly identified. Using method-II, all the 49 buildings are detected but many of the scattered points can also be considered as buildings.

Figure 6.18: (a) Image after PCA (b) Image after Area Analysis

Both methods are applied in two more PS-MS IKONOS test images shown in fig. 6.20. Fig. 6.21 and fig.6.22 show the building extraction results using both methods compared with the manually detected reference for the test images.
Figure 6.19 (a) Extracted Buildings considering Regions of Interest (b) Extracted Buildings by Masking Regions out of Interest (c) Manually Labelled Buildings

Figure 6.20: (a) IKONOS PS-MS Satellite Image (1m) (b) IKONOS PS-MS Satellite Image (1m)

Figure 6.21: Building Extraction Results of Fig.6.19 (a): (a) Considering Regions of Interest (b) Masking Regions out of Interest (c) Manually Labelled Buildings
6. Building Extraction from Satellite Images

In method-I, building extraction is done only after implementing a segmentation algorithm with wavelet and watershed transforms. Therefore, buildings even from the high dense areas can be detected with considerable degree of accuracy. Also, by incorporating neural networks with differential morphological profiles, insignificant buildings which cannot be distinguished from the background features are able to extract. Though the variety of the type of buildings and shapes of rooftops makes the building extraction complicated, acceptable accuracy for the extracted buildings is obtained. However, the output image may not provide the exact shape of the building structures. Also some non-building structures are wrongly identified as buildings.

In method – II, we have to detect even the irrelevant elements other than buildings in the scene and have to find a solution to mask it which seems to be a tedious procedure. But as seen from the results obtained from both methods and manually labelled reference in fig. 6.19, fig. 6.21 and fig. 6.22 it is concluded that buildings detected using method – II outperforms compared to that of method – I for many reasons. For the image shown in fig. 6.20(b), buildings are not only placed closely but also having no specific shape for most of the buildings where the neural network algorithm using method-I is challenged. When compared to method – I the detection rate of method – II is very much higher and provides exact shape of
buildings. Building results obtained using method – II is very close to the manual reference image. The above inferences are mathematically verified and given in table 6.1.

For performance evaluation, we use the evaluation measures widely accepted for building extraction (Shufelt, J. A., 1999; Lee, D. S. et al., 2003). The extracted buildings and the manually detected buildings are compared pixel-by-pixel. The pixels in the image are categorized into four types:

(1) True positive (TP): Both manual and unsupervised methods label the pixel belonging to the buildings.

(2) True negative (TN): Both manual and unsupervised methods label the pixel belonging to the background.

(3) False positive (FP): The unsupervised method incorrectly labels the pixel as belonging to a building.

(4) False negative (FN): The unsupervised method does not correctly label a pixel belonging to a building.

Based on these categories the system performance is evaluated using the following measures:

\[
\text{Branching Factor} = \frac{FP}{TP} \quad (6.17)
\]

\[
\text{Miss Factor} = \frac{FN}{TP} \quad (6.18)
\]

\[
\text{Detection Percentage} = 100 \times \frac{TP}{TP + FN} \quad (6.19)
\]

\[
\text{Quality Percentage} = 100 \times \frac{TP}{TP + FP + FN} \quad (6.20)
\]

The detection percentage is the percentage of building pixels correctly labelled by the proposed method. The quality percentage measures the quality of the extraction process. Performance evaluations for three different test images of various complexity levels are shown in table 6.1.
The results of the performance evaluation for all the test images is higher for method-II compared to method - I. In method – II, the detection percentage reaches 95.3% which is very high compared to the conventional method of building detection with algorithms based on region of interest segmentation.

In method-II, masking is performed on areas other than the region of interest which helped to achieve improved performance of the algorithm. It became possible to incorporate both spatial and spectral properties of the image through the use of mean shift segmentation. Utilizing statistical properties of the image with the help of principal component analysis, it is possible to achieve better adaptability for different complexity levels of urban areas. Finally extensive analysis with respect to area is done, which removes most of the unwanted pixels classified as buildings. In this approach, building extraction is completely automatic and unsupervised. Here we only provide an input which is a pan-sharpened multispectral image and we get an output which is the candidate buildings.

From the analysis it can be concluded that masking maximum unwanted components using multispectral images will give very good results for building extraction, instead of developing methods focusing only on region of interest having buildings alone.

### Table 6.1: Performance Evaluation of Extracted Buildings

<table>
<thead>
<tr>
<th>Reference Images</th>
<th>Branching factor</th>
<th>Miss factor</th>
<th>Detection percentage</th>
<th>Quality percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Method-I</td>
<td>Method-II</td>
<td>Method-I</td>
<td>Method-II</td>
</tr>
<tr>
<td>Fig. 6.2 (a)</td>
<td>0.36</td>
<td>0.12</td>
<td>0.31</td>
<td>0.09</td>
</tr>
<tr>
<td>Fig. 6.19 (a)</td>
<td>0.33</td>
<td>0.24</td>
<td>0.37</td>
<td>0.31</td>
</tr>
<tr>
<td>Fig. 6.19 (b)</td>
<td>0.31</td>
<td>0.38</td>
<td>0.51</td>
<td>0.47</td>
</tr>
</tbody>
</table>
6.7 Summary

Two methods of building detection algorithms are discussed in this chapter – One by considering Region of Interest and the other by considering Region Out of Interest. For each method a dedicated algorithm is developed and discussed the merits and demerits. Performance evaluation is done for both methods by comparing with the manually labelled building reference.