

Automatic Building Extraction using a distance from a variant mean Classification and Hough transform

A. Shaker

A.shaker@feng.bu.edu.eg

Mahmoud M.M. Hasan

mahmoud.hassan@feng.bu.edu.eg

Surveying Department
Shoubra Faculty of Engineering
Benha University

1. Abstract

Automatic building extraction from digital ortho - photos remains an open research area in digital photogrammetry. This paper proposes a sequence of classification and linearization processes for the extraction of building features with different rooftops from high resolution Multispectral satellite images (e.g., IKONOS and Quickbird) in Middle East countries. In this paper the minimum distance from a variant mean classification followed by a linearization approach using Hough transform is used. The variant mean is performed using a part of the image as a training data set for each building. This training set is iteratively updated at each process by adding a new color data to the training data set. A linearization process is then performed using the ability of Hough transform to detect straight lines.

The software has been designed and developed in Matlab programming environment and designed to be a user friendly since little interaction is required from the users is required. The proposed system works in four different phases: the first phase is the user interaction phase, the second phase is the colors classification phase, the third phase is Hough transform phase, and the fourth phase is the coordinate classification and polygons generalization phase. An image from IKONOS sensor is used for testing this system. The extraction results are compared with manually digitized ones. The comparison illustrates the efficiency of the proposed algorithm in which it can extract approximately 80% of buildings in the image properly.

2. INTRODUCTION

Since manual extraction of buildings from imagery is a very slow process, automated methods have been proposed to improve the speed and utility for urban map's production process.

Automatic building extraction has been an active research topic in the field of digital photogrammetry and computer vision for many years. Some useful applications are automation of information extraction from images and updating of geographic information system (GIS) databases. A wide range of techniques and algorithms have been proposed for automatically constructing 2D or 3D building models from satellite and aerial imagery.

Most of the recent work on building extraction from high resolution satellite images is based on supervised techniques. These techniques either require a classification based on initial training data to provide hypotheses for the positions and sizes of the candidate building features [1,6,7], or they use training sets or model databases to classify or match the buildings[8,9].

In this paper, an automated system for extraction of buildings from high-resolution satellite imagery is proposed it is based on the supervised classification (using training data) interactively. The key feature of this paper is the user friendly feature since it needs little interaction from the user to choose the approximate buildings polygons. The area inside these polygons will be the training data for their respective buildings and will not be processed. The only part will be processed is the buffering area around these points not the hole image as we will see in item 4. The chosen points will serve also as a training set when classifying the edge points of the detected buildings. In other words which points on the edges will belong to which side on the buildings?. For this reason, the choice of the points on any building should nearly represents the building outer shape.

This system is different than the system introduced at [1] in two process. The first difference is in the color classification process. It is based here on measuring the distance of each image pixels from two different classes (building and background) rather than several classes in the first system. At the same time, the distance used as a classifier here is computed from a variant mean rather than one mean in the first system. The use of a variant mean eliminate the existence of misclassified points because if their distances from both classes mean is equal in one iteration , it will be close to either means in the subsequent iterations. The second difference is on the straight lines forming the building polygons detection. The first system depends on detecting the points that form straight lines using some geometric criteria mentioned at [1], eliminate the nodes presented at these lines, then restrict the lines within a predefined length and azimuth to follow a predefined azimuths (45° or 90°) to their nearest neighbor lines . In other words, it is a point by point detection process. While in this paper the Hough transform is used to detect some possible lines on the edge of each building. These lines' pointes are subjected to a classification process to determine to which line on the building polygon these lines are belonging. Connecting the lines nearly belong to one straight line. The intersection of the resulting straight lines will then form the building polygon.

The proposed automatic system consists of four main p phases as follows:

1. **User interaction phase:** At which the user will choose polygons around the buildings as will described later. The RGB colors inside each building polygon will represent the training data set for this respective building. A buffering area around this polygon by a pre-specified width will be formed and the RGB colors inside this area will be the training set of the background class.
2. **Colors Classification Phase:** In this phase the mean of two discriminate classes' colors is computed. The distance of each pixel in the background class from each class mean will give the pixel the class number it is closest to(1 for buildings and 0 for background). These two means are variant i.e. they are updated iteratively as

will be discussed later. The classified buildings are then subjected to clearing boarder operation and edging.

3. **Hough Transform Phase:** At which some straight lines are detected on the edge of each polygon using Hough transform
4. **Coordinate classification and polygons generalization phase:** At this stage the coordinates of the Hough lines are classified according to the order of the clockwise polygon represent each polygon. The points that follow a straight line equation to a certain threshold are then fitted to straight line. The intersection points of these lines are then computed and will be the least number of points represent the shape of the polygon

The detected polygons areas are then compared to the digitized ones to asses the accuracy of the detection accuracy.

The algorithm is tested on an IKONOS image ‘assumed to be an orthogonal’ as a test image to clarify the effectiveness of the algorithm to extract straight lines edged buildings

3. Methodology

3.1 User interaction phase

At this phase the user should click on the points that best represent the shape of the buildings. This step is of crucial importance because the classification of the detected building outer lines points (phase four) will depends on it. For this reason the points chosen should not randomly chosen they should follow either a clockwise or anti-clock wise order. This procedure is a sequential procedure The program will keep prompting with "**Do you want to choose another building**" if the answer is yes the site will be displayed and the user should click on the points that best represent the potential building shape until an answer with 'n' or 'no' will terminate the user interaction and the program will start processing the remaining phases for each polygon separately.

Getting the approximate building polygons a buffering area around the selected polygons by a fixed distance (chosen 30 pixels) will be formed. **If** any line in the buffering area polygon intersects the image frame, the intersection points of these lines with the image frame will included in the buffering polygons not the end points of these lines. Figure (1) shows the site and the chosen polygons displayed on it. Because of the running speed the chosen polygons are chosen near the building polygon but the procedure will be the same if it is not so close, but the running time will be increased.

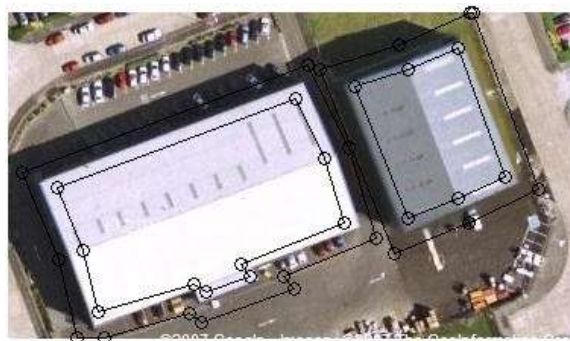


Figure 1: a. Initially chosen polygons and buffer polygons

3.2 Colors Classification Phase

The trend followed here is rather than dividing the image to several classes and get some sample training sets for each class. Formerly, this was done either by clicking the mouse on some color samples representing the discriminate colors for each class or taking a signature around some parts of each building. This process is tedious and time consuming.

In this paper only two discriminate classes will be used, namely the building class and the background class. Two mask images will be formed with the two above mentioned polygons. Figure 2 shows the two black and white images representing the two masks used for one of the buildings in the site seen. The training sets for the building class will be all the RGB colors under the mask formed (Figure 2a) is used as a preliminary training data for this building and the RGB color ranges under the second mask image (figure 2b) are used as a training data for background even if they contain color data for the building itself.

The mean for both the building and background colors are computed and used as initial two means classifiers for both the building and background. Through an iterative process the distance from each point under the background mask images (not the whole image) are computed. Any pixel takes either 1 or zero according to which mean its distance is closest (1 for building and 0 for background class). A new set of training set is then formed using the previous set and the new points RGB colors classified as building class. This process is repeated until the number of building class points becomes zero. At this instant the remaining colors in the buffer zone will only represent the background colors. Figure 3a & b shows the results of the building and background classes for the two buildings at each iteration ordered from the upper left to lower right. It shows that the left building takes six iterations to be completed while the second one takes five iterations.

The last stage at this phase is the clear boarder and area threshold operations. The purpose of these two operations is to set the zero pixels at boarder of the building takes the label of the majority of the neighborhood pixels and removes detected objects less in area than a given threshold (taken 200 pixels). Figure 4 a & b shows the resulting black and white images representing the shape detected for each building after these two processes.

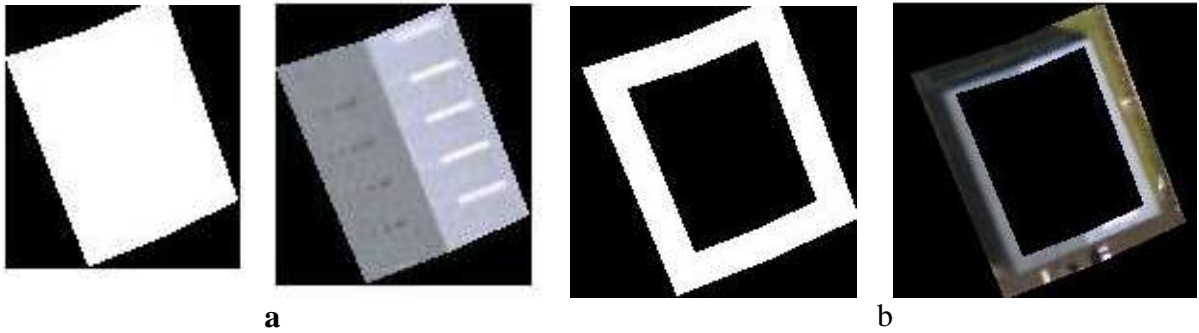


Figure 2: a. Building mask image, b. Background mask image

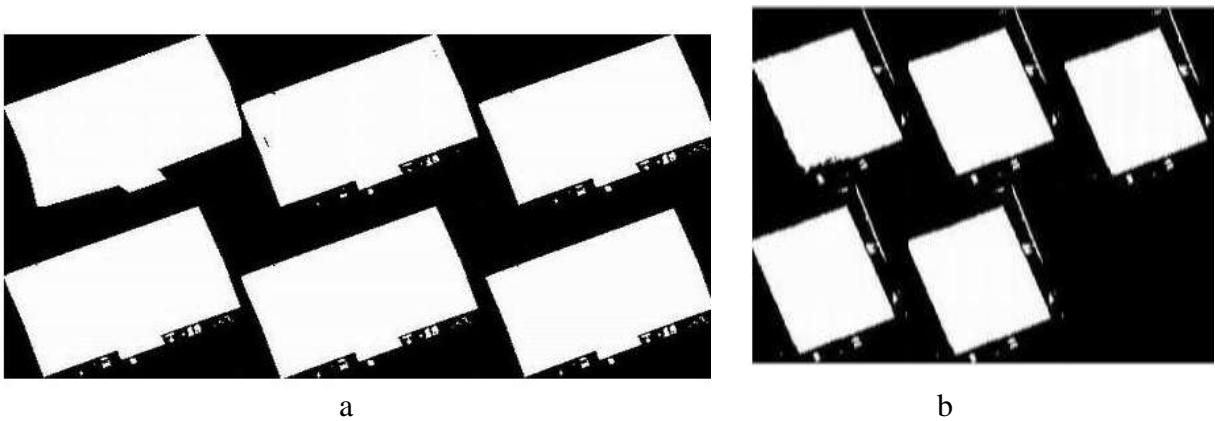


Figure3: Montage of the iterations results



Figure 4: Building class after morphological operation and area threshold

3.3 Hough Transform Phase

The principle of the Hough transform is that there are an infinite number of potential lines that pass through any point, each at a different orientation. The purpose of the transform is to determine which of these theoretical lines pass through most features in an image - that is, which lines fit most closely to the data in the image.

In order to determine that two points lie on the same potential line, it is necessary to create a representation of a line that allows meaningful comparison in this context. In the

standard Hough transform, each line is represented by two parameters, commonly called r and θ (theta), which represent the length and angle from the origin of the normal to the line in question. Using this parameterization, an equation of the line passes through a point (x_0, y_0) can be written as:

$$r(\theta) = x_0 \cdot \cos\theta + y_0 \cdot \sin\theta$$

It is therefore possible to associate to each line in the image, a couple (r, θ) which is unique if $\theta \in [0, \pi]$ and, or if $\theta \in [0, 2\pi]$ and $r \geq 0$ as shown at figure 5. The (r, θ) plane is referred to as Hough space.

The above equation corresponds to a sinusoidal curve in the (r, θ) plane, which is unique to that point.

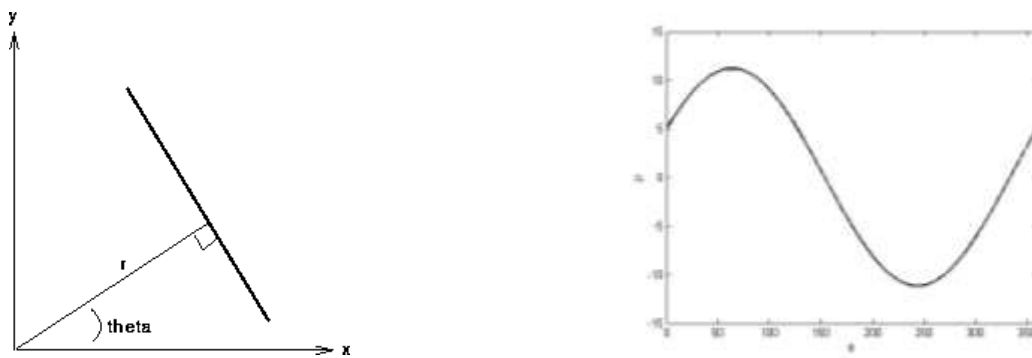
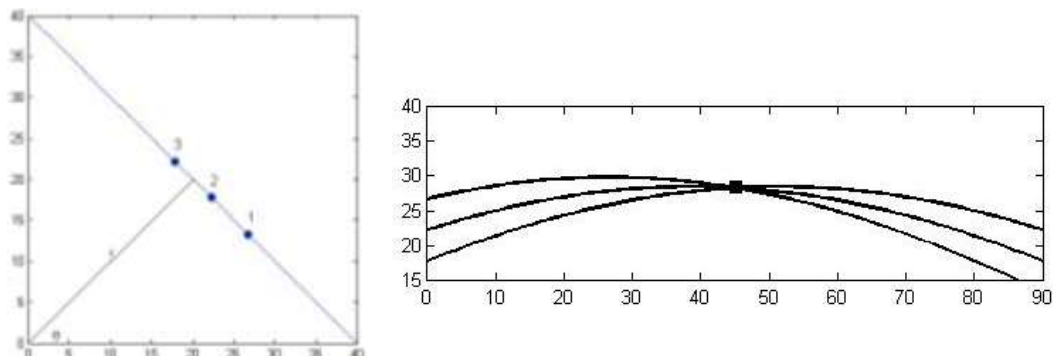


Figure 5 Parametric description of a straight line.

If the curves corresponding to two or more points are superimposed, the location (in the *Hough space*) where they cross correspond to lines (in the original image space) which pass through these points. Figure 6 shows a representation of three points forming straight line (left) and their corresponding Hough space plot (right) plotted to $\theta \in [0:\pi/2]$. This figure shows that the three points forms a straight line the radial distance of its normal is 28.4 and $\theta = 45^\circ$ shown by a filled square.

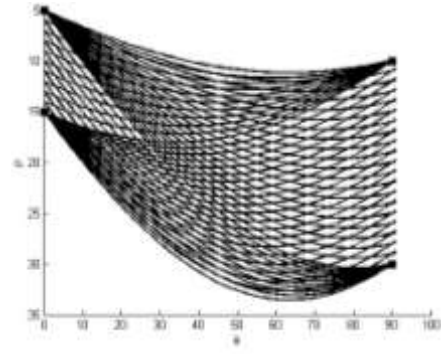
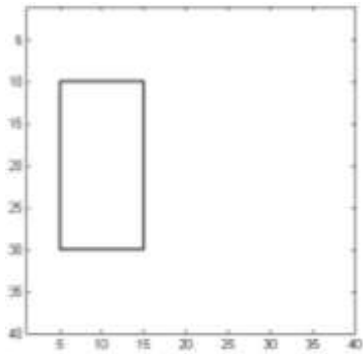
Figure 7 shows an ideal case edge of a rectangular shaped building where each side edge points' fall in the same straight line. The four lines in the Hough space at this case will be represented by a four points represented by bold squares figure 7b.



a

b

Figure 6: Hough transform for a line represented by 3 points



a

b

Figure 7: Hough transform for a square shaped building

Hough transform is used in this paper for detecting the straight lines on the building at the site seen. For this reason, the edge of the building classified object using Canny's edge detector is used. Due to errors on the sensor and some straight line edge points might not be a straight line. For this reason one straight line edge might not be detected as one straight line, it might be segmented to several straight lines taking approximately the same orientation of this straight line. Let there intersection in the Hough space being nodes (peaks) rather than an intersection point (as in the ideal case above) as that shown at figure 7. For this reason a threshold is required for choosing the circles most probably forming straight lines in the edge image and decrease the huge number of possible straight lines between the points in the edge image. Figure 8a shows the detected straight lines superimposed on the edge of this building and figure 8b shows the Hough space of the detected building straight lines.

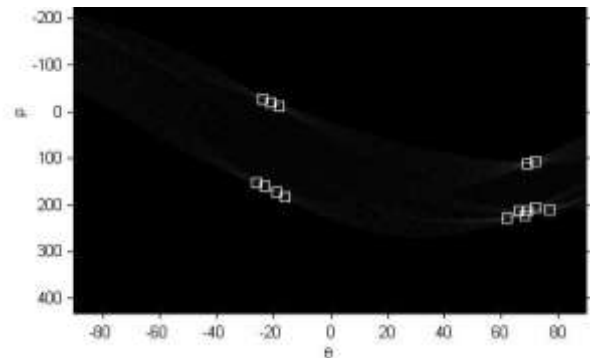
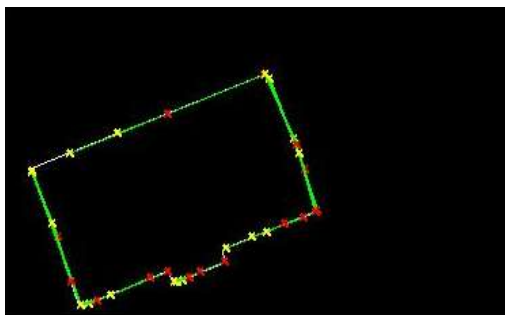


Figure 8: Hough transform for some possible lines for one building

a: Overplayed on the edge image

b: In the Hough space

3.4 Coordinate classification and polygons generalization phase:

At this stage we have some start and end points on the edge of the building polygon. A filtering process is implemented to remove any identical points (points that have nearly the same x and y coordinates). Any group of points has the same x-y coordinates ± 2 pixels are considered as identical points and the mean coordinates will be their respective coordinates. The white square figure 9 shows the position of the new coordinates of these points.

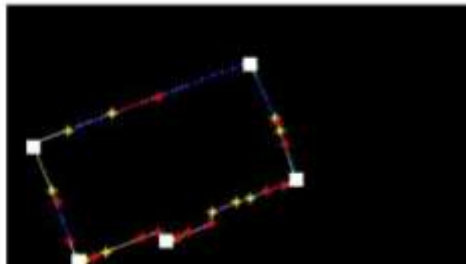


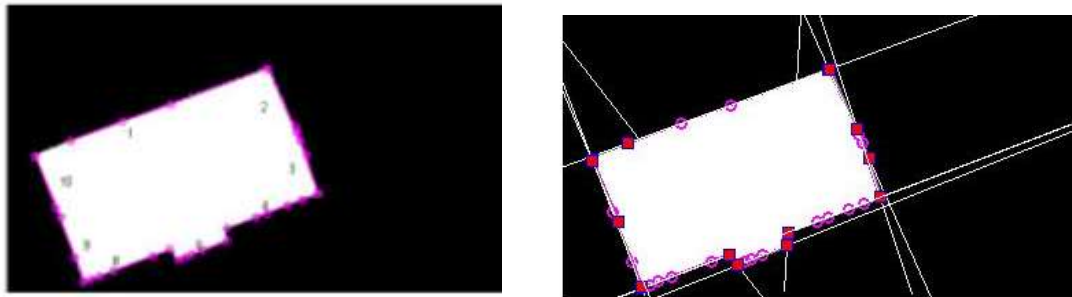
Figure9: Merging nearly equal coordinates points

The Hough line function gives the start and end points sorted in ascending order according to their x-coordinates and the chosen building polygons are chosen even in a clock wise or a counter clock wise order. In order to match all Hough lines with their respective lines on the chosen polygon, a nearest neighbor classification is implemented for this purpose. The chosen polygons are also sorted in an ascending order according to their x-coordinates. The sorted coordinates and their respective azimuths are used as training points for the classification process while the Hough lines points and their respective azimuths are used as the sample data to be classified. The result of this process is to which of the building polygons lines the Hough lines points are belong. Figure 10a shows the class number which each group of points had had. This stage is followed by a robust straight line fit to each of these classified groups of points as shown by filled squares figure 10b.

As shown in figure 10a no one of the detected lines has either class 5 or 7 this is due to that this is due to the minimum length threshold applied during the Hough transformation process. To avoid this, the algorithm will consider the end point of the previous line and the first point of the next line as the two end points of this missed line. For doing this a clock wise contouring to the building is done. The purpose of this process is to find a pixel by pixel index in a clock wise direction. A nearest neighbor classification is performed again to classify the end points of the fitted straight lines mentioned above to order their end points to a clock wise manner.

Having this done, the last coordinates of the previous and the first point on the next line are taken as the end points coordinate of the missed lines.

The intersection points of these lines with the image frame are then calculated as shown in figure 10b (white lines).and the intersection of these white lines will give the final least number of points representing the building shape.



**Figure 10: a- classified points after robust straight line fit
b- best fitted straight lines intersection with image frame**

To assess the results of the detected building boundaries (figure 11a), their areas are compared with the digitized ones (figure 11b). The resulting areas computed from both the predicted and digitized process are shown in table 1.

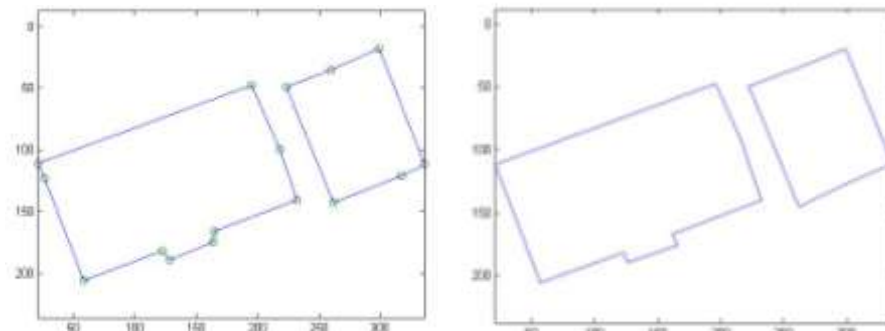


Figure 11: a-Predicted polygons (left) b: digitized polygons (right)

Table 1 Predicted and digitized polygons areas

	Building 1 (left)	Building 2 (right)
Digitized polygons areas	19236	8256
Predicted polygons areas	19169	8208
Percentage	99.7%	99.4%

4. CONCLUSION

In this paper, a system is proposed for automatic building detection from satellite images based on a new classification and line simplification (generalization) process.

In this system, the complexity of building classification process due to the interference of classes' colors is reduced by processing on a building by building manner. The processing speed is also developed by processing only the building area and a small buffering zone around it. The efficiency is also increased by sequentially updating the mean of the two discriminate classes (building and background) by iteratively adding an update to the training sets which make no pixel to be misclassified.

For delineating polygonal shapes of buildings Hough transform is used to find some lines on the edge image of each building. The end points of these lines are subjected to merging nearly equal coordinate's points, grouping using nearest neighbor classification and robust straight line fitting for each group. The intersection of the final robust straight lines will then form the buildings polygons.

The overall accuracy of the developed building extraction system was evaluated by comparing the detected polygons with their respective digitized ones.

Future work should be directed towards the generalization of the algorithm to extract circular and complicated shapes buildings

5. REFERENCES

- [1] A. Shaker. M. Hasan M. Eid & K. Zak, " **Classification and Generalization Approach for the Automatic Extraction of Buildings from High Resolution Satellite Images**"; Civil Engineering Research Magazine Faculty of Engineering El_Azhar University, Cairo, Egypt. Vol.26 No.4
- [2] Avrahami, Y., Y. Raizman, and Y. Doytsher, 2004, " Semiautomatic 3D Mapping of Buildings from Medium Scale (1:40,000) Aerial Photographs", In Proceedings of the 20th ISPRS Congress, Istanbul, Turkey, July 2004.
<http://www.isprs.org/istanbul2004/comm3/comm3.html>
- [3] C.J. Bellman, M.R. Shortis (2002) " A Machine Learning approach to Building Recognition in Aerial photographs"- ISPRS Commission III, Symposium, Graz, Austria
- [4] E. Michaelsen, U. Soergel, U. Thoennessen (2005) " Potential Of Building Extraction from Multi Aspect High Resolution Amplitude SAR Data " In: Stilla U, Rottensteiner F, Hinz S (Eds) CMRT05. IAPRS, Vol. XXXVI, Part 3/W24 --- Vienna, Austria, August 29-30, 2005
- [5] Lee J.Y., Warner T.A (2004).," Image Classification with a Region Based Approach in High Spatial Resolution Imagery", In Proceedings of the 20th ISPRS Congress, Istanbul, Turkey, July 2004,
- [6] Mori M., Hirose Y., Akamatsu Y., Li Y(2004)," Object-based Classification of Ikonos Data for Rural Land Use Mapping", In Proceedings of the 20th ISPRS Congress, Istanbul, Turkey, July 2004,

- [7] S.D Mayungaa , Dr. Y. Zhanga, and Dr. D.J. Colemana (2005) "Semi-Automatic Building Extraction Utilizing Quick bird Imagery", In: Stilla U, Rottensteiner F, Hinz S (Eds) CMRT05. IAPRS, Vol. XXXVI, Part 3/W24 Vienna, Austria, August 29-30, 2005
- [8] Sohn G.,2004" Extraction Of Buildings from High-resolution Satellite Data and Lidar",.In Proceedings of the 20th ISPRS Congress, Istanbul, Turkey, July 2004,
- [9] The Math Works, Inc (2007),” Matlab Image Processing Toolbox”, the Math Works, Inc., Natick, MA
- [10] Wang, S.D. and Y.H. Tseng, 2004, Semi-Automated CSG Model-Based Building Extraction from Photogrammetric Images, In Proceedings of the 20th ISPRS Congress, Istanbul, Turkey, July 2004,