Classification and Generalization Approach for the Automatic Extraction of Buildings from High Resolution Satellite Images

Prof. Dr A. Shaker
Prof Of surveying and Geodesy
Dr. M.M. Eid
Lecturer of Mathematics

Dr. Mahmoud M. M. Hasan,
Lecturer of surveying and geodesy
Dr. Khaled Mohamed Zaky
Lecturer of surveying and geodesy

Faculty of Engineering - Banha University

ملخص

لقد استحوذ التعرف الالي على المباني و المنشآت المدنية من الصور الرقمية المأخوذة بواسطة القمر الاصطناعي IKONOS كثير من البحوث في هذا الشأن. نظرا لاهميته البالغة في مجال المساحة التصويرية و عمل الخرائط. و هذا البحث يقدم طريقة لاستخلاص الحدود الخارجية للمباني من صورة رقمية ملونة لمنطقة السد العالي بأسوان. و الاسلوب الذى استخدم لهذا الغرض هو عمل برنامج للحاسب الالى يستخدم أسلوب التصنيف الموجه لفصل المنشآت المدنية عن باقي التصنيفات الاخرى كالرمال و الاشجار الخ. ثم بعد ذلك تحويل الصورة الرقمية الثنائية الناتجة الى متجهات خطية تمثل المضلعات الخارجية للمباني ثم تعميم هذه المضلعات لتماثل الاشكال المتعارف عليها للمباني باستخدام المتجهات. و قد استخدمت طريقة اقل مسافة من المتوسط في هذا البحث لعمل التصنيف الموجه للمباني لاعطاء الحاسب المعلومات اللازمة لعملية التصنيف.

1. ABSTRACT

The automatic recognition and reconstruction of buildings polygons from IKONOS imagery has become an important research topic with widespread applications in digital photogrammetry, urban planning, and environmental studies. In recent years IKONOS imagery has been used as a major source of information in mapping fields and became popular as any other type of mapping sources

This study presents a method to explore a new procedure for building outer boundary extraction from high resolution satellite imagery such as IKONOS.

The test area is a part of a one meter multi-spectral IKONOS image to the High Dam at Aswan. The buildings chosen for the job are in purpose chosen of different roof reflectance and relatively complex shapes to test the accuracy of the proposed algorithm (classification and generalization) to extract them. The measure of the accuracy will be the difference in area between the interpreted generalized buildings and the visually interpreted (digitized) buildings.

The procedure proposed for building extraction involves the use of supervised classification technique to the test area to segment the building classes from the background and any other classes. The images are classified using minimum distance from the mean to provide the data for the machine learning phase.

The final binary image is subjected to a, holes filling, then contouring process to extract the outer edges of the final filled image that forms the polygons of the buildings. Finally these polygons are reconstructed or refined by applying a proposed generalization algorithm. The algorithm is designed to keep the general shape of the polygons and perform sequential refinement of the building polygons by removing the unrealistic behavior caused by the sensor noise, teethes (nodes) caused when the edge line is inclined to the horizontal or the vertical and uneven reflection of the building roofs.

CERM Vol. (27) No. (3) October, 2005 Page 658-668

2. INTRODUCTION

The commercial satellite IKONOS provides an important data source for urban mapping applications. IKONOS relatively very high resolution data makes it possible to map urban buildings from satellite imagery. However, due to the complexity of image data, the automatic extraction of buildings in urban areas composes many problems. For satellite imagery in urban area, brightness values are not only controlled by surface material and texture but also by the object geometry, and illumination. The large amount of information contained in images and the numerous factors influencing a brightness value make it very difficult to separate important information from irrelevant details [5].

Regarding the fundamental methodology in spatial feature extraction from aerial and space images, a lot of experiences have been gained in the past few years. Several methods have been developed and the corresponding accuracies have been reported. Jie Shan D ET. Al.(2002) ,[4], discussed methods for building classification and generalization using the Doglas Peucker approach to remove nodes from high resolution satellite imagery. Tao Guo.et al.(2001) ,[9], carried out an experiment using a snake-based approach for building extraction from high-resolution IKONOS satellite images by combining height data. C.S. Fraser (2002), [2], uses a machine learning approach in the form of a support vector machine. The images are characterized using wavelet analysis to provide multi-resolution data for the machine learning phase.

Our study is aimed at developing a new procedure to extract urban building boundaries from IKONOS image. The procedure is expected to be advantageous in urban map updating, especially for buildings map. The next two sections will be focused on the classification and the proposed generalization approach. The study area is located at Aswan near the high Dam Figure (1). The image used for this study is a multi-spectral three band (R, G &B) IKONOS image. Since our concern is only buildings polygons and due to the lack of ground control points, the image will be used without converting it to ortho image. The buildings chosen for this study satisfy two conditions, first, they should have some complexity in shapes, and second the roofs should not have surfaces of equal reflectance. If after these two conditions the buildings are extracted with a certain accuracy it will be a judge for the convenience and generality of the proposed classification and generalization processes.

Figure (1): Chosen scene.

3. CLASSIFICATION STAGE METHODOLOGY

The first step of this study is the supervised classification of the multi-spectral image. The objective of this stage is to automatically categorize all pixels in the image into land cover classes or themes. Normally, multi-spectral data are used to perform the classification for each pixel according to its numerical spectral data (DNs) [6]. Four thematic classes are chosen for this purpose namely, Buildings, Trees, March and Sand.

The first stage of the proposed classification and generalization algorithm is the categorization process. This process is performed by specifying to the computer algorithm, numerical descriptors of the various land cover presented in the scene by taking samples of known cover types in different sites in the scene. These representative samples "training data set" are used as an interpretation key that describes the spectral attributes for each land cover type. Each pixel in the image data set is categorized into the land cover class it most closely resembles. If the pixel is insufficiently similar to any training dataset, it will be labeled as "unknown". The judge for the closeness of each pixel to the land cover classes used in this study is the minimum distance from the mean classifier. Figure (2) shows the resulting image from this stage

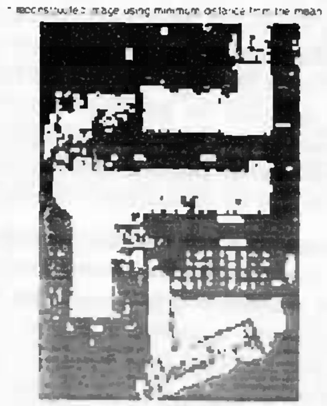

Figure (2): Binary image after categorization process

In the second stage, and to overcome the correlation between the four land cover classes as shown in the scatter plot shown in Figure(3) specially between building, marsh and Sand classes, the unknown pixels are re-categorized, but this time using its 3x3 nearest neighbor's classifier. Each unknown pixel will be assigned the class that is mostly repeated in this 3x3 kernel. This means, set a pixel to 1 if five or more pixels in its3-by-3 neighborhood are 1's

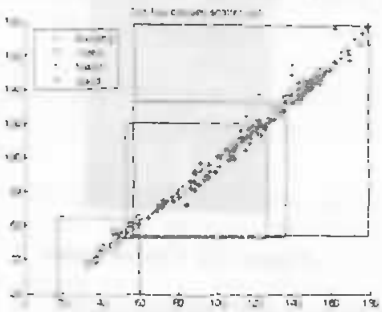

Figure (3): Scatter plot for the four land covers
X-axis: Band1 DN's & Y-axis: Band2 DN's

There are nine kernels used for this classification step depending upon the position of the pixel under classification in the resulting binary image matrix. In other words, edge pixels will have kernels different than those forming the building body pixels as shown in the table (1)

Table (1): Kernels used for Nearest Neighbor Classification

i,j	i,j+1
i+1,j	i+1,j+1

a Corner kernel

i,j-1	i,j	i,j+1
i+1,j-1	i+1,j	i+1,j+1

b edge kernel

i,j-1	i,j
i+1,j-1	i+1,j

a Corner kernel

i-1,j	i-1,j+1
i,j	i,j+1
i+1,j	i+1,j+1

b edge kernel

i-1,j-1	i-1,j	i-1,j+1
i,j-1	i,j	i,j+1
i+1,j-1	i+1,j	i+1,j+1

c- Body kernel

i-1,j-1	i-1,j
i,j-1	i,j
i+1,j-1	i+1,j

b edge kernel

i-1,j	i-1,j+1
i,j	i,j+1

a Corner kernel

i-1,j-1	i-1,j	i-1,j+1
i,j-1	i,j	i,j+1

b edge kernel

i-1,j-1	i-1,j
i,j-1	i,j

a Corner kernel

The buildings after this step may contain holes "0's edged by 1's in the building matrix" due to the sensor noise and uneven reflection of the building roofs which causes that some pixels could be misinterpreted by the algorithm.. These unrealistic holes need to be detected and removed from the building body. The algorithm will detect and fill these rings so that only the building body is remained for further generalization processing

The result of using the above mentioned decision windows and filling the polygons holes is shown in figure (4). One should notice that any hole in this stage will be closed even if the hole is a part of the building.

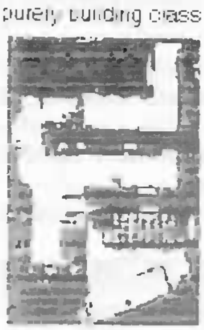

Figure (4): The resulting binary image after using 3x3 kernels and filling holes

To extract the building boundaries, the combined classification resulting image are processed to detect building boundaries using some facilities supported by Matlab. In this process the exterior boundaries of the final objects in the binary image are traced where the none zero pixels belong to an object and 0-pixels constitute the background. The initial point to begin search for the boundary is chosen at the upper left corner of the image and the direction of the search is west clock wise direction.

The result from this process is a pixel by pixel coordinates of each point in the boundary of each object in the binary image in the pixel coordinate system. Figure (5) shows the predicted shapes of the buildings in vector format. A close look to this figure shows that the building at lower right part of figure (1) is not completely interpreted but rather it is interpreted as some closed polygons. This is due to that some parts of the building are interpreted as a Marsh class while the others are interpreted as a building class. For this reason the program after the classification stage will prompt the user [*Are you want to merge some polygons: y/n[n]*], the default answer is no. But if the user answer with "y", figure (5) will be displayed again with the prompt [*Click on the polygons you want to merge, press Enter when finish*]. The user should keeps clicking on the polygons he wants to merge or clicking on the polygons that have the maximum x, minimum x maximum y, and minimum y coordinates. An Enter keystroke will close the choice of the polygons. The chosen polygons will then be merged in one polygon representing the coordinates chosen above. The program will continually toggles between the above mentioned two prompts until an answer with 'n', default answer, is performed to the first prompt

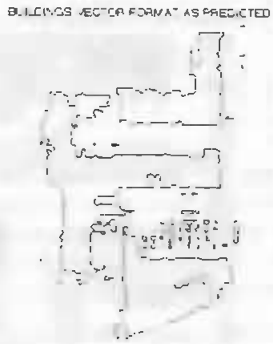

Figure (5): Predicted buildings in vector format.

After the building class is converted to vector form, each building will be represented by a polygon object. The area of each building polygon is then calculated. The predicted polygons after this stage are subjected to an area threshold. Building polygons below a given area threshold caused by misclassification are then regarded as a non building objects and eliminated from further processing. The program will show figure (5) again prompting the user with (*What is the minimum building area in pixels to be processed ? [ ]*). An enter keystroke will process all the interpreted areas or the user should writes the minimum areas in pixels should be processed

4. GENERALIZATION STAGE METHODOLOGY

The objective of building generalization is to obtain accurate building boundaries resembles in shapes to those in the original image Figure (1) than those obtained from the vectorisation results Figure (5).

As shown in Figure(5), the original vectorisation of building classes contains many unexpected shapes this is due to that the contouring of the buildings is performed in the pixel format by connecting the center of each pixel in the boundary to the center of the next boundary pixel. This makes building polygons always contains interior angles a devisor of 45°. This makes us always faced with several line connections represented by solid lines as shown in figure (6)

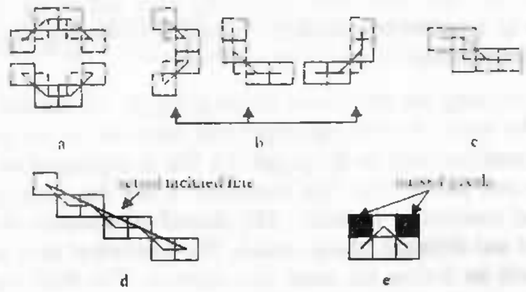

Figure (6) Unrealistic building connections

As shown in Figure (6-a, b, c, e), corners of building polygons are connected unrealistically or not completed correctly forming a 45° or 90° nodes (interior angles >180 followed by interior angles< 180 or visa versa in a one pixels distance).This due to that some pixels are missed or wrongly interpreted as a background pixels due to the shadow of buildings and trees. Some other nodes are due to the phenomenon mentioned before (working with pixel format). Another reason of this nodes are due to the fact that inclined straight boundaries at building polygons are represented as a zigzagged line due to that its inclination is not a devisor of 45° Figure(6d).

The value and sign of each node in a sample building polygons (figure (7)) are determined using the Douglas-Peucker (1977) formulations as mentioned in [4]

$$\Delta_i = \frac{p_{i-1}p_i \cdot p_i p_{i+1}}{|p_{i-1}p_i| \cdot |p_i p_{i+1}|} \qquad (1)$$

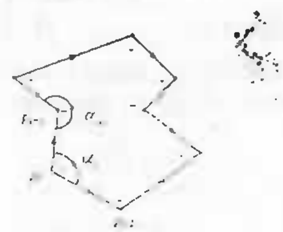

Figure (7): Sample polygon.

Where

$P_{i-1} P_i = [(x_i - x_{i-1}), (y_i - y_{i-1}), 0]$

$P_i P_{i+1} = [(x_{i+1} - x_i), (y_{i+1} - y_i), 0]$

α is the internal angle at each point computed from the azimuths using the usual surveying conventions

To solve the above mentioned problems, a vector based generalization technique is proposed. This technique is outlined as follows.

As a fist step, the formulations mentioned above are used to determine the value of Δ at each node in the building polygons. The algorithm will then track each two consuctive points having the same sign Δ (two positive or negative sign Δ) like those shown with solid lines at figure(6 a&b) and change their coordinates to form right angle between the two lines connected to each point of them(shown by dashed lines). The correction for each point will be based on the azimuth of the previous and next lines connecting to it and using their computed distance in pixels. This step is an iterative step until the condition of finding two consutive points having the same sign Δ is not satisfied.

After reconstructing the 90°corners of the polygon, the second step of the generalization algorithm will be the search for the zigzagged line segments in the polygon like that shown in figure (6d). Three condition will be the judge if a line is zigzagged or not, and if it is, what are the positions of its end points. The first condition is, are the nodes (polygon points) are with different sign Δ and conductive in order. The second is, compute the azimuths from the first point of this ordered and different of sign nodes. The points that have azimuth differences within a given threshold will be follow the same line segment. The third condition is, if the distance between each two successive nodes in each line segment are within a distance threshold "<=√2 pixels" classify these points to follow the same line segment. The intersection of the results of the above mentioned three conditions will be the order of the points in the building vector that follow a zigzagged line segment.

Each point at any zigzagged line segment, will then be subjected to a straight line fit function to project the x or y coordinates of it "depending on the azimuth between the first and end point in the line segment" on the straight line equation represented by the end points of each line segment. There are two straight line equations used for this purpose. The first, when the azimuth of the end points is within the rang (45° < az <135° or 225°< az <315°) the y- coordinate of ant point (y_i) will be fixed and the x-coordinate (x_i) will be changed as follows:

$$m = \frac{x_2 - x_1}{y_2 - y_1}$$
$$c = x_1 - my_1 \qquad (2)$$
$$x_i = my_i + c$$

Where, m is the slope of the straight line

If the azimuth of the objective line is in the rang (45° > az >315° or 225°> az >135°), the change will be in the y - coordinate and the x - coordinate will be fixed as follows:

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$
$$c = y_1 - mx_1 \qquad (3)$$
$$y_i = mx_i + c$$

Where

x_1, y_1 and x_2, y_2 are the coordinates of the end points of the line segment

x_i, y_i are the coordinate of the point to be projected

The third step of the generalization algorithm is the reconstructing of the line combination like that shown by solid lines " two successive points having different sign Δ." to form perpendicular angles to match those shown by dashed lines figure(6c). The same functions used in the second step will be also used here. But the experience we have gained during the processing shows that this step should be after the smoothing of the inclined lines.

The fourth step is projecting any two successive lines having internal angles close to 180° or 90° within an angle threshold to be straight lines or perpendicular lines respectively. This angle threshold should not be big enough to change the smoothness of the circular or curved objects "not tested in this study". The equations used to make the azimuth az_2 perpendicular to the other azimuth az_1 are:

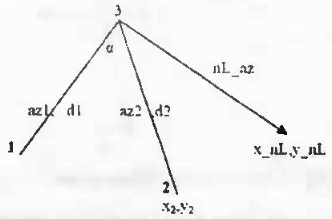

Figure (8): Perpendicular angle between 3 points.

$$nL\_az = az_2 - (\alpha - a) \tag{4}$$

Where a =90° or 270° according to sign Δ

$$x\_nL = x2 + d2.\sin(nL\_az) \tag{5}$$

$$y\_nL = y2 + d2.\cos(nL\_az) \tag{6}$$

The resulting building vectors from the above described generalization process is shown in Figure (9)

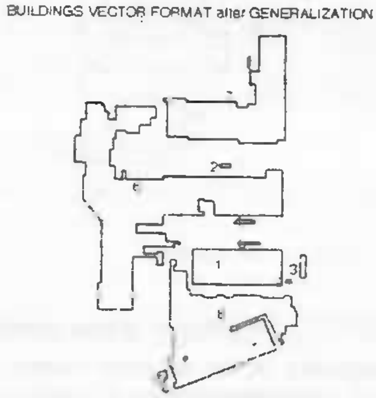

Figure (9): Building vector format "in pixel coordinates".

The last stage in the generalization process is to convert the building vectors from the pixel format (dotted lines) to the special format (bold lines) as shown in figure (10). In converting to spatial coordinates the x&y coordinates of each point in the building polygon will be increased or decreased according to the sin and cos of the azimuths of lines connected to each point A function to perform that purpose is created using also the normal surveying conventions for computing the coordinates differences from the azimuth and the length of the line as shown in figure (10).

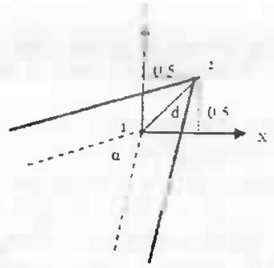

Figure (10): Converting from pixel to spatial coordinates.

The formulae used for the conversion from pixel to spatial coordinates are:

$$d=0.5\ \text{pixel}/\cos az_{12} \qquad (7)$$

$$x_2=x_1+d.\sin(\text{external angle}/2) \qquad (8)$$
$$y_2=y_1+d.\cos(\text{external angle}/2)$$

The resulting boundary polygons after conversion to spatial coordinates is shown in figure (11)

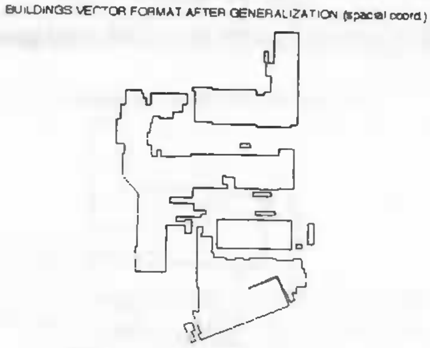

Figure (11): Buildings in spatial coordinates.

For measuring the accuracy of the extracted building polygons, the same polygon extracted are also digitized by visual interpretation. For refining the digitized data from error induced incorporated by the digitization process, the generalization algorithm is also used for this purpose. The resulting polygons from both digitization and generalization processes are shown in Figure (12). The measure to the accuracy will be the difference in area between the extracted and digitized polygons as shown in Table (2)

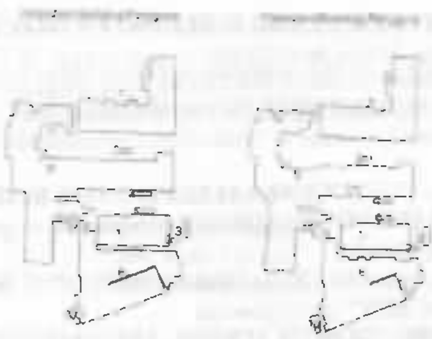

Figure (12): Digitized and vectorized building class.

Table (2): Polygons areas " in pixels ".

Digitized and Predicted polygons area										
polygons	1	2	3	4	5	6	7	8	9	10
Digitized (d)	510	11	22	18.4	18	2362.3	927	1283.5	6	21.7
Predicted(p)	390.7	9	25.2	16.5	16.4	2215.7	929.3	127	5.3	22.7
d-p	119.3	1.8	-3.20	1.91	1.64	146.55	-2.31	16.51	0.7	-1
percentage	0.23	0.16	-0.14	0.10	0.09	0.06	-0.002	0.07	0.12	-0.04

5. CONCLUSION AND FUTURE WORK

A method for the automatic extraction of buildings from high-resolution IKONOS data urban areas has been presented. The site chosen for this purpose is at the City of Aswan at the High Dam area. This research shows that this approach is quite promising building extraction processes. A difference in areas of 0.2% up to 23% pixels between the digitized and the predicted polygons has been attained. A lot of these errors are due to that some pixels are missed or wrongly interpreted as background pixels due to the shadow of buildings or the color of the roof itself (polygon 1 & 6). It seems that the confusion between buildings and ground will be unavoidable, the discrimination between them still requires some user interaction to add missed data like that missed corner at polygon 6. For these reason, the accuracy obtained from these two polygons shouldn't affect the high potential of the proposed algorithm.

Future work will be directed towards the assessment of this algorithm quality by comparing it with other software. In addition, improving the capabilities of this building extraction algorithm by additionally using other classifiers such as maximum likelihood, Mahalanobis distance …etc classifiers

6. REFERENCES

[1]C.J. Bellman, M.R. Shortis (2002)"A MACHINE LEARNING APPROACH TO BUILDING RECOGNITION IN AERIAL PHOTOGRAPHS"- ISPRS Commission III, Symposium, Graz, Austria

[2]C.S Fraser, E. Baltsavias, A. Gruen (2002)' **PROCESSING OF IKONOS IMAGERY FOR SUBMETRE 3D POSITIONING AND BUILDING EXTRACTION**" ISPRS Journal of Photogrammetry & Remote Sensing 56 177-194

[3] F. Rottensteiner, Ch. Briese "**A NEW METHOD FOR BUILDING EXTRACTION IN URBAN AREAS FROM HIGH RESOLUTION LIDAR DATA** ".Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Gußhausstraße 27-29, A-1040 Vienna, Austria – (fr,cb)@ipf.tuwien.ac.at

[4] Jie Shan D. Scot Lee (2002) "**GENERALIZATION OF BUILDING POLYGONS EXTRACTED FROM IKONOS IMAGERY**" Symposium of geospatial theory, Processing and application, Ottawa, Canada

[5Jie Tian, Jinfei Wang, Peijun Shi (2004) **URBAN BUILDING BOUNDARY EXTRACTION FROM IKONOS IMAGERY**" jtian5@uwo.ca,

[6]Lillesand & Kiefer, (1994). "**REMOTE SENSING AND IMAGE INTERPRETATION THIRD EDITION**", John Wiley sons, Inc.

[7]M. Gerke, C. Heipke and B.-M. Straub (2001)' **BUILDING EXTRACTION FROM AERIAL IMAGERY USING A GENERIC SCENE MODEL AND INVARIANT GEOMETRIC MOMENTS** 'Proceedings of the IEEE/ISPRS joint workshop on remote sensing and data fusion over urban areas Rom-Italy ISBN 0-7803-7059-7, pp85-89

[8]M. Gerke, Bernd-Michael Straub & Andreas Kochi (2002)"**AUTOMATIC DETECTION OF BUILDINGS AND TREES FROM AERIAL IMAGERY USING DIFFERENT LEVELS OF ABSTRACTION**" {gerke, straub,koch}@ipi.uni-hannover.de

[9]Tao Guo, Yoshifumi Yasuoka (2001)"**SNAKE-BASED APPROACH FOR BUILDING EXTRACTION FROM HIGH-RESOLUTION SATELLITE IMAGES AND HEIGHT DATA IN URBAN AREAS**", Institute of Industrial Science, University of Tokyo

[10] The Math Works, Inc (2004)," **MATLAB IMAGE PROCESSING TOOLBOX**", the Math Works, Inc., Natick, MA