AUTOMATIC FEATURE EXTRACTION FOR MAP REVISION

Dr. Hiroshi MURAKAMI
Deputy Director
International Affairs Division
Ministry of Construction, JAPAN

and

Dr. Roy WELCH
Director
Center for Remote Sensing and Mapping Science
University of Georgia
Commission IV

ABSTRACT:

Automatic feature extraction techniques were developed for use with digital images and map data to assess the feasibility of employing expert systems for map revision. The map and image data were placed in register to create a cartographic database suitable for use with a prototype expert system optimized for the extraction of building features. The expert system approach permitted control of image processing routines applied to the cartographic database for feature extraction. The accuracy of feature extraction increased as the image pixel resolution was improved.

KEY WORDS: Cartographic Database, Change Detection, Expert System, Feature Extraction, Image Processing, Map Revision.

1. INTRODUCTION

Most developed countries have completed national mapping programs that provide topographic map coverage at scales of 1:25,000 or smaller, and map revision is now the main task. Urban expansion, however, causes maps to become out-dated rapidly, while funds allocated to mapping have been reduced. Consequently, there is a need for more efficient and cost effective methods for labor-intensive map revision tasks, particularly for change detection, in which differences between newly acquired images and old maps are determined.

Feature extraction studies have mainly focused on objects such as roads and buildings included in a digital image (Bajcsy and Tavakoli, 1976; Nagao and Matsuyama, 1980; Nevatia and Babu, 1980; Fischler et al., 1981; McKeown et al., 1985; Huertas and Nevatia, 1988; Wang and Newkirk, 1988).

The objective of this study was to develop a method of detecting changes of buildings in SPOT images. Since change detection method needs the photo interpreters' knowledge to identify each detected change, expert system approach was employed to deal with human expertise (Murakami, 1990).

2. ISSUES IN FEATURE EXTRACTION

This study focused on the following three important points out of the problems encountered in feature extraction (Nagao and Matsuyama, 1980; Hanson and Riseman, 1988; Matsuyama and Hwang, 1990).

2.1 Initial Parameter Value Selection in Image Segmentation

Computers can not reliably extract specific objects directly from gray-scale images. Consequently, the original gray-scale image must be transformed first to an image in which each ground feature is independently labeled. In image segmentation, initial parameters (e.g., threshold values) must be employed to distinguish ground features from their background. Appropriate threshold values, however, may differ from feature to feature - even in a single image. Consequently, developing a method to select appropriate threshold values in an a priori manner will be required.

2.2 Extraction of Descriptor Values

Interpretation of individual labeled regions requires descriptors of the characteristics of each ground feature. Most descriptors are related to the seven elements of photo interpretation, i.e., tone, shadow, pattern, size, texture, shape, and association (Paine, 1981; Lillesand and Kiefer, 1987). In theory, extraction and proper processing of all the information concerning these elements would provide the same understanding of the input image as human interpreters. Hence, selection of the most important elements for a particular kind of features, i.e., building, will be necessary.

Consequently, there must be a procedure for establishing values for descriptors related to each of the interpretation elements. Of course, it must be understood that "human perception" does not necessarily correspond with "machine perception".

2.3 Uncertainty Management and Inference Method

Some uncertainty is associated with descriptor values derived from segmented regions. Thus, knowledge or guiding rules must be applied to establish the identity of each object. Unfortunately, these rules may also contain some uncertainty. For example, "A bright, elongated object (20 m x 40 m) in a satellite image is a building," may be true in most instances. However, a road or agricultural field may exhibit similar characteristics. Hence, there must be a method for managing uncertainties in an inference process.

3. APPROACHES FOR BUILDING EXTRACTION IN CHANGE DETECTIN

Solutions for the above mentioned problems are proposed, and approaches to building extraction for map revision are described in this section.

3.1 Region Growing Method and Threshold Value Selection

In this study, the region growing method was employed for image segmentation. This method assigns the same label to the pixels with relatively uniform digital numbers (DN) in a region. In satellite images, buildings usually have larger DNs than the surrounding ground. Hence, in classifying the segmented regions as building candidates or background, a threshold value was employed to the average DN of each segmented region.

As discussed above, it is difficult to derive an appropriate threshold value in an a priori manner. In map revision, however, old maps are available for locating areas containing buildings. These areas can be located in the image data, and sub-image of the buildings extracted. The histograms of such sub-images will form a bi-modal structure as shown in Figure 1. The DN value at the valley point of the histogram is considered as the most appropriate threshold value for the sub-image. The image in Figure 1 was segmented with the region growing method, and then divided into a building candidate and background using the threshold value as discussed above (Figure 2). The sub-images of all the existing buildings were examined with this method to derive an average threshold value which was assumed to be applicable to the entire input image.

3.2 Selection of Descriptors

Of the seven elements employed in human photo interpretation, shadow, pattern, texture, and association are not useful for building extraction from satellite images. However, the rest of the elements, tone, size, and shape are all useful for building extraction. The first two elements were defined as the average DN value and the area of each segmented region,

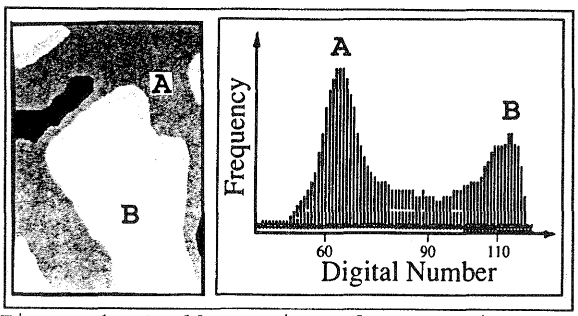


Figure 1. Small section of a SPOT image around an existing building and its histogram showing bi-modal structure.

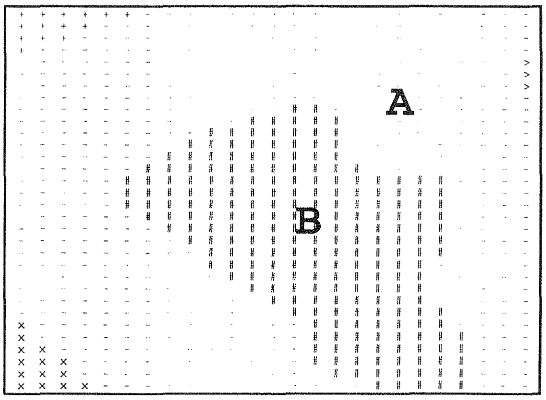
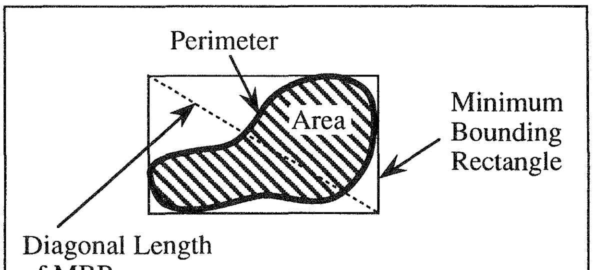


Figure 2. Result of image segmentation with the threshold value derived in Figure 1.

respectively. Since there is no one parameter which can properly describe shape, several descriptors were employed to indirectly define shape. These descriptors include elongatedness, perimeter length, and diagonal length of minimum bounding rectangle (MBR) (Figure 3).

In order to examine the utility of these descriptors, the test pattern shown in Figure 4 was recorded in the laboratory at two equivalent image resolutions, 10 m (Test Image A) and 2.5 m (Test Image B). Test Image A was then rotated from 0 to 90 degrees in 10 degree steps and resampled to resolutions of 10 m and 2.5 m. Test Image B, on the other hand, was also rotated from 0 to 90 degrees, but resampled only to 2.5 m pixel resolution. The effects of rotation and pixel resolution on descriptor values are further discussed below.



$$\text{Elongatedness} = \frac{(\text{Perimeter})^2}{4\pi\,(\text{Area})}$$

Figure 3. Definition of descriptors for shape.

The descriptor values of the features in these rotated images are shown in Figure 5 through 7. The shape of the features in the images of 2.5 m pixel resolution (Figure 6 and 7) can be distinguished using the calculated elongatedness and area. The graphs of elongatedness for the 10 m pixel resolution image, however, intersect one another, indicating the difficulty of providing proper shape information. These graphs clearly show that the pixel resolution of the input image is important when defining shape with descriptors.

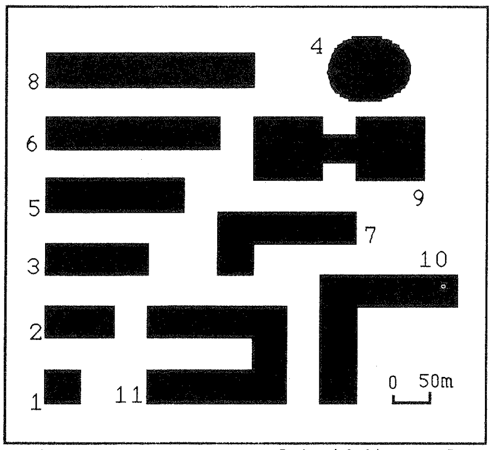


Figure 4. Test pattern of building of different shape.

In order to further clarify the relationship between pixel resolution and fidelity of descriptor value, another test image with 10 rectangles (Figure 8) was created and processed in the same way as above. Figures 9 through 11 show the results of calculated descriptor values. Those calculated from 10 m pixel resolution image show the expected values for the 5 largest rectangles (6-10), whereas descriptor values derived from 2.5 m images are correct for rectangles 3-10 indicating the superiority of the higher resolution images. Similar results were derived for the other descriptors.

A 10 m pixel resolution image is displayed in Figure 12. Human interpreters can easily define shape of rectangles as small as rectangle No. 3. In order to obtain equivalent results by machine interpretation, however, descriptors must be derived from 2.5 m pixel resolution images. This indicates that building recognition with shape descriptors may require an image of four times smaller pixel resolution (16 times larger data volume) to be comparable to human interpreters.

3.3 Inference with Uncertainty

In this study, the expert system approach was employed since the interpretation of building candidates by applying human knowledge to their descriptor values is similar to the process in diagnostic expert systems. In implementing an expert system, uncertainty management and knowledge representation are the two most important factors to be specified. This study employed the approach of MYCIN (medical diagnostic expert system), i.e., the certainty factor model for uncertainty management and production rules for knowledge representation (Buchanan and Shortliffe, 1984).
An example of simple production rules developed in this study for map revision is shown below.

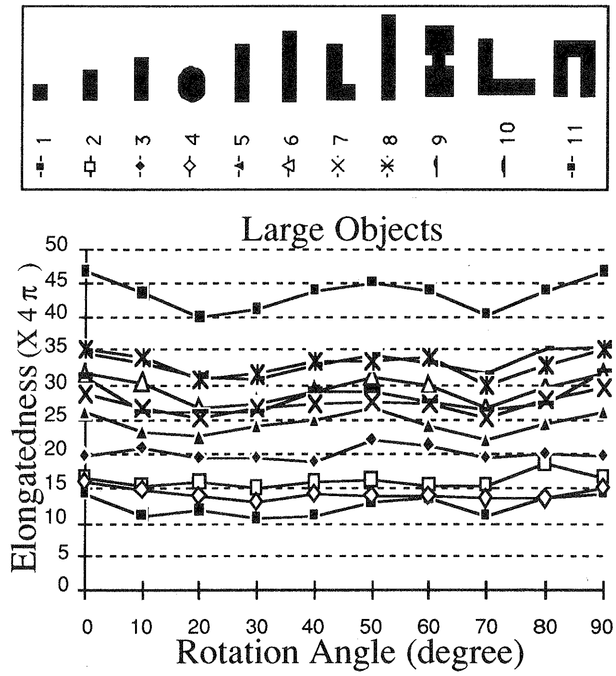


Figure 5. Elongatedness plotted against the rotation angle for the image of 10 m pixel resolution.

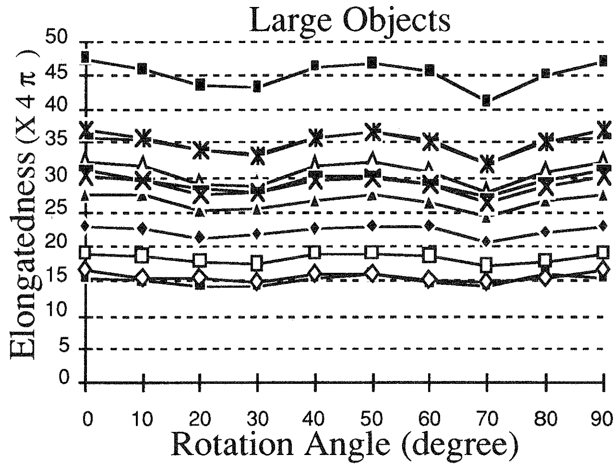


Figure 6. Elongatedness plotted against the rotation angle for the image of 2.5 m pixel resolution resampled from the 10 m test image.

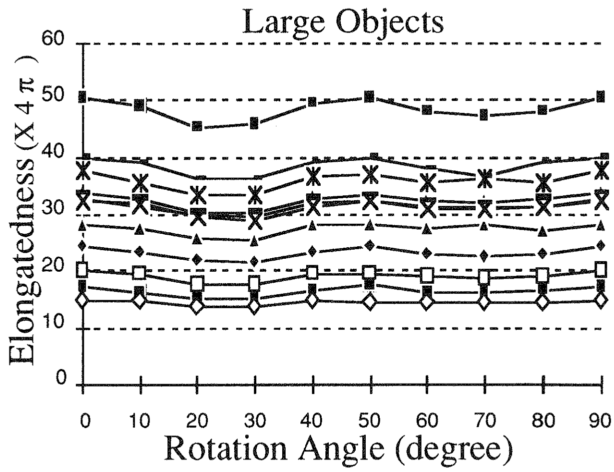


Figure 7. Elongatedness plotted against the

rotation angle for the image of 2.5 m pixel resolution resampled from the 2.5 m pixel resolution test image.

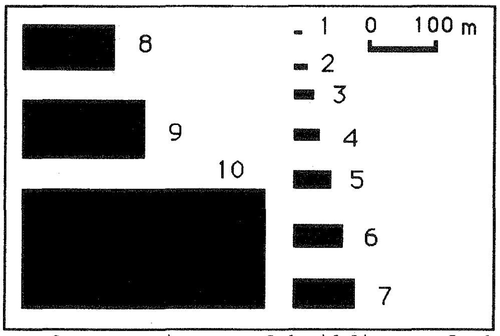


Figure 8. Test image of buildings of the same shape (2:1 rectangles).

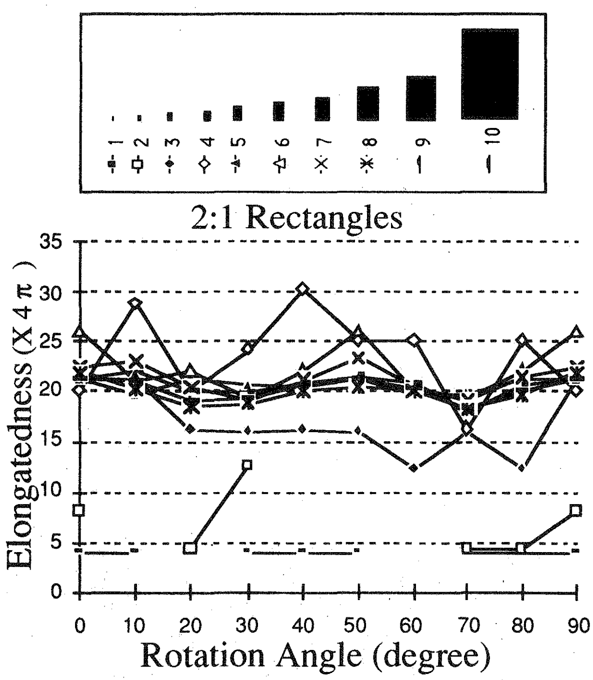


Figure 9. Elongatedness plotted against the rotation angle for the image of 10 m pixel resolution.

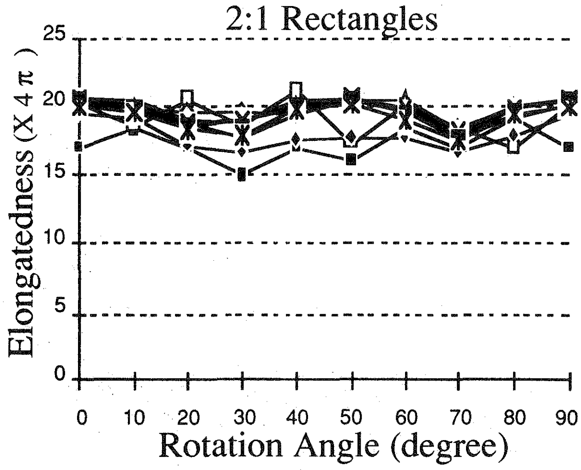


Figure 10. Elongatedness plotted against the rotation angle for the image of 2.5 m pixel resolution resampled from the 10 m pixel resolution test image.

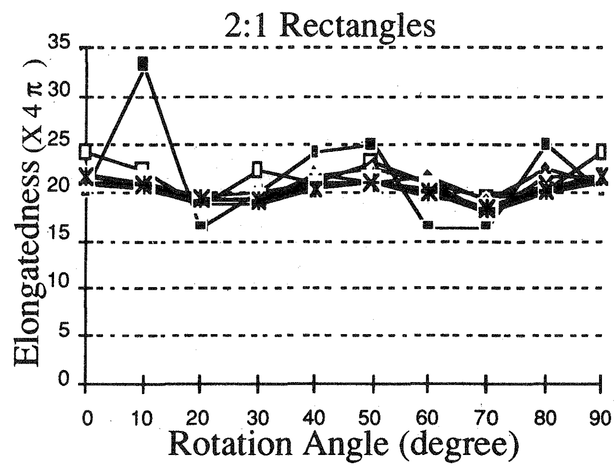


Figure 11. Elongatedness plotted against the rotation angle for the image of 2.5 m pixel resolution resampled from the 2.5 m pixel resolution test image.

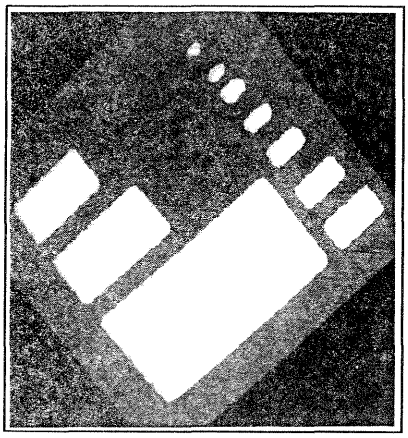


Figure 12. Test image of 10 m pixel resolution displayed on a computer screen.

RULE 1
IF (area) IS (medium) AND
(elongatedness) IS (small) AND
(perimeter straightness) IS (medium) AND
(diagonal length of MBR) is (medium)
THEN
(object is a building) CF = 0.5

The role of the rules is to identify a portion of the n-dimensional space spanned by n descriptors as shown in Figure 13 for a particular building candidate, to calculate its probability as building, and to conclude whether or not it is a building.

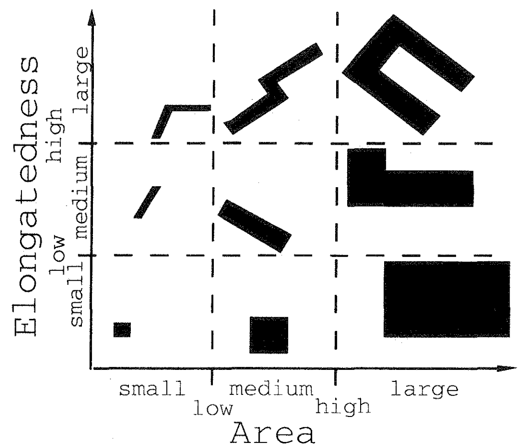


Figure 13. Two-dimensional descriptor space spanned by area and elongatedness.

Building candidates with insufficient probability need to be processed again from the original image with more suitable parameter values since further processing of the segmented image would add little useful information. Once building candidates are found and analyzed in the image in the first processing pass, however, their location and appropriate parameter values can be given for each of the candidates in the following processing passes (focusing mechanism). Each processing pass outputs newly identified buildings which are accumulated in a "building file" as shown in Figure 14. Several processing passes are required to record all the detectable buildings in the original image in the file.

4. EXPERIMENT WITH SPOT IMAGE

The method described in the previous section was implemented with the programming language C in a VAX Station 3500 installed in the Center for Remote Sensing and Mapping Science of the University of Georgia, and applied to a SPOT image. The rules were recorded in a text file separately from the expert system inference routine.

4.1 Data Used in the Experiments

A SPOT panchromatic image recorded on May 4, 1986 covering Atlanta, Georgia was used as the source of the original image for the experiments. A test area was selected from the USGS 1:24,000 topographic map, "Chamblee, Ga." After image rectification, the test area was cut out from the original image and resampled to an image of 300 by 400 pixels with 5 m pixel resolution as shown in Figure 15.

The black and purple separates of the topographic map were also digitized with a linear array CCD camera and rectified for the same area and pixel resolution as Figure 15. Noise in the resultant map separate image was eliminated with image processing techniques (Figure 16).

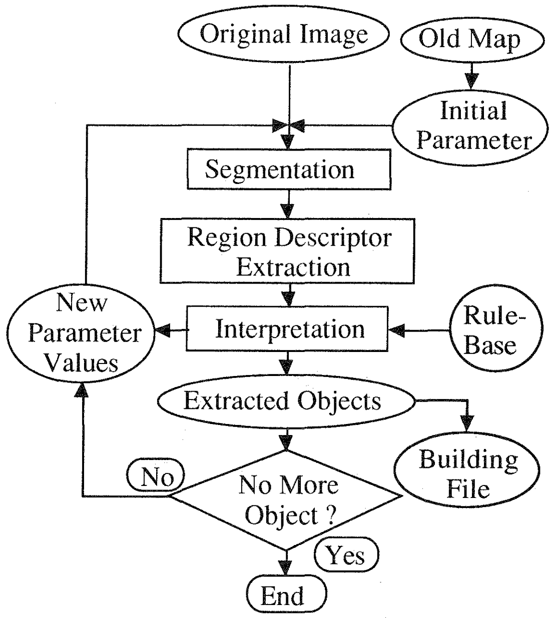


Figure 14. Flow of new building extraction process.

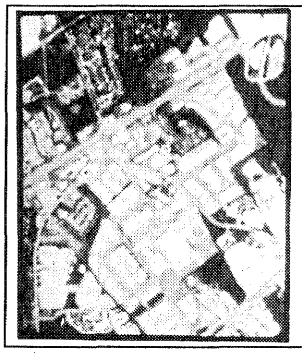


Figure 15. SPOT image of test area.

Figure 16. Map separate image of test area.

4.2 Result of Experiments

The test area has a number of existing large industrial buildings and factories. The segmentation result from the first pass is shown in Figure 17. Four categories of segmented regions may be noted: 1) correctly segmented regions; 2) regions which contain more than one feature (multiple feature regions); 3) bright background (mostly bare ground); and 4) regions for which only the edges were segmented. Multiple feature regions were caused by the bright background surrounding buildings. Also edges tend to form where the intensity gradient of an object edge is not uniform relative to the background.

The segmented images were then processed with the expert system to extract building-like regions using shape, size and tone descriptor values of each region. Figure 18 shows the regions which were confirmed as new buildings (from the SPOT image) in the first pass. The remaining uncertain regions were then expanded as shown in Figure 19 to make a mask image for the next pass. In the second and later passes, the input SPOT image was processed only within these masked areas with specific processing instructions such as new threshold values. If a region was likely to include more than one feature connected through narrow channels, the shrinking-and-expanding method was employed to isolate each independent feature.

Figure 17. Segmentation result of the SPOT image from the first pass.

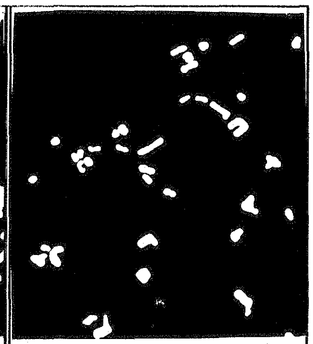


Figure 18. New buildings extracted from the SPOT image from the first pass.

The segmentation result for the second pass on the SPOT image is shown in Figure 20. Regions which were incorrectly segmented in the first pass now appear to have more building-like shapes. For example, the buildings indicated by the arrow were successfully decomposed by the shrinking-and-expanding technique. In addition, existing regions for which only the edges were segmented in the first pass, were successfully segmented with new threshold values. New buildings extracted in the second pass were merged with those extracted in the first pass. In this way, newly extracted buildings in each pass were accumulated in the building file. The final result for the test area after four iterations is shown in Figure 21.

Figure 19. Mask image of focused areas where further processing is required.

Figure 20. Segmentation result of the SPOT image from the second pass.

The accuracies of this result were calculated by comparing Figure 21 with the buildings identified by manual techniques on the original SPOT 10-m pixel image (Figure 22). The ratio of correctly extracted features to manually detected changes (i.e., accuracy) was 82 % for the experiment. The omission errors were mainly small or dark features which might not have been extracted by photo-interpreters.

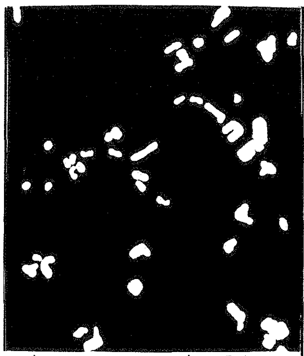


Figure 21. Final result of new buildings extracted from the SPOT image.

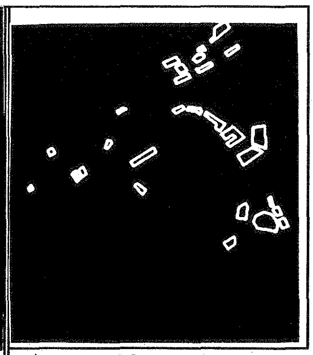


Figure 22. New buildings manually extracted from the SPOT image.

The same feature extraction procedure was applied to other areas of the SPOT image. In addition, the SPOT 10 m data were resampled to 2.5 m pixel resolution and feature extraction attempted with these "higher resolution" data (Figure 23). Machine feature extraction of buildings showed definite improvement using these resampled 2.5 m pixel resolution data (Figures 24 and 25).

Figure 23. SPOT image of 2.5 m pixel resolution for building extraction.

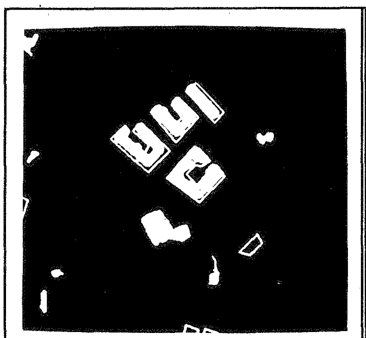


Figure 24. Result of new building extraction from the SPOT image in Figure 23.

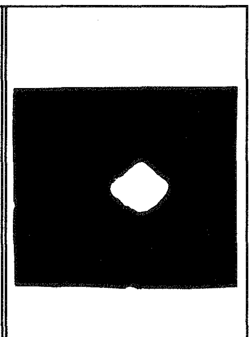


Figure 25. Result of new building extraction from the SPOT image of the same area as Figure 23 but of 10 m pixel resolution.

5. CONCLUSION

The map data used in this study provided the initial threshold values for image segmentation. The map data also demonstrated the capability of detecting large changes in existing buildings in the map data which did not match well with the source images. However, this change detection failed when the change was relatively small compared to the building size.

An expert system approach permitted control of the iterations required for feature extraction and the refinement of threshold values. Transparent nature of the knowledge coded as rules in text format allowed easy access and understanding. Processing for uncertain regions and merged features was well-controlled by the expert system using the focusing mechanism. Another advantage of expert system approach was an ability to efficiently find solutions from the large descriptor space.

The experiments of this study demonstrated that two to four times smaller pixel resolution was required to achieve machine feature extractions comparable to those of human interpreters. This relationship implies the requirement for small pixels for automatic feature extraction. As the USGS 1:24,000 scale maps show buildings as small as 12 x 12 m (USGS, 1961), and an original image pixel resolution of about 5 m is necessary for human interpreters to extract these buildings, the pixel resolution required for automatic feature extraction will be on the order of 2.5 to 1.25 m.

According to the estimation by Light (1986) the optimum pixel resolution for a cartographic database for 1:24,000 topographic maps and digital gray-scale images is about 2.0 m. Hence, it will be possible to extract most buildings required for the maps from the images in such cartographic databases using automatic feature extraction. However, there are other map features with smaller dimensions than buildings, e.g. narrow roads and creeks. For these features, smaller pixel resolution images may have to be resampled from the images in the database. Consequently, automatic extraction of such small features will require larger data storage and longer processing time.

ACKNOWLEDGEMENTS

Authors would like to acknowledge the use of the SPOT image data employed in this article. These data are copyrighted by CNES, Toulouse, France.

REFERENCES

Bajcsy, R., and M. Tavakoli. 1976. "Computer Recognition of Roads from Satellite Pictures," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-6, No.9, pp. 623-637.

Buchanan, B.G., and E.H. Shortliffe. 1984. Rule-Based Expert Systems, Addison-Wesley Publishing Co., Reading, MA.

Fischler, M.A., J.M. Tennenbaum, and H.C. Wolf. 1981 "Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique," Computer Graphics and Image Processing, Vol. 15, pp. 201-223.

Hanson, A., and E. Riseman. 1988. "The Visions Image Understanding System," Advances in Computer Vision : Vol. 1, Ed. by C. Brown, Lawrence Erlbaum Associates, Hillsdale, NJ.

Huertas, A., and R. Nevatia. 1988. "Detecting Buildings in Aerial Images," Computer Vision, Graphics, and Image Processing, Vol. 41, pp. 131-152.

Light, D. L. 1986. "Planning for Optical Disk Technology with Digital Cartography," Photogrammetric Engineering and Remote Sensing, Vol. 52, pp. 551-557.

Lillesand, T.M., and R.W. Kiefer. 1987. Remote Sensing and Image Interpretation, John Wiley and Sons, Inc. New York, NY.

Matsuyama, T., and V.S-S. Hwang. 1990. SIGMA: A Knowledge-Based Aerial Image Understanding System, Plenum, NY.

McKeown, D.M., W.A. Harvey, and J. McDermott. 1985. "Rule-Based Interpretation of Aerial Imagery," IEEE Trans. Pattern Recognition and Machine Intelligence, Vol. PAMI-7, pp. 570-585.

Murakami, H. 1990. Automatic Feature Extraction for Map Revision, Ph.D. Dissertation, Department of Geography, University of Georgia, Athens, Georgia, 236p.

Nagao, M., and T. Matsuyama. 1980. A Structural Analysis of Complex Aerial Photographs, Plenum, NY.

Nevatia, R., and K.R. Babu. 1980. "Linear Feature Extraction and Description," Computer Graphics and Image Processing, Vol. 13, pp. 257-269.

Paine, D.P. 1981. Aerial Photography and Image Interpretation for Resource Management, John Wiley & Sons, New York, NY.

USGS. 1961. "Building and Urban Areas," Chapter 3A2 of Topographic Instructions of the United States Geological Survey, U.S. Geological Survey, Reston, Virginia.

Wang, F., and R. Newkirk. 1988. "A Knowledge-Based System for Highway Network Extraction," IEEE Trans. on Geoscience and Electronics, Vol. GE-14, pp. 37-44.