MAP/IMAGE CONGRUENCY EVALUATION KNOWLEDGE BASED SYSTEM

Gordon W. Plunkett and David G. Goodenough
Canada Centre for Remote Sensing
Ottawa, Ontario, Canada

Morris Goldberg
University of Ottawa
Ottawa, Ontario, Canada

ABSTRACT

A Knowledge Based System (KBS) for analyzing LANDSAT MSS images and comparing this analysis to corresponding geocartographic data is presented. This paper discusses the preprocessing requirements for the LANDSAT and the geocartographic data for a uniform representation of the data. The segmentation of the LANDSAT data and the interpretation of the segments are presented. The preprocessed data are read into the Map/Image Congruency Evaluation (MICE) KBS where the image segments are classified and then compared with the map data, based on class, segment size, shape, and location. Results of the map/image congruency analysis are output and converted to image form. This paper presents the MICE KBS and reviews the results generated for a LANDSAT MSS scene of the Prince George area of British Columbia.

KEYWORDS: LANDSAT, computer cartography, image analysis, artificial intelligence, knowledge based systems

SUMMARY

Remotely sensed data, particularly from the LANDSAT series of satellites, are being used for a wide variety of useful applications. One of the more challenging applications is the data integration of remote sensing data with existing cartographic data bases. It was found that simple algorithmic data integration methods did not provide satisfactory results due to various geometric irregularities in the remote sensing data and in the cartographic data. These spatial irregularities could be due to factors such as temporal differences between the data, spatial errors in the map data or topographic effects in the remote sensing data. Algorithmic techniques break down in this data integration [BILLINGSLEY82]. Therefore, we have tried to solve the integration problem with a knowledge based system approach.

The Map/Image Congruency Evaluation (MICE) knowledge based system was developed to study the spatial differences between maps and images. Map/-image congruency evaluation means the determination of the spatial agreement of features in the map with the corresponding feature in the image. The data integration problem study included three basic operations. These operations were: (i) preprocessing the data for uniform representation of both the image and the map data; (ii) spatial reasoning on the data using the PROLOG-based MICE system; (iii) output of a congruency evaluation map from the results of the MICE analysis.

The MICE system was evaluated using LANDSAT MSS data for the Prince George area of British Columbia and a BC provincial forest cover map. Various levels from the BC digital map were selected. These levels corresponded to single-line creeks and rivers; double-line rivers and lakes; road and utility systems; and the forest cover. Each level was gridded to a 50x50 metre grid. The LANDSAT data were geocoded by the Canada Centre for Remote Sensing (CCRS) Digital Image Correction System (DICS) to a UTM coordinate grid with 50x50 metre pixels. The sub-area of the image corresponding to the map was selected.

The LANDSAT image was then segmented to highlight the various features. Numerous properties such as the segment shape, size, location and spectral means were evaluated. The map data were similarly processed to determine properties such as shape, size and location. These data were then read into the knowledge based MICE system.

The MICE system then evaluated the matching of the various segments from the map and image by examining the identification of the segments and the structure of the segments. The identification of the segments was done to determine if the structurally corresponding segments have corresponding identifications. For instance, a segment that has been identified as a lake in the map data must correspond to a segment in the image that has a spectral signature that corresponds to a lake. If the LANDSAT segment does not have a corresponding spectral signature, then the segment is only weakly identified. Finally, the exact positions of the remaining segments are determined and all location differences are reported.

The MICE system, which is currently under development, uses a variety of meta-level rules and object-level rules. These rules and some of the internal workings of the knowledge based system will be given, as well as suggested enhancements.

INTRODUCTION

For many years, human photo-interpretors have been analyzing air-photos, deciding on the classification of various objects in the photo and then transcribing the classification and location of these objects onto a map or more recently into a geographic information system [ZARZYKI82]. Since this map making procedure is primarily a human endeavour, it is prone to human error. In addition, the world land-mass is a constantly changing entity. For example, rivers meander further, forests burn or are cut, and subdivisions and roads are built. The cartographic data on the other hand, is relatively static and is only updated periodically.

For some time, the remote sensing community has been extoling the virtues of the integration of remote sensing data with Geographic Information Systems (GIS). This data integration problem has been researched and solutions developed, which are used operationally by some agencies [HEGYI83]. However, the automatic integration of remote sensing data with geographic information systems is not yet possible as it still requires human interpretation and assistance.

One of the first steps in the integration of remote sensing data with GIS data is simply to evaluate how similar or different are the map and image data. It has been shown that algorithmic techniques such as differencing and correlations simply don't work very well [PARSONS84] [GOODENOUGH85]. Also many rule based systems for image interpretation have shown promising results [MCKEOWN85]. Thus, a knowledge based system for the comparison or congruency evaluation of maps and images was developed.

The MICE system was developed on a VAX 11/780 system running VMS. The VAX system hosts a suite of software from Intergraph Corp. for processing and manipulating cartographic data. Also, the VAX hosts a large suite of image processing software, that was developed in VMS Fortran at CCRS. It was decided that the existing software base be used for some of the processing programs. Additional processing and reformatting programs were developed in Fortran and the results fed into the MICE KBS. MICE itself, was written in M-PROLOG from Logicware Inc.

The Prince George area of British Columbia was selected as a test area as a number of data sets from a variety of sources were available. The digital cartographic data from the BC Ministry of Forests was obtained. These data contained the hydrography, cadastral, forest, roads, railroads and other cartographic information required for a forest cover map. The map scale was 1:20,000 and corresponded to the UTM map number 93G096.

The LANDSAT MSS geocoded image for the Prince George area (93G15) was obtained from CCRS. This DICS product [GUERTIN81] consists of the LANDSAT data scaled and projected onto a 50 metre grid, in a UTM projection.

GIS PREPROCESSING

The BC forest cover map was received and stored as an Intergraph design file. This file contained a variety of cartographic information, but for the purposes of this experiment, only the following information was processed:

Information	File Level
a) single-line creeks, rivers	5
b) double-line rivers, lakes	6
c) utility systems	8
d) forest cover typelines	9

The levels were extracted and any spurious information or text was deleted. Each level was edited using an automated technique for ensuring that all line intersections were cartographically sound. Next, the levels were individually converted from vector format to grid format, based on a presence/absence algorithm onto a 50x50 metre grid. These grid files were then converted to CCRS standard imagery files. Each polygon (such as a lake), which was not fully filled was filled. The image was precision registered to a UTM grid and each entity of the map was identified uniquely and its location was run length encoded. Finally, each unique element, along with its run length encoded location was converted into symbolic object form. The file containing these symbolic objects were read into the MICE system. The procedure for preprocessing the cartographic data is given in Figure 1.

IMAGE PREPROCESSING

The LANDSAT MSS image (frame Id: 50458-18360) used for this experiment was imaged on June 25, 1985. The MSS image was then precision geocoded on the CCRS DICS system. The area of the image corresponding to the BC forest cover map 93G096 was extracted and precision registered to the rasterized 93G096 map data. The MICE system then requires the segmentation and statistical analysis parameters of these segments.

The image subscene is first operated on by a specified gradient operation. The resulting file is then segmented [BOILEAU85]. Each segment is uniquely identified and its location run length encoded. Next statistical information on each segment is generated. This statistical information is shown in Table 1. The segment locations and segment statistical values are converted into symbolic object format for input to the MICE KBS. The procedure for preprocessing the image data is given in Figure 2.

MAP/IMAGE CONGRUENCY EVALUATION KNOWLEDGE BASED SYSTEM

The map/image congruency evaluation knowledge based system is implemented in PROLOG using a shell for developing hierarchical expert systems for remote sensing [GOLDBERG85] [BRAUN85]. The implementation is primarily divided into two rule types. These are the meta-rules, which are rules about what MICE should do next, based on information deduced to that point. The other type of rule is the object rule, which is a rule that has been input to the MICE system, or has been deduced by the MICE system.

Meta-Level Rules:

The meta-rule consists of four items. These items are: 1) condition predicate; 2) action procedure; 3) phase number; 4) rule number.

FIGURE 1:

THE PROCEDURE FOR PREPROCESSING THE CARTOGRAPHIC INFORMATION

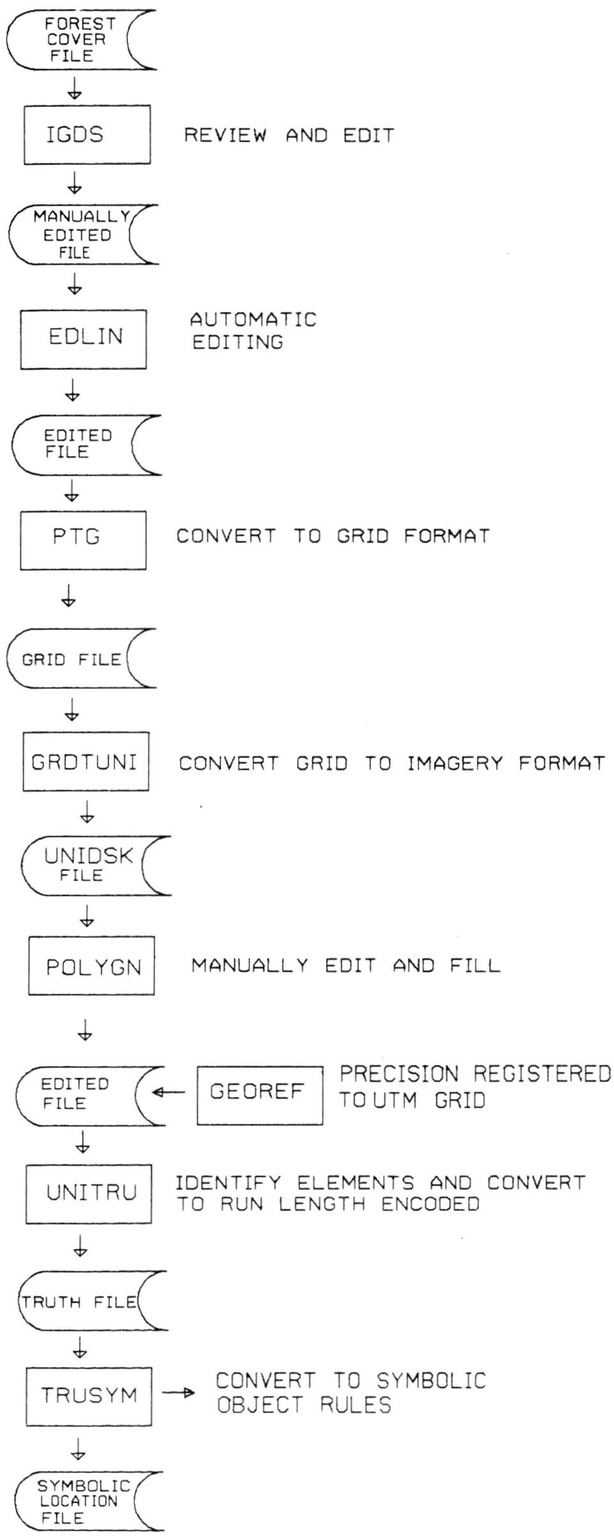
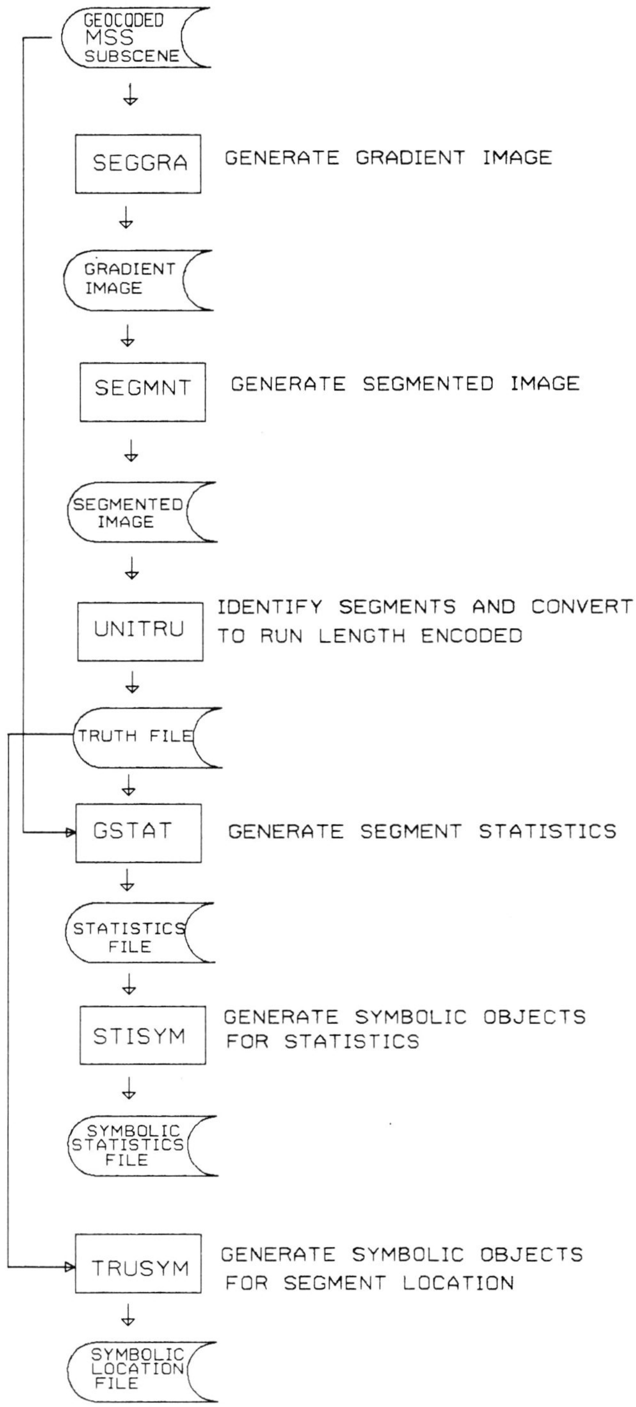


FIGURE 2:

THE PROCEDURE FOR PREPROCESSING LANDSAT MSS IMAGE



The condition predicate (or "if" part of the rule) is evaluated by MICE to determine if the condition predicate is true. The action procedure (or "then" part of the rule) may then be executed if the condition predicate is true. The phase number is the strategy level within the meta-level procedure in which this rule is to be evaluated. The rule number is simply to uniquely identify each rule in the meta-level procedure. An example of two meta-level rules for one phase from MICE is as follows:

if: the image segments were identified (ok)
then: compare map and image segment sizes
phase: 9
rule #: 17
if: not (the image segments were identified (ok))
then: write the string ("no image segments identified")
phase: 9
rule #: 18

Object Rules:

Object rules also consist of four items. These are: 1) condition predicate; 2) action procedure; 3) rule number; 4) certainty factor. These rules deduce object values based on the values of the objects in the condition predicate. The rule number uniquely identifies the rule number and the certainty factor is a value from 0 to 100.

Objects:

Objects are the basic manipulation element of MICE upon which deductions are made. Objects consist of four values, which are: 1) object context or description; 2) object attribute; 3) object value; 4) measures of belief and disbelief. MICE uses context values such as: source (image-MSS), source (map-bcfs), segment (segment-number) and class (class-name). Attribute values such as location, size, mean-channel, and shape etc. are used with the corresponding value of the attribute in the object element. The measures of belief and disbelief for each element are included. The measures of belief and disbelief range from 0 to 100. A belief/disbelief value of 100 means that this object is very believable/disbelievable. A smaller value indicates less belief/disbelief in this object. A sample object element is as follows (in PROLOG notation):

obj([[*, source (image-MSS), *, segment (2), *, class (hydrography), *], size, [2160], [75,25]]).

This Prolog statement means:

1) The source of the segment is the LANDSAT MSS image.
2) The segment is segment number 2.
3) The segment has been classified into the class hydrography.
4) The attribute of this object element is the size of the segment.
5) The attribute value or the size of the segment is 2160 pixels.
6) The measure of belief for this object is 75 and the measure of disbelief is 25.

A simplification of the agenda that MICE uses to perform the congruency evaluation is as follows:

1) load the map object elements
2) load the image object elements
3) perform preliminary classification on image objects
4) get the next map segment
5) find all image segments near map segment (focus)
6) compare class values
7) compare segment sizes
8) compare segment shapes
9) compare segment locations
10) output results

The output generated by MICE is in the form of object elements, that were deduced by the KBS. These object elements indicate where the map and image segments overlap, where they are partitioned, where they are hierarchical and where they are bipolar. These elements are then converted from symbolic form to run length encoded format. Finally the results are converted to imagery format, which can be displayed and reviewed.

SAMPLE OUTPUT

An experiment using the MICE KBS was performed using LANDSAT MSS and digital map data from the BC Ministry of Forests. The sample outputs are for the double-line rivers and lakes data from the BC map. The figures 3 to 10 show the input and the output from various phases within the MICE system. They all correspond to the processing of the map or image data from figure 3 or figure 5, respectively.

Figure 3 shows the input LANDSAT MSS image for band 7 (infrared 0.8μm to 1.1μm) for the Prince George area of BC. The image has been geocoded to the UTM projection and resampled to 50m x 50m pixels. The date of the image is June 2, 1985. Figure 4 shows the same image following the application of the Sobel gradient operator, segmentation, and grey level coding. The coding algorithm is for display purposes only. It reviews the segments and assigns each segment a digital value that ensures that no neighbouring segment has the same value (grey-level). However, non-neighbouring segments may have the same grey-level value. Figure 5 shows the input map vector data for the double-line hydrology level of a BC forest cover base map. Figure 6 shows the same map data after it was cleaned and rasterized. Cleaning means removing annotation, processing vector and points (overshoot/undershoot conditions), processing vectors against themselves (knot conditions), and processing vectors against other vectors (lobe conditions). Rasterization uses the presence/-absence algorithm.

Figure 7 shows the segments of the image that were identified as being in a map segment window (focused) and were also classified as hydrology. A focused image segment means that the image segment is "near" the map segment. The map segment window is the smallest rectangle that encloses the segment. An image segment is near (focused on) the map segment if any part of it is within the window or half the window length in any direction.

Figure 8 shows the focused image segments that are of similar size. Figure 9 shows the focused image segments with similar shape and Figure 10 shows the focused image segments which are determined to be overlapping segments. An image segment that is of similar size to a map segment satisfies the following rule

$$50 < \frac{\text{map segment size}}{\text{image segment size}} \times 100 < 150$$

An image segment that is of similar shape to a map segment satisfies the following rule

$$50 < \frac{\text{map segment shape}}{\text{image segment shape}} \times 100 < 150$$

An overlapping image segment is one where any pixel of the image segment overlaps any pixel of the map segment.

CONCLUSIONS

The results reported thus far are very encouraging for the use of knowledge based systems for performing visual tasks, such as verifying the congruency of maps and images. Obviously, this is the first step in automating the process of integrating remote sensing data with geographic information systems. Further work is required and more rules must be added to enhance the functional performance of the congruency verification procedure, but the same techniques should also apply then to the extraction of selected areas in the image and including this information in the GIS system. Future work will include experiments with LANDSAT Thematic Mapper data and federal topographic maps.

REFERENCES

[BILLINGSLEY82] Billingsley, F.C., "Edited oral presentation", in Proceedings of the NASA Workshop on Registration and Rectification, Bryant, N.A. (ed), NASA-JPL, 1982.

[BOILEAU85] Boileau, J.M., Goldberg, M., "A Selection of Segmentation Similarity Measures for Hierarchical Picture Segmentation", Graphics Interface, p.179-186, Montreal, May, 1985.

[BRAUN85] Braun, F., University of Ottawa, MASC Thesis, 1985.

[GOLDBERG85] Goldberg, M., Goodenough, D.G., Alvo, M., and Karam, G.M., "A Hierarchical Expert System for Updating Forestry Maps with LANDSAT Data", Proceedings of the IEEE, Vol.73, No.6, June, 1985.

[GOODENOUGH85] Goodenough, D.G., Fung, K.B., Hegyi, F., Robson, M., and Swanberg, N.A., "Integration of Geographic Information Systems with LANDSAT Thematic Mapper Data", paper presented at IGARSS'85.

[GUERTIN81] Guertin, F.E., Shaw, E., "Definition and Potential of Geocoded Satellite Imagery Products", Presented at the 7th Canadian Symposium on Remote Sensing, Winnipeg, Manitoba, Sept. 1981.

[HEGYI83] Hegyi, F., Quenet, R.V., "Integration of Remote Sensing and Computer Assisted Mapping Technology in Forestry", Paper presented to The Canadian Institute of Surveying Congress, Victoria, B.C., 1983.

[MCKEOWN85] McKeown, D.M., Harvey, W.A., McDermott, J., "Rule-Based Interpretation of Aerial Imagery", PAMI-7, No.5, Sept. 1985.

[PARSONS85] Parsons, T.J., "Towards Robust Image Matching Algorithms", SPIE, Vol.504, Applications of Digital Image Processing, VII (1984).

[ZARZYCKI82] Zarzycki, J.M., and Allam, M.M., "Canadian Council on Survey and Mapping - National Standards for the Exchange of Digital Topographic Data", Topographical Surveys Division, Surveys and Mapping Branch, April, 1982.

TABLE 1: Image Statistical Information Generated for MICE Input

1) segment size (number of pixels)
2) segment MSS channel 1 (band 4) mean
3) segment MSS channel 2 (band 5) mean
4) segment MSS channel 3 (band 6) mean
5) segment MSS channel 4 (band 7) mean
6) segment MSS channel 1 (band 4) maximum value
7) segment MSS channel 2 (band 5) maximum value
8) segment MSS channel 3 (band 6) maximum value
9) segment MSS channel 4 (band 7) maximum value
10) segment MSS channel 1 (band 4) minimum value
11) segment MSS channel 2 (band 5) minimum value
12) segment MSS channel 3 (band 6) minimum value
13) segment MSS channel 4 (band 7) minimum value
14) segment shape (perimeter^2/area)
15) segment window (smallest rectangle containing segment)

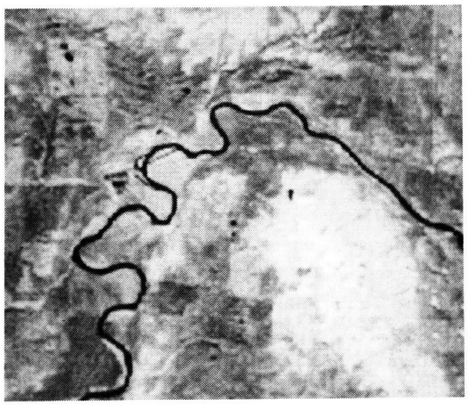


FIGURE 3 LANDSAT MSS band 7 image for Prince George area BC

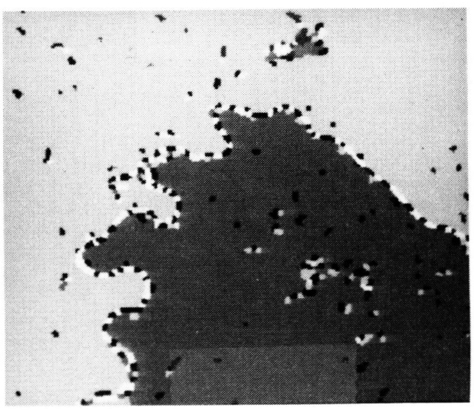


FIGURE 4 LANDSAT image following Sobel gradient operator, segmentation and coding for display. This image corresponds to the LANDSAT image of Figure 3.

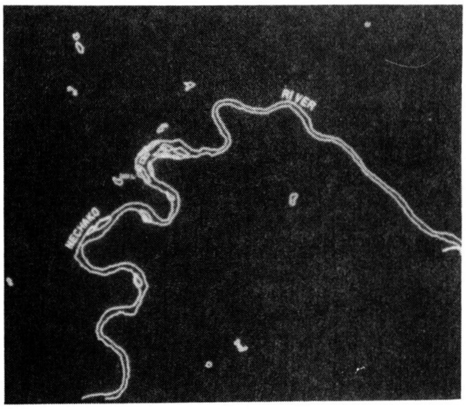


FIGURE 5 BC Ministry of Forests base map for doubleline hydrography vector data.

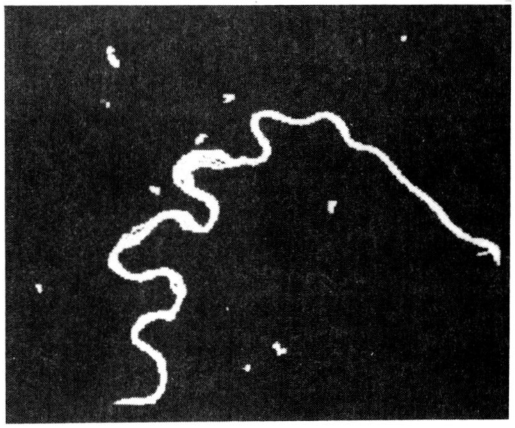


FIGURE 6 Map data following cleaning and rasterization of vector data from Figure 5.

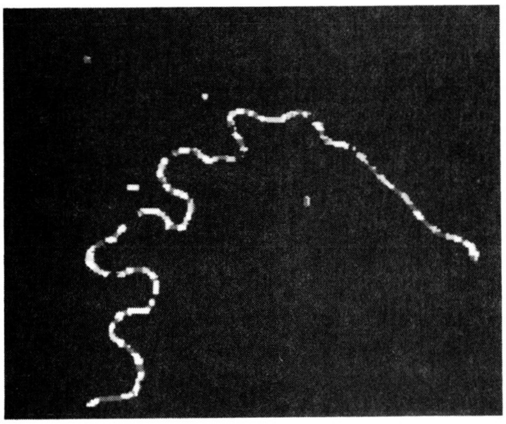


FIGURE 7 Image segments (from Figure 4) that were classified as hydrography and are near any map segment.

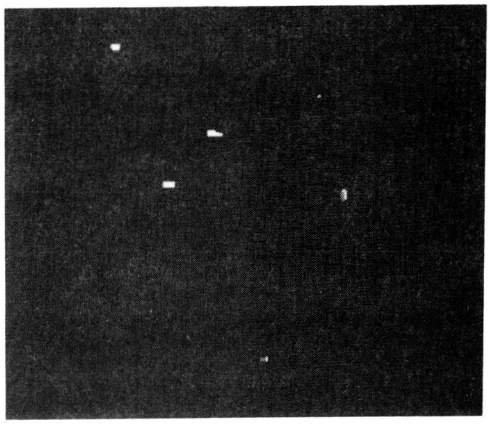


FIGURE 8 Image segments (from Figure 4) that were of similar size to any map segment that was near.

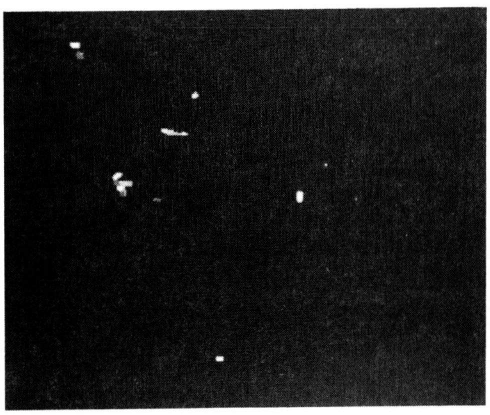


FIGURE 9 Image segments (from Figure 4) that were of similar shape to any map segment that was near.

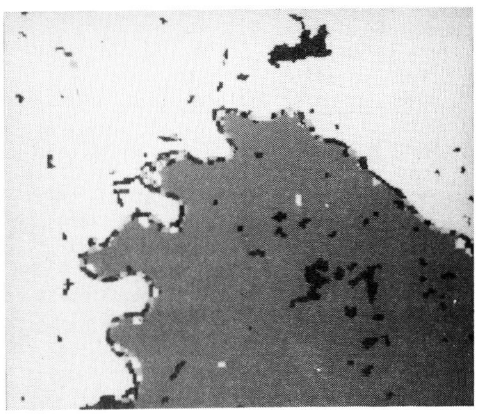


FIGURE 10 Image segments (from Figure 4) that were overlapping any map segment.