Conference on

Machine Processing of

Remotely Sensed Data

October 16 - 18, 1973

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

Copyright © 1973
Purdue Research Foundation

This paper is provided for personal educational use only,
under permission from Purdue Research Foundation.

A COMPUTER RECOGNITION OF BRIDGES, ISLANDS, RIVERS AND LAKES FROM SATELLITE PICTURES

Ruzena Bajcsy and Mohamad Tavakoli

The Moore School of Electrical Engineering
University of Pennsylvania
Philadelphia, Pennsylvania

ABSTRACT

This paper describes a computer program which recognizes bridges, rivers, islands and lakes from satellite pictures. The program is structured into three basic parts: the world model, the low level operators, and the higher level program. The recognition process is conceived as a process of continuously refined verification of the hypothesized descriptions of objects. We use conceptual identification of objects during the recognition process as soon as we can; we equip these concepts with meanings in the three-dimensional world. We present several concrete examples as a demonstration of the capabilities of our program.

1. INTRODUCTION

The purpose of this paper is to describe a computer program that can recognize bridges, rivers, lakes, and islands occuring in scenes that are captured by satellite pictures.

The motivation for this work is the need to automatize the recognition process of hundreds of pictures obtained by satellite photography. People are better performers than machines in separating relevant from irrelevant information in a picture, especially in complex visual situations; however, if the amount of analyzed pictures exceeds a certain (rather small) number, people tire and their recognition capabilities rapidly decrease, while the machine maintains a consistent capacity. Thus, the task is to design a computer system which behaves "intelligently" during the picture recognition process, creating scene descriptions which approximate human descriptions (although not necessarily identical!). By the word "intelligently" we refer, in particular, to the following capabilities of the computer system; the ability to distinguish the relevant and irrelevant visual information in a given context, the ability to interpret the objects and their relationships seen in a two-dimensional picture as three-dimensional (real) objects, and the ability to make inferences about the objects and their relationships, even from incomplete visual data. These are the main problems that we shall concentrate on in this paper.

Historically, the use of pattern recognition methods produced many interesting results in the recognition of agricultural areas, vegetation, etc. obtained from satellite pictures (See the review in Nagy, 1972). However, these methods neglect some very important spatial features such as spatial relationships and the fact that the recognition process is essentially an interpretation of a 2-dimensional picture within the 3-dimensional world. The usefulness of spatial relationships has been presented recently in a paper by Fischler and Elschlager (1973).

In work (Bajcsy, 1972) we have suggested a possible approach to computer recognition of real outdoor scenes; here, we shall further develop those ideas in the context of the world seen from a satellite.

2. THE DESCRIPTION OF THE SYSTEM

The organizational scheme of our approach is displayed in Figure 1.

2.1 INPUT DATA

The satellite pictures are given in a digitized form with spatial resolution, one point corresponding to 4503 m^2 on the earth. Moreover, each scene generates 4 pictures corresponding to 4 different spectral bands. The amount of light reflected in each band is represented at each point by its gray value (we use 128 possible gray values). Under ideal conditions, different materials, such as water, soil, vegetation, concrete, etc., should be distinguishable in several bands by their different spectral signatures. The signature values are tabularized in tables according to material, such as: water, dry soil, vegetation, and others (Leeman, 1972). However, the amount of reflectance is a function of several parameters, such as the differences in material, the relative amount of sky light present, the altitude of the sun, light scattered by the material volume into the atmosphere. In the case of water, the depth of the watery area and the pollution could confuse the expected reflectance. In the case of relatively small objects, such as bridges and roads, the surrounding environment changes greatly the otherwise expected reflectance. The reasons for the influence of the environment on small objects with respect to their absolute reflectivity are twofold:

a) The given resolution of satellite pictures, i.e., the gray value of each point on the picture, corresponds to the integral of reflectivity over an area 57 x 79 m^2.

b) The picture is a two-dimensional projection of the three-dimensional surface. Thus, in addition to the reflectance of a material, a depth information is encoded (though in an unknown fashion) in every gray value of the picture.

Finally, we mention the flactuation noise of the measured reflectance, that is, the gray values, which are due to the noise of the sensor, as measured by the multi-spectral scanner subsystem installed on the satellite. This noise is between 0.3% and 0.4% of the full scale radiance. The radiance is linearally proportional to the gray value in respective bands (Data User's Handbook, NASA, 1972). From this information, it follows that our expected error is less than one unit of the gray scale. Thus, certain values of gray scale can suggest the existence of some particular material in the picture, but they do not by any means form sufficient criteria for recognition purposes. What one has to do is to extract more geometric features and perhaps use some additional knowledge in order to be able to make the proper interpretation of observed objects.

2.2 LOW LEVEL OPERATORS

The low level operators are procedures that extract from the input data the relevant features and descriptors, such as texture, shape, size, etc. Currently, we use the following low level operators:

A global region finder and a local region expander (both of these operators include a homogeneous texture identifier).
Size measure (the area measure and the length of the boundary).
Location measure (the coordinates of initial and end points of a boundary, of the center of gravity).
Shape descriptor (skeleton operator and the boundary description).
Relationship identifiers: a) connectivity in a given direction
b) surrounded by

In what follows, we shall describe in detail each of the above mentioned operators. The global region finder uses the region growing algorithm described in Bajcsy (1973). This algorithm assumes:
Definition of a local structure and an equivalence relationship on two adjacent (overlapping) structures. The local structure can be texture, color, and/or brightness values. In this work, we have implemented only an homogeneous texture operator, which is no more than a spatial low pass filter. The equivalence rela-

tionship is a threshold function that partitions the picture into mutually exclusive regions. The threshold is derived from signature tables as the most conservative estimate of the expected reflectance for a given material. The strategy here is to begin with a material that sharply contrasts in brightness in relation to everything else in the picture. In our case, the water in band 3 has this property.

The local region expander assumes that the global region finder has already been applied; therefore, the regions to be expanded are marked by "M."

Then the algorithm is implemented in two steps:

First, we define a neighborhood of a point (i,j) as shown in Pfaltz and Rosenfeld (1967). Let us denote: NTH is the conservative threshold set in the global region finder; CX is the relaxation coefficient representing the amount of variation of NTH.

a(i,j) is the gray value of a point (i,j).

Secondly,

If $a(i,j) \leq NTH + CX$,

and there exists a point (k,l) such that it is within the neighborhood of (i,j) and $a(k,l) \leq NTH$,

THEN mark point (i,j) by "M."

We use two kinds of Shape Descriptors: skeleton operator and the boundary descriptor.

For the skeleton operator, we have basically implemented the skeleton operator described in Pfaltz and Rosenfeld (1967) with a small addition. Rosenfeld's algorithm derives the skeleton as a maximum of minimal distances from the boundary of a region. Since our pictures are just windows from the scene, the boundaries imposed by windows are artificial ones. This fact leads us to ignore the window boundaries for the skeleton purposes.

The skeleton technique is very suitable for recognition of thin and elongated shapes. However, if the shape is very irregular and complex (part elongated and other parts convex-like with different radius), then additional shape measures are necessary.

The boundary operator first detects the boundary and then it follows the boundary. In addition, the operator records:

a) the closed boundary, or,
b) the almost closed boundary (with a small distance between the initial and end points), or,
c) the open boundary,
d) in all cases (a, b, c) it stores the initial point and the end point of a boundary.

The size and the location measures are simple operators and their implementation follows directly from their description.

In this paper we describe two spatial relationship operators. The connectivity relationship in a given direction is very simple. One looks for adjacent points in a given direction that have similar gray values.

The second relationship, "surrounded by," is implemented in the following procedure:

Find all regions (R), in approximation convex and simply connected, which satisfy the following conditions:

(1) The neighbor points of region R are all first adjacent points to region R, which do not belong to region R.
(2) Region R has neighbor points with expected gray values of the surroundings.

We could think of several other useful low level operators, such as a more detailed and elaborate shape and/or size recognizer, a spatial high pass filter, more elaborate texture operators, and others. However, in our limited context, we did not need them to this point.

2.3 THE WORLD MODEL

The world model is a network where each node represents a description of objects that one expects to encounter in the scene. The arcs correspond to relationships between the objects.

The description of objects occurs on two levels: (1) conceptual (river, lake, bridge, road, city, etc.) and (2) quantitative (in terms of measurable features such as shape, size, color, texture, etc.). The relationships can be spatial (above, left, right, next to, etc.), topological (continuous, proximal positions, connected, etc.), and quantitative (larger, smaller in size, darker, lighter in color, etc.).

Our world model at present consists of five types of objects: rivers, lakes, bridges, land, and islands (see Figure 2).
The objects are further described as follows:

Water: Gray value (brightness) in band 3 $\leq$ 10% of the maximum gray values. In this band watery areas form an homogeneous texture.

Land: Gray value (brightness) in band 3 < 10% of the maximum gray values. The texture and shape are not important.

Rivers: Gray value of the water
Texture: homogeneous
Boundaries: open
Contrast: large
Spatial relationships to bridge: below
Topological relationships: continuous
Spatial relationships to land: surrounded by

Lakes: Gray value of the water
Texture: homogeneous
Boundary: mostly closed; if open, then lakes are connected to the river.
Contrast: large
Spatial relationship to land: surrounded by
Spatial relationship to island: surrounding

Bridge: Gray value of the land
Texture: homogeneous
Shape: thin, elongated, and smoothly curved
Relationship to land: connected to its two shorter sides
Relationship to water: surrounded by on its two longer sides

Island: Gray value of the land
Texture: homogeneous
Boundary: closed
Spatial relationship to water (lakes or rivers): completely surrounded by
Spatial relationship to land: must not be connected

Given this model it is clear what kind of objects we have in mind. However, we can also see where the model will miss certain objects in recognition. For instance, we will not recognize bridges which are built over valleys because they are not surrounded by water; on the other hand, we will recognize as a bridge a narrow piece of land which separates a river from a lake. Clearly, this is a natural bridge as opposed to a man made bridge. In this sense our program does not distinguish between these two kinds of bridges. We shall misidentify all ships, constructions in the water as either islands or peninsulas, depending upon whether they are attached to the land or not. We will make mistakes on recognition between the lakes and rivers, if the window cuts (thereby introducing artificial boundaries) into the watery area in such a way that it causes confusion for shape and/or boundary recognition. An example is Figure 12, where using just the picture itself, it is unclear whether the watery area represents a part of a river or of a big lake.

In spite of the above mentioned difficulties, we can use for most of the pictures several of these features for identification. In general, one can say that the more features that we can measure and test, the more accurate our procedure will be (Poppelbaum, 1971). One wishes, however, to have some quantitative measure of the correctness of the procedure. For that purpose we would need a large sample of the recognized windows, compared with human recognition, a process which is currently impossible for us.

2.4 THE HIGHER LEVEL PROGRAM

The higher level program organizes the whole recognition process. It decides when and which low level operator will be applied and matches the newly obtained descriptors with the world model. After the matching process, the higher level program comes up with a hypothesis about the object and decides what other measures (if any) to consider for verification of the hypotheses.

In our case the higher level program uses the following algorithm:

1. Find the watery areas based on the gray scale value estimates for water in band 3. (Apply the global region finder) If there is no watery area, then stop the program.
2. Apply the local region expander and find the missing points where the water could continue.
3. Apply the skeleton operator in order to find the shape of the land areas.
4. Find the thin elongated areas of the land; go to the world model and make an hypothesis about the object (in this case, bridge). If we do not find any elongated areas, then go to step 11.
5. If the object is an hypothetical bridge, then verify it by checking the connectivity between the hypothetical bridge and the land on its two shorter sides. If this condition is satisfied, then go to step 7; otherwise, go to step 6.
6. Apply the connectivity operator in a given direction to check whether the previously missing connectivity is just due to a threshold error, or is an indication that the object is not a bridge. In any case, go to step 7.
7. Verify the bridge by checking the relationship "surrounded by" the water on its two longer sides.
8. If all the evidence, after steps 5, 6, and 7 shows that the hypothesis of a bridge has been verified, then mark all of the bridges on the picture and continue to step 9. Otherwise, discharge the hypothesis about the bridges and form new ones which, in this case, are small islands or peninsulas. From here continue in step 11.
9. Make a list of all recognized bridges associated with their starting and end point coordinates for the output purposes.
10. Remove all the bridges from the picture.
11. Find the water regions which have closed boundaries; if there is no such region, go to step 12. Check the world model for the possible hypothesis. In this case, the only possibility is the lake; therefore, name the regions as lakes, and make a list of lakes associated with the area measure and the coordinate of the center of gravity.
12. Find the land regions which have closed boundaries. If there is no such region, go to step 14. Check the world model for the hypothesis. In this case, the only possible interpretation is an island.
13. Verify the hypothetical islands by checking that the region is a land (and not water).
14. If a watery region has two open boundaries and the width between the boundaries is in a certain threshold, then it is hypothesized that this region is a river; otherwise the region will be hypothesized as a possible candidate for part of a river or part of a lake or ocean.
15. Stop.

The above process is an example of a possible higher level program. Indeed, one can change the calling sequence of the low level operators and of the world model for verification purposes. We have chosen this way because it seems to work the most effectively on the given data.

Nevertheless, we can suggest some improvement in our higher level program. For instance, the sequence of the recognized objects can be given by the best contrast in all bands. The band with the best contrast is chosen manually. This can be automatized by making statistical measurements on the gray levels from all bands and taking the band which has the highest contrast (similar considerations can be made for other parameters besides contrast). The next thing is to find the edge value which separates the two contrasting ranges of gray values. This value (its most conservative estimate) will be the threshold used in the global region finder.

The interpretation of the gray value, in terms of water, land, etc., must be found in the world model.

2.5 THE OUTPUT

The output is a description of the scene in terms of the names of the objects and their relationships. We do not expect any elaborate English sentences. Rather, we have a list of objects like River 1, River 2,...Lake 1, Lake 2,...Bridge 1, Bridge 2,.... Each object is associated with some parameters, such as the area measure, the length of the boundary, the coordinates of its center of gravity, the coordinates of initial points, end points, etc. A complete set of parameters for each object is shown in Table 1. The relationships between the objects will be described in the output as well. An example of such a description is "Bridge 1 is above River 2," etc.

3. THE RESULTS AND THEIR EVALUATION

In our experiments we use a digital picture (thanks to the courtesy of IBM in Gaithersburg, U.S.A.) of the Chesapeake Bay (Washington, D.C. area).

What follows next is that we present the experimental results from one window (in Washington, D.C. - Potomac River) after almost each step in the higher level program.

In Figure 3 we show the result after step 1 and Figure 4 after step 2. The stars in both figures represent the water. Notice that after the second step the watery regions are more complete than just after the first step. Figure 5 (the stars) displays the thin, elongated areas of the land (steps 3 and 4). The output after steps 5, 6, 7 and 8 is in Figure 6. The names of the bridges were put in manually. During the process of verification of bridges, we lost a portion of the Theodore Roosevelt Bridge. The reason for this is that the bridge continues over a land which is beyond our definition of a bridge. Once we recognize the bridges, we remove them (steps 9 and 10). See Figure 7. Finally, the end result of the recognition (completing steps 11 until 15) is shown in Figure 8. The stars denote the recognized bridges. The number "1" denotes all points which belong to the land. The blank areas represent the water. The names of bridges and lakes were inserted manually. Our program recognized not only all of the bridges that we expected, but also some of the bridges whose sizes were below our spatial resolution. The explanation of this fact goes as follows:

a) Let us recall that the gray value of each point corresponds to the integral over an area 57 x 79 m^2.
b) If an area happens to be a mixture of a watery area and a bridge, then this will show in the corresponding gray value of those points in such a way that the gray values will be greater than those of a water but less than those of a land.
c) Since our recognition is not based purely on threshold values, we can afford to have rather relaxed thresholds and thus pick up points which lie on the boundary of the gray scale of the water and land.

Because of the above points, the sizes of recognized bridges are only estimates of the real sizes.

The program distinguishes the lakes from the river. However, it described two watery spots on the picture as lakes while, in reality, one of them is part of a lake - the Georgetown Reservoir - and the other is part of a river.

We do not describe a peninsula (which would be a natural extension of our world model) because of the difficulty in defining the shape of a peninsula and also the difficulty of recognition of the relative size between the land and the peninsula. The point is that this definition seems to us very subjective to make good use of it for computer recognition.

Until now, we have tried to explain all the successes and faults of our recognition program on the picture of the Potomac River. To demonstrate the advantage of the artificial intellignece approach, we compare the recognized bridges by our program in Figure 6 with the "bridges" obtained by only a certain bridge gray value in Figure 9. Needless to say, error in recognition appears in Figure 9.

In well shaped rivers (where the boundaries are reasonably smooth) one can use the skeleton (see Figure10) of the river for making an hypothesis about the possible flowing direction of the water. This part we have not implemented yet.

The next two figures, 11 and 12, display the final recognition of another window of the Potomac River and of a window of the Chesapeake Bay. In both pictures the star stands for the bridge points, and the "one" stands for the land points. We recognize on Figure 11 the river, two lakes (Oxon and a nameless one), one island - Goose. The picture from Chesapeake Bay shows (and our program identifies) a bridge - the Plane Jr. Memorial Bay Bridge. The dots on our display in Figure 12 show some hypothetical islands or peninsulas, but we cannot verify them.

The last picture is a very good example of the confusion in classification caused by seeing the world through a window. In this case, for example, the watery region can be part of a big river, or part of a big lake, or part of a bay. This problem can be resolved only by looking for the continuity at the adjacent windows and seeing the world more globally than just locally.

All the presented results were obtained on the computer (Spectra 70/40) in the Moore School Computer Center. The programs are written in Fortran. The time for processing one window (128 x 60 points) is around 250 sec. CPU time.

4. CONCLUSION

In this paper we presented a procedure for recognizing bridges, rivers, lakes and islands occurring in scenes. The recognition finding process is a continuous interaction between the world model and the low level operators, directed by a higher level program. The world model embodies the knowledge about the world (for instance, the fact that the river is floating contintuously in an area) as well as the observer's position, angle, etc. (in our case, the observer is above the scene) Thus, in spite of the fact that the low level operator finds two pieces of river separated (though nearby), our program will unite them due to the knowledge from the world model about continuity of a river. The region that separates a river will be interpreted as an object <u>above</u> the river. This object could be a bridge or a cloud depending on other properties (such as shape, spectral characteristics of the material, and so on).

The comparison between our findings and the map shows that we are able to find all the bridges which are within the range of the spatial resolution (even beyond it) of the picture. We correctly identify most of the lakes and rivers, unless the watery area is part of a larger lake or river or a bay. Finally, we properly find all the islands.

The limitations of our program, as we pointed out before, follow

a) From the restricted definitions of objects (examples are the bridge only above the water).

b) From the weaknesses of some low level operators (skeleton technique in a complex region).

c) From the limitation of seeing the continuous world through a restricted window; in other words, seeing the world only locally instead of globally.

Many of these limitations will be considered in our future research.

The contribution of this research is that we conceive the recognition process as a process of continuously refined verification of the hypothesized descriptions of objects. We use conceptual identification of objects during the recognition process as soon as we can; we equip these concepts with meanings in the 3-dimensional world. This fact together with the 2-dimensional data allows us to infer 3-dimensional relationships and to identify some of the new objects.

ACKNOWLEDGMENTS

The authors are indebted to Dr. H. Kritikos from the Moore School of Electrical Engineering, University of Pennsylvania, who called our attention to the problem of recognition from satellite pictures. We also acknowledge with gratitude the help obtained from Dr. Ralph Bernstein and Mr. S.W. Murphy from IBM - Gaithersburg, who supplied us with the input data.

REFERENCES

Bajcsy, R., "Computer Identification of Textured Visual Scenes," AIM-173, Artificial Intelligence Project, Stanford University, Stanford, California, 1972.

Bajcsy, R., "Computer Identification of Visual Surfaces," Computer Graphics and Image Processing, No. 6, 1973.

Fischer, M. A. and Elschlager, R. A., "The Representation and Matching of Pictorial Structures," IEEE Transactions on Computers, Vol. C-22, No. 1, January 1973.

Leeman, V., "The NASA Earth Resources Spectral Information System," N72-30345, 1972.

Nagy, G., "Digital Image-Processing Activities in Remote Sensing for Earth Resources," Proceedings of the IEEE," Vol. 60, No. 10, October 1972.

NASA Earth Resources Satellite Technology. Data User's Handbook, 1972.

Pfaltz, J. L. and Rosenfeld, A., "Computer Representation of Planar Regions by their Skeleton," Communication of the ACM, Vol. 10, November 1967.

Poppelbaum, W. J., AEC Proposal, 1969,cited in Simons, A., "Eidolyzer: A Hardware Realization of Context-Guided Interpretation," Department of Computer Science, University of Illinois at Urbana, Champaign, Urbana, Illinois, 1971.

OBJECT	AREA	STARTING POINT	END POINT	CENTER OF GRAVITY	LENGTH OF BOUNDARY
Bridge	--	X	X	--	X
Island	X	--	--	X	--
Lake	X	--	--	X	--
River	X	X	X	--	X

Table 1. Parameters of Objects

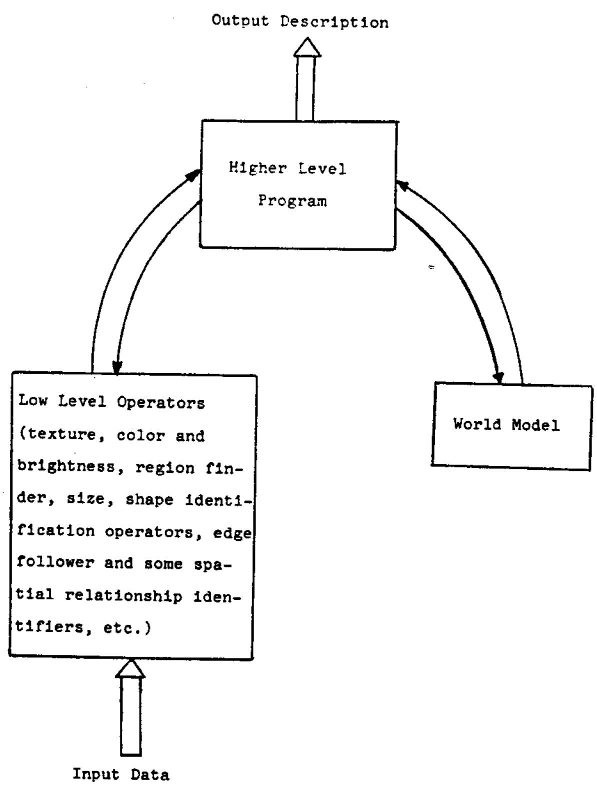

Fig. 1: Organizational Scheme

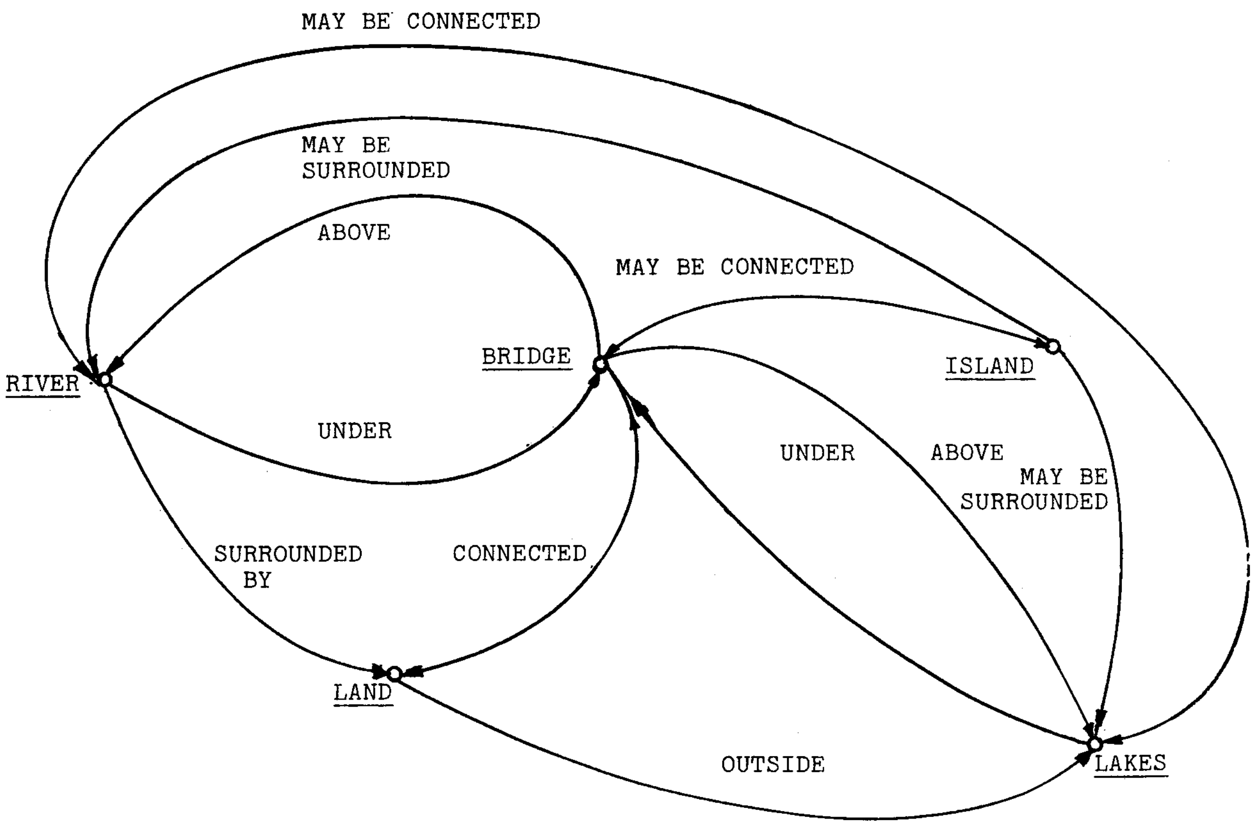

Fig. 2: The World Model

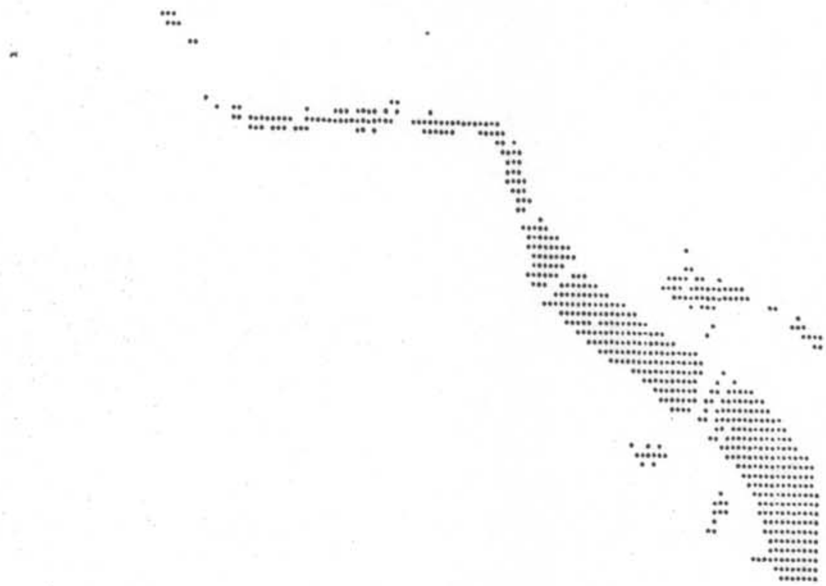

Fig. 3 Water I

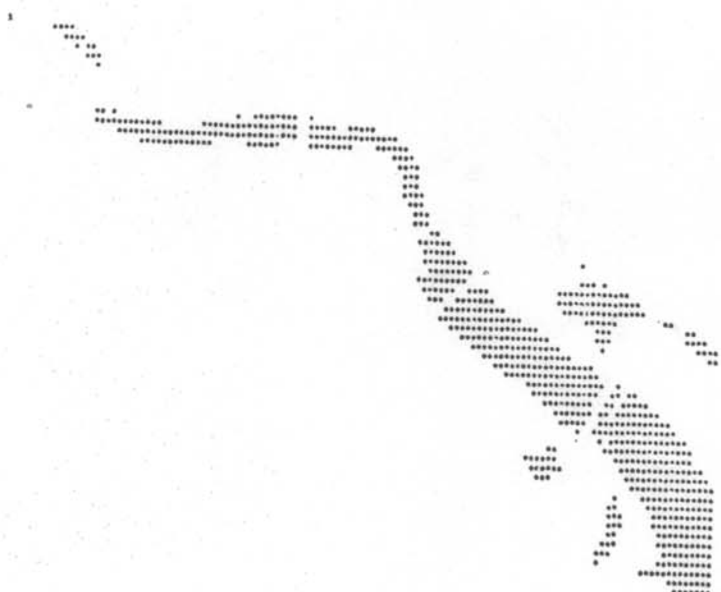

Fig. 4 Water II

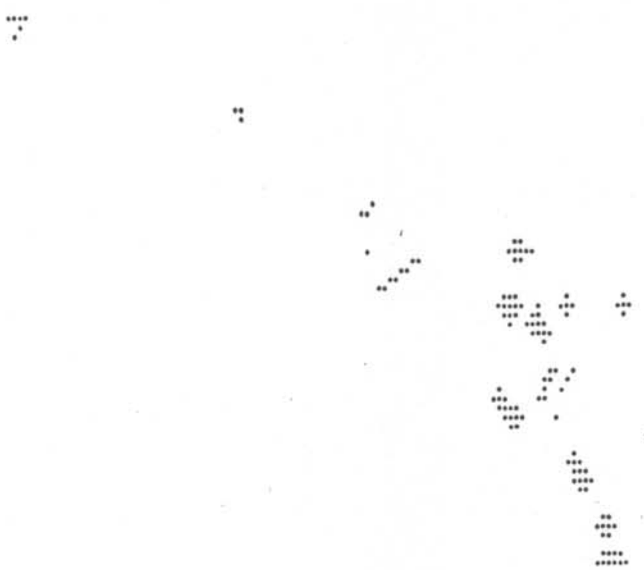

Fig. 5 Hypothetical Bridges

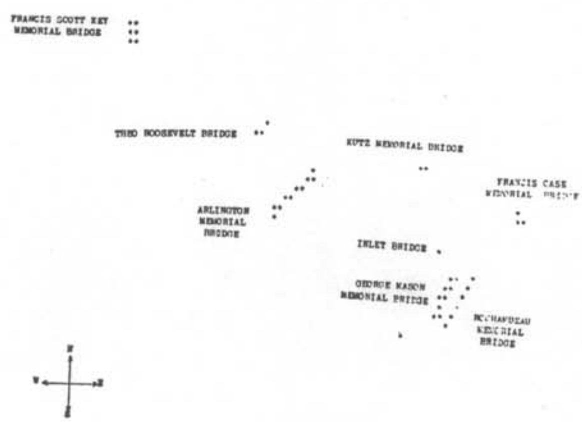

Fig. 6 Recognized Bridges

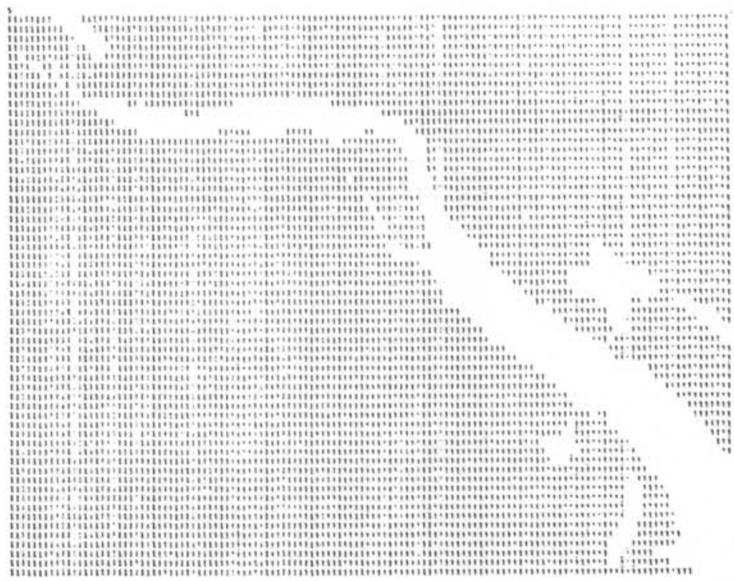

Fig. 7 Potomac River Without Bridges

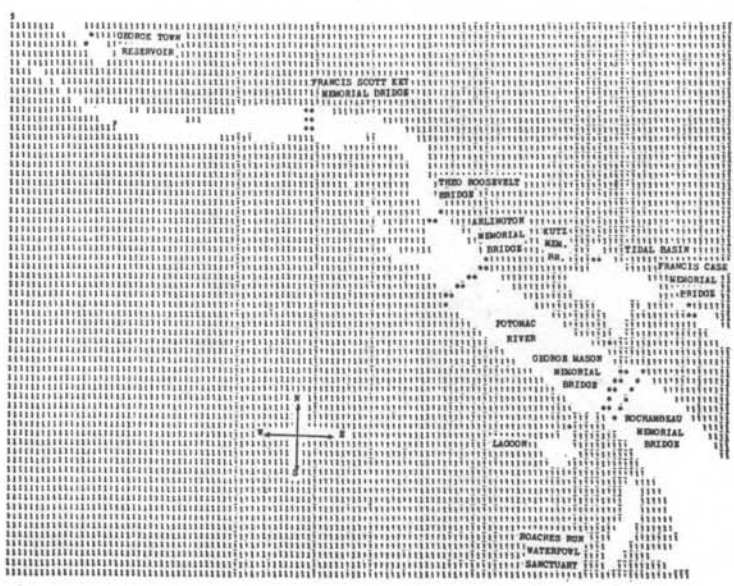

Fig. 8 Potomac River I

Fig. 9

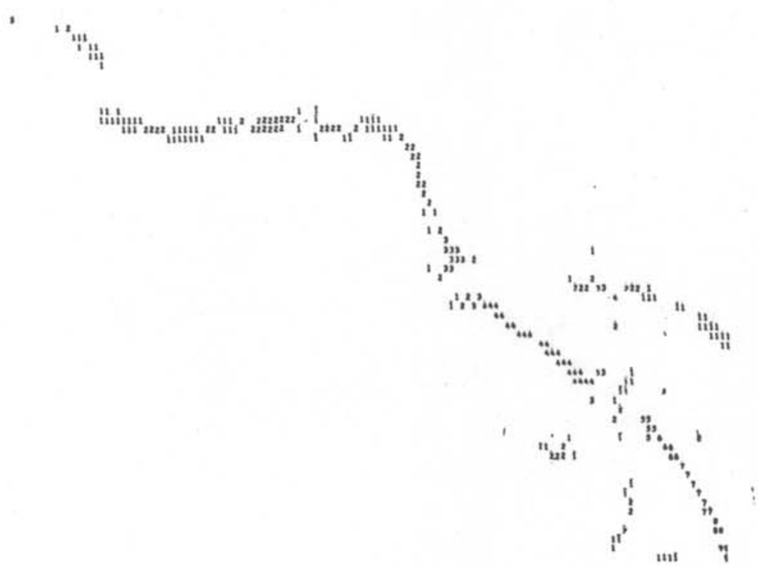

Fig. 10 Skeleton of the River

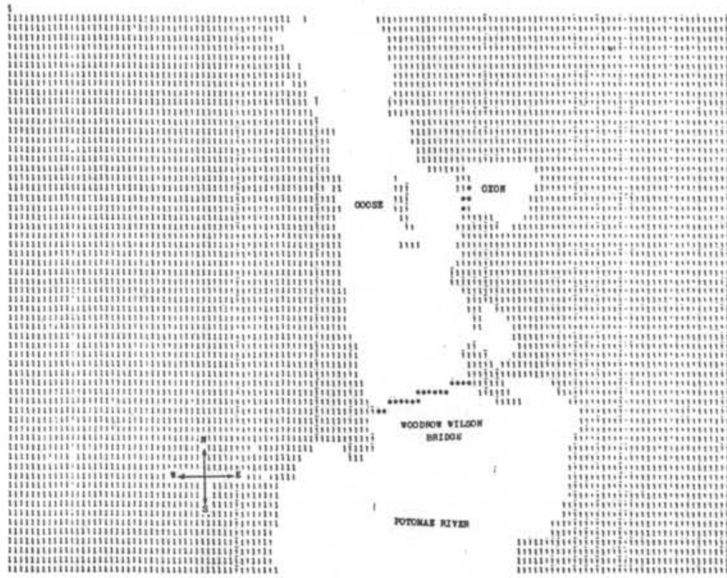

Fig. 11 Potomac River II

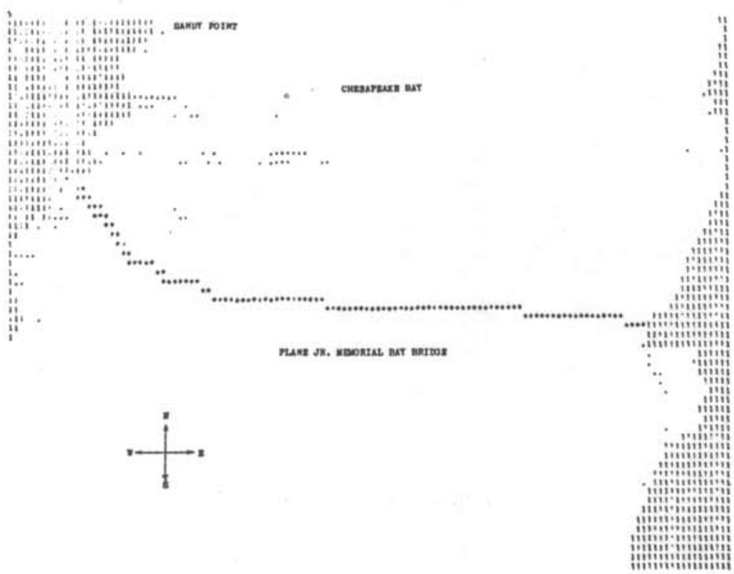

Fig. 12 A Window in Chesapeake Bay