MACHINE VISUAL GUIDANCE FOR AN AUTONOMOUS UNDERSEA SUBMERSIBLE

Hoa G. Nguyen, Peter K. Kaomea* and Paul J. Heckman, Jr.

Undersea Artificial Intelligence and Robotics Branch
Naval Ocean Systems Center
San Diego, California 92152-5000

ABSTRACT

Optical imaging is the preferred sensory modality for underwater robotic activities requiring high resolution at close range, such as station keeping, docking, control of manipulator, and object retrieval. Machine vision will play a vital part in the design of next generation autonomous underwater submersibles.

This paper describes an effort to demonstrate that real-time vision-based guidance and control of autonomous underwater submersibles is possible with compact, low-power, and vehicle-imbeddable hardware. The Naval Ocean Systems Center's EAVE-WEST (Experimental Autonomous Vehicle-West) submersible is being used as the testbed. The vision hardware consists of a PC-bus video frame grabber and an IBM-PC/AT compatible single-board computer, both residing in the artificial intelligence/vision electronics bottle of the submersible.

The specific application chosen involves the tracking of underwater buoy cables. Image recognition is performed in two steps. Feature points are identified in the underwater video images using a technique which detects one-dimensional local brightness minima and maxima. Hough transformation is then used to detect the straight line among these feature points. A hierarchical coarse-to-fine processing method is employed which terminates when enough feature points have been identified to allow a reliable fit. The location of the cable identified is then reported to the vehicle controller computer for automatic steering control. The process currently operates successfully with a throughput of approximately 2 frames per second.

1. INTRODUCTION

Traditional methods for guidance of submersibles employ sonars, magnetic sensors, acoustic transponders and optical sensors. Acoustic transponders and sonars are long-range devices, their useful operations are limited to outside the 3-meter range. Magnetic sensors have poor resolution and are only effective in the vicinity of relatively large ferro-magnetic objects. Optical imaging sensors (e.g. TV camera), on the other hand, are most effective at shorter distances, where the effects of forward and back-scattering are less dominant. Optical imagers are thus the systems of choice for applications that require high image resolution at close range, such as station keeping, control of manipulators, object identification, cable following or salvage and retrieval. Computer vision will play an important role in achieving underwater autonomous systems.

Visual guidance of submersibles is currently accomplished by relaying the video data to topside operators via high bandwidth data links such as optical fibers or hard cables. Merely replacing the operators with powerful topside computers will not be enough. To achieve the highest degree of freedom and truly enjoy the full advantages of an autonomous submersible, the information processing must be accomplished aboard the submersible itself.

This is a report on an effort to demonstrate that automatic vision-based guidance can be accomplished in real-time aboard a free-swimming submersible.** Vision and control software have been developed to solve a simple underwater guidance problem. The hardware has been kept to a minimum, and consists of low-cost, low-power, off-the-shelf single-board products. The success of this demonstration will suggest much more complex applications given the more powerful processors currently in existence.

The Naval Ocean System Center's EAVE-WEST (Experimental Autonomous Vehicle-West) submersible [2] is being used as the testbed (see Figure 1). The vision hardware resides in the artificial intelligence/vision electronics bottle of this submersible (see Figure 2) and includes an IBM-PC/AT compatible 80286 single-board computer, PC-bus frame grabber receiving input from an underwater video camera, and a hardcard for program storage.

---

*Currently with the Cognitive Sciences Branch.
**Preliminary results have previously been published [1].

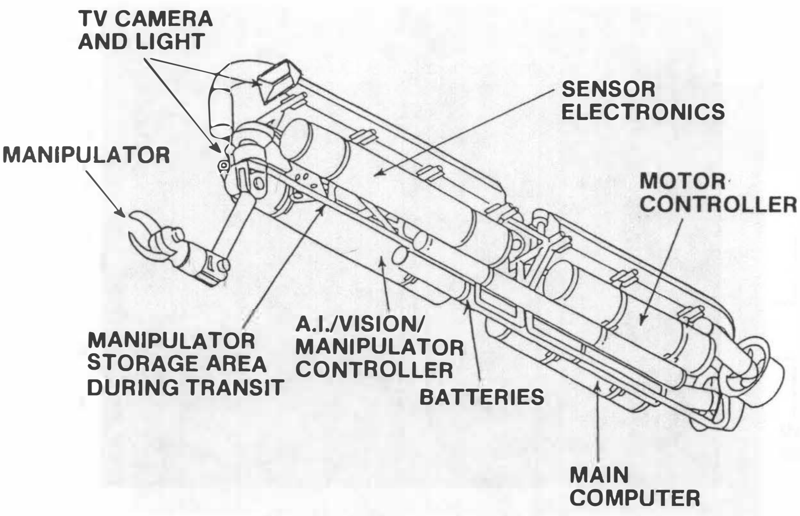

Figure 1. The NOSC Experimental Autonomous Vehicle (EAVE-WEST)

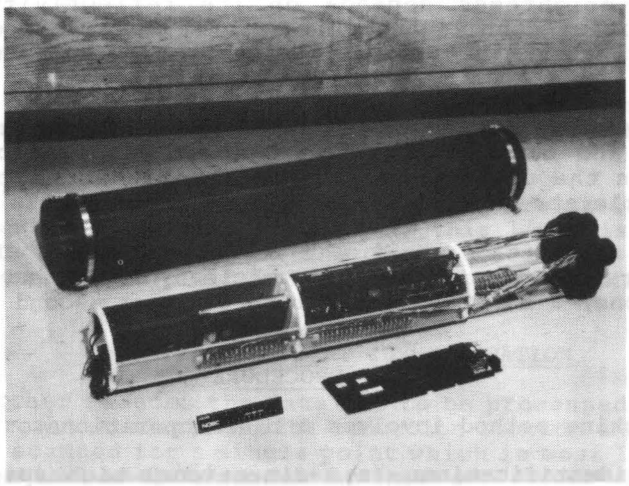

Figure 2. The AI/Vision electronics bottle

2. OPERATIONAL SCENARIO

The application we have chosen to demonstrate machine visual guidance involves the recovery of moored objects. The targets selected are the vertical cables and chains which are often connected to inflatable buoys, intrumentation buoys, or acoustic transponders (see Figure 3). The submersible is guided to the moored object by sonar or directional hydrophones. The image recognition process takes over when the object and its cable are visible, and guides the vehicle along the cable to a point where the recovery process can be initiated. The vision computer keeps the vehicle centered on the cable as the vehicle descends by sending periodic steering information to the vehicle controller.

Useful constraints and guidelines derived from the target description above include:

a. Straight and elongated shape. The width of the target in the image is dictated by the type of cable or chain used, the field-of-view of the lens, and the distance from the target to the camera. However, the minimum width should always be greater than 1 to eliminate single-pixel noise.

b. Approximately vertical major axis. Arbitrary limits of +/- 30 degrees from the vertical were used for this initial effort. These can be refined by calculations using specific buoy buoyancy, cable weight and water velocity.

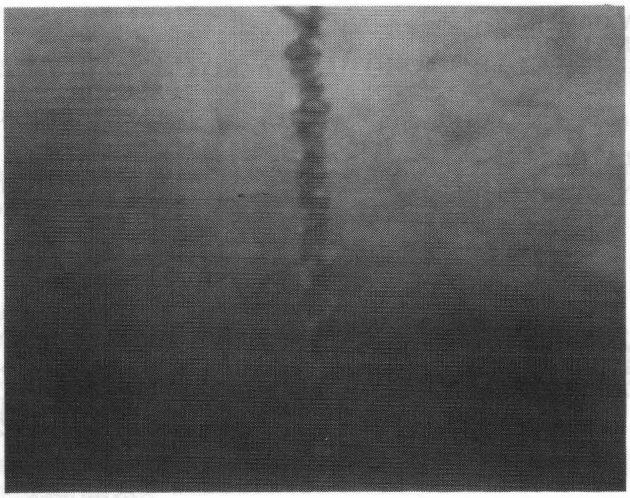

Figure 3. Buoy chain

c. Gray-level segmentable target. Figure 3 shows that in daylight the background is lighter than the target due to the scattering in the background water. On the other hand, the target's relative brightness depends on its reflectivity when it is directly illuminated by an artificial spot light. The vehicle controller computer must inform the vision computer whether natural or artificial lighting is being used and the approximate target reflectivity.

d. Blurred boundaries. The images will tend to be blurry due to the physical properties of forward and back scattering of light propagation through water. This constraint necessitates the use of recognition algorithms which do not require nicely defined edges and can tolerate gaps.

e. Minimum target recognition speed of 1 image per second. The necessary update rate for controlling a submersible depends on the vehicle dynamics and operational environment. For low speed applications, a minimum update rate of 1 per second is adequate.

3. PROCEDURE

Our buoy chain tracking method involves 3 basic operations:

a. Feature point identification: a 1-dimensional high speed operation which scans selected rows in the image for possible centers of the cable or chain. One-dimensional operations are possible because only approximately vertical lines are being sought. Hence, most of the useful information is contained in the horizontal direction.

b. Line identification: an operation which searches for colinear feature points in the image. Given the feature points identified in (a), this routine identifies the slope and intercept of a line which passes through the largest number of points.

c. Adaptive data reduction: a coarse-to-fine search routine which iteratively selects the rows on which steps (a) and (b) are performed.

4. ADAPTIVE DATA REDUCTION

To achieve a high throughput, the image data should be significantly reduced before the more computation intensive stages. An adaptive data reduction algorithm has been developed as a means for reducing the 512-line image through hierarchical coarse-to-fine search. First, a number of horizontal rows evenly spread over the height of the image are selected. Feature points are identified in these rows (step 3a). These feature points are then passed on to the line identifying procedure (step 3b). If a target has not been identified with enough confidence, the adaptive data reduction procedure further bisects the intervals between previously selected rows and continues sending more image rows to the feature point identification and line identification routines. The line identification procedure keeps a cumulative record of the feature points it receives, thus a coarse-to-fine search routine is implemented. Figure 4 gives an example of this iterative process.

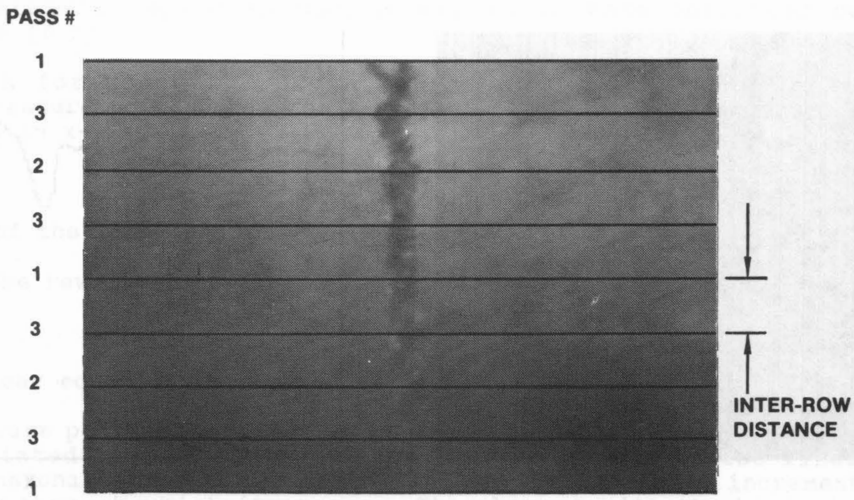

Figure 4. Adaptive data reduction method

This iterative search routine stops under 2 conditions:

a. A target has been identified to a reasonable level of confidence. The confidence factor is based on the number of feature points which contribute to the determination of the line.

b. The proximity of the horizontal rows reaches a preset limit. This limit prevents costly searches of poor (or no-target) images. If this limit is reached before the target is identified, the image is deemed poor and discarded. A new image is taken and a new search started. At present this limit is set to an inter-row distance of 5 pixels.

5. FEATURE POINT IDENTIFICATION

This function further reduces the data set to be processed. Each time it is evoked, it receives as input from the adaptive data reduction routine a single horizontal row from the image. The row is scanned for a single point which is most likely to correspond to the center pixel of the vertical chain or cable of interest.

The feature point identification consists of 2 sub-processes:

5.1. Contrast region identification

Against the more homogeneous background, the target cable or chain exhibits marked contrast. In natural light, a horizontal scan across the image would show a distinct dip in brightness representing the cross section of the target (which may be a peak in images taken with artificial light). However, the high scattering property of water prevents this dip from being square. The edges of the pulse therefore tend to be more slanted, as can be seen in Figure 5.

To find this region of interest, the horizontal row is scanned from left to right for:

a. the region of largest continuous increase in grey level and

b. the region of largest continuous decrease in grey level.

These two regions mark the boundaries of our target pulse.

Occasionally, camera noise or marine particles can cause large variations in grey level. Such variations normally have very narrow width, and can be filtered by imposing a minimum width threshold.

On the other hand, on scans with no target present, the most promising region of contrast can be very shallow. A minimum height threshold is similarly used to eliminate false feature points in these cases.

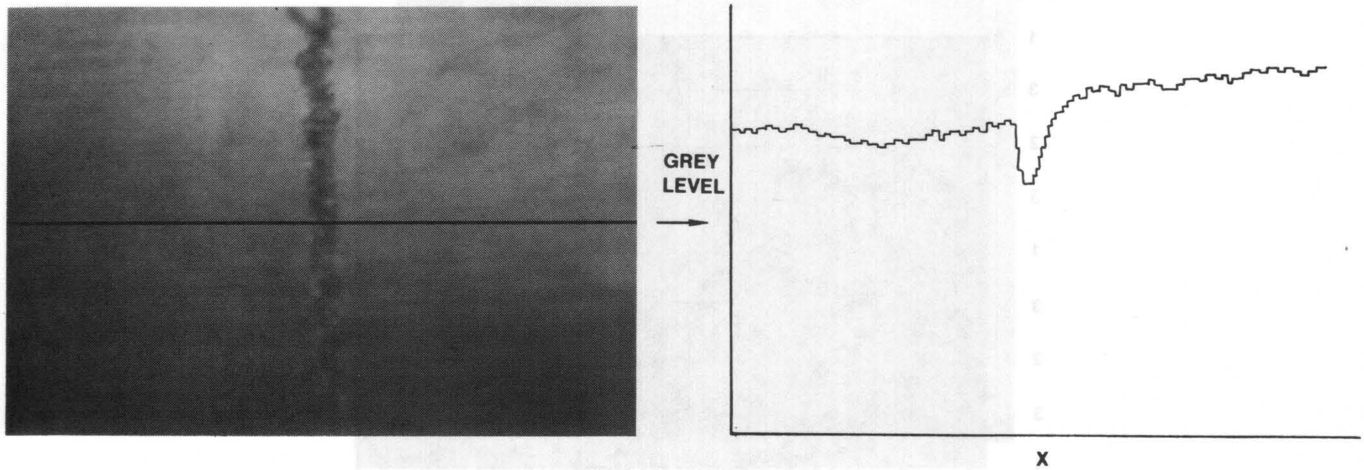

Figure 5. Feature point identification

5.2. Contrast region to feature point

The midpoint of the region bounded by the inner ends of the above edges is designated a feature point, a pixel with high probability of lying at the center of the desired cable or chain. Figure 6 shows a collection of feature points obtained after several passes through the feature point identification routine.

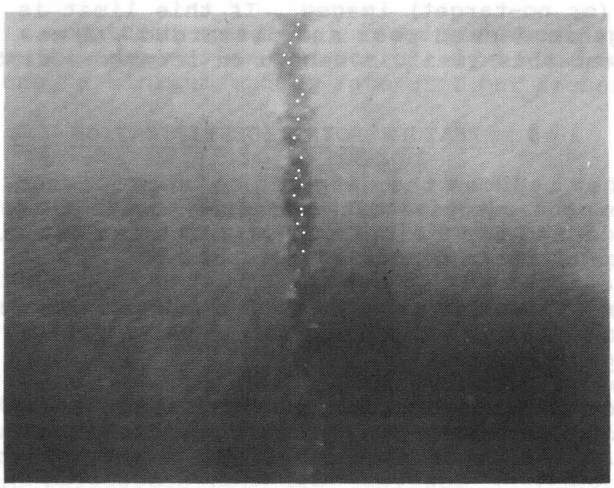

Figure 6. Feature points identified

6. STRAIGHT LINE IDENTIFICATION

Several methods for linking points into straight lines were investigated. Chain coding [3,4] was found to be not effective since it is too susceptible to noise and does not handle gaps efficiently. Least squares fitting [5] is a fast and effective way of linking points into a line if the points to be linked represent a spreading of the line by normal-probability-distribution error. However, in our present application, variations in the brightness of the background may contribute feature points that are not merely a noisy spread of the target. It is desirable to have only those points which form the longest linear cluster contributing to the determination of the line.

The Hough transform was found to be a better method for linking feature points in this application. The Hough transform maps each feature point in the image space into a line in a new parameter space in such a way as to make collinear points map into intersecting lines [6,7].

One approach for using the Hough transform to find straight lines involves transforming the feature points from the x-y space into the slope/intercept space [8]. The equation of a line in x-y space is

$$y = mx + c \tag{1}$$

where m = slope of the line, and c = y-intercept.

This equation can be rewritten as

$$c = -xm + y \tag{2}$$

This is also a linear equation in the m-c space, with x = slope and y = c-intercept.

For each feature point identified in the x-y space, the coordinates (x_k, y_k) are used to find the associated line in the m-c space (see Figure 7). These lines are kept in a cumulative 2-dimensional array (m,c). Each line in the m-c space increments the elements in the (m,c) array through which it passes. The element with the highest value--at (M_0, C_0)--is a result of the intersections of the largest number of lines in the m-c space. It also represents the longest linear cluster of feature points in the image, which has slope M_0 and y-intercept C_0. The accuracy and noise tolerance depend on the resolution chosen for m and c. Presently m is the slope of angles at 1-degree intervals, and the resolution for c is 8 pixels.

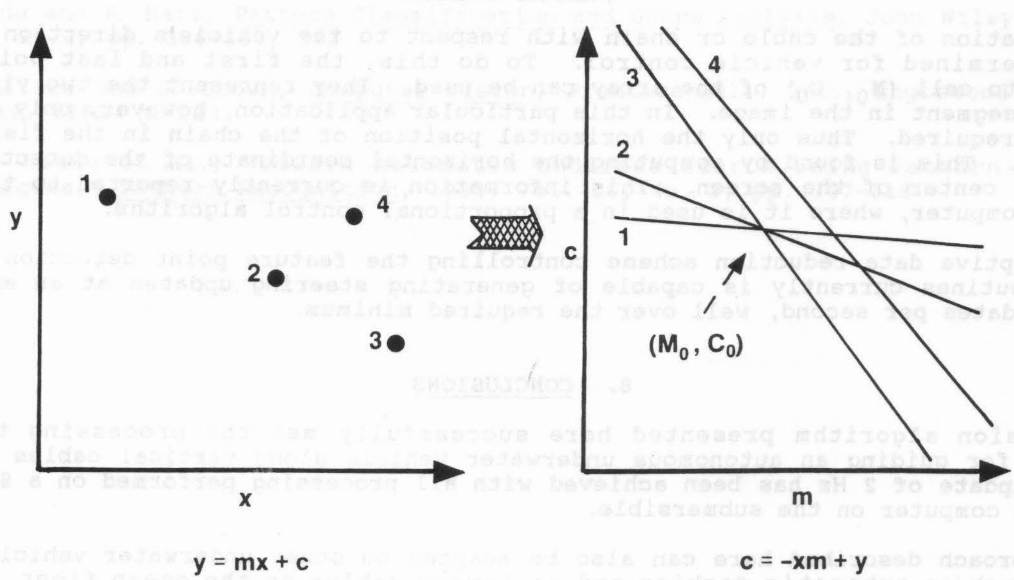

Figure 7. The Hough transform

This transform is simple (linear); can tolerate gaps; and can accommodate noisy, jagged boundaries (by adjusting the resolution of c). Furthermore, the points which are not in the vicinity of the cable or chain do not influence the formation of the line. It is thus appropriate for this application. However, the x- and y-axes have been switched from conventional notation (x is now down, and y is across) to prevent infinite slopes since approximately vertical lines are being sought. The slope m is computed for angles at 1-degree intervals between +/- 25 degrees from the vertical in pixel space (approximately +/- 30 degrees on the screen, due to the aspect ratio of the pixels). Given the feature points in Figure 6, the resulting line with slope = M_0 and y-intercept = C_0 is depicted in Figure 8.

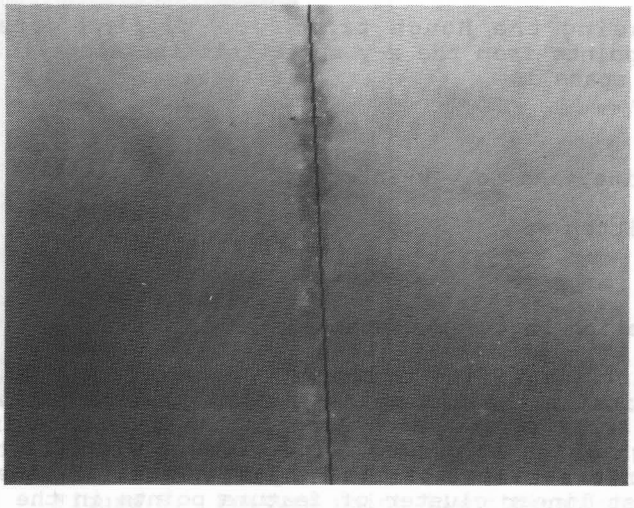

Figure 8. Chain detected

7. VEHICLE CONTROL

The location of the cable or chain with respect to the vehicle's direction of travel must be determined for vehicle control. To do this, the first and last points which contributed to cell (M_0, C_0) of the array can be used. They represent the two visible ends of the line segment in the image. In this particular application, however, only left/right steering is required. Thus only the horizontal position of the chain in the field-of-view is necessary. This is found by computing the horizontal coordinate of the detected line at the vertical center of the screen. This information is currently reported to the vehicle controller computer, where it is used in a proportional control algorithm.

The adaptive data reduction scheme controlling the feature point detection and Hough transform routines currently is capable of generating steering updates at an approximate rate of 2 updates per second, well over the required minimum.

8. CONCLUSIONS

The vision algorithm presented here successfully met the processing throughput requirement for guiding an autonomous underwater vehicle along vertical cables or chains. A steering update of 2 Hz has been achieved with all processing performed on a 80286-based single-board computer on the submersible.

The approach described here can also be adapted to other underwater vehicle guidance problems, such as automatic docking and following cables on the ocean floor. Extended visibility is possible with LIBEC (Light Behind Camera) and range gating techniques [9]. Preliminary studies indicate that our vision technique is also applicable to acoustic imaging, which is necessary for extended range in turbid water.

This research effort has demonstrated that for specific, well-defined vision problems, autonomous real-time performance can now be achieved with on-board processing using off-the-shelf, low cost, and imbeddable hardware. This makes possible long-range, application-specific autonomous undersea robots, or supervisory-controlled robots with autonomous low-level visual tasks.

We have selected a simple application for demonstration. However, the hardware chosen to solve the problem was correspondingly simplistic. Our success suggests that more complex vision problems can also be solved in real time with more advanced imbeddable hardware currently available (including transputers, single-board parallel or array processors). Combined with new developments in underwater imaging, this opens up a dynamic area of applications for further exploration.

ACKNOWLEDGEMENTS

The authors would like to thank Ron Reich for his help in obtaining underwater video images, and Dr. A. L. Pai for his valuable contribution to the earlier stage of the project [1]. This project was supported by the Independent Exploratory Development (IED) program at the Naval Ocean Systems Center.

REFERENCES

[1] H. G. Nguyen, P. J. Heckman, Jr., and A. L. Pai, "Real-Time Pattern Recognition for Guidance of an Autonomous Undersea Submersible," Proc. 1988 IEEE International Conference on Robotics and Automation, Philadelphia, PA; pp. 1767-1770

[2] P. J. Heckman, Jr., "Free-Swimming Submersible Testbed (EAVE WEST)," NOSC Technical Report 622, 1980

[3] H. Freeman, "On the Encoding of Arbitrary Geometric Configurations," IRE Transactions on Electronic Computers, June 1961; Vol. 10, pp. 260-268

[4] A. Pugh, Robotic Technology, Peter Peregrinus, Ltd., London, 1983; pp. 67-81

[5] W. Cheney and D. Kincaid, Numerical Mathematics and Computing, Brooks/Cole Publishing Co., Monterey, CA, 1980; pp. 267-271

[6] P. V. C. Hough , "Method and Means for Recognizing Complex Patterns," U.S. Patent 3069654, 1962

[7] R. Duda and P. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons, New York, 1973; pp. 335-337, 373

[8] D. Ballard and C. Brown, Computer Vision, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982; p. 123

[9] T. S. Huang et al., "Toward Automatic Undersea Search Using Pattern Recognition Techniques," Pattern Recognition, 1983; Vol. 16, No. 3, pp. 307-318