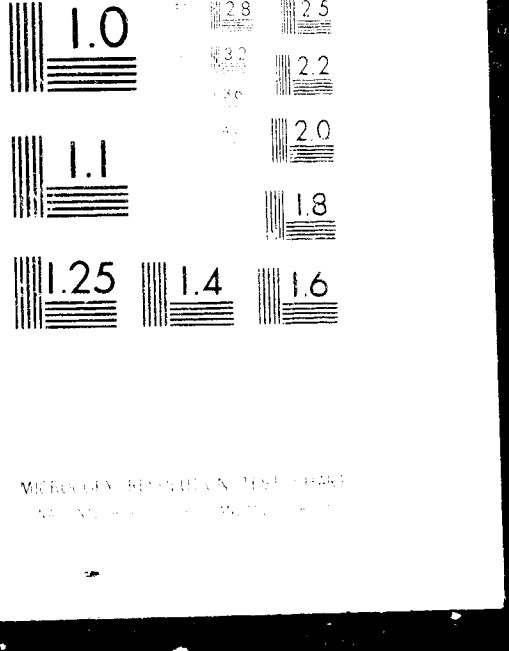


# A SCENE-ANALYSIS APPROACH TO REMOTE SENSING

Final Report

June 1978



By: Jay M. Tenenbaum, Program Manager
Martin A. Fischler, Senior Computer Scientist
Helen C. Wolf, Computer Scientist

Artificial Intelligence Center

Prepared for:

National Aeronautics and Space Administration
Washington, D.C. 02546

Attn: Dr. W.B. Gevarter, Code RES

Contract NASW-2865

SRI Project 4683



Approved:

Peter E. Hart, Director
Artificial Intelligence Center

David H. Brandin, Executive Director
Computer Science and Technology Division

DRA

# A SCENE-ANALYSIS APPROACH TO REMOTE SENSING



Final Report

June 1978

By: Jay M. Tenenbaum, Program Manager
Martin A. Fischler, Senior Computer Scientist
Helen C. Wolf, Computer Scientist

Artificial Intelligence Center

Prepared for:

National Aeronautics and Space Administration
Washington, D.C. 02546

Attn: Dr. W.B. Gevarter, Code RES

Contract NASW-2865

SRI Project 4683

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025
(415) 326-6200
Cable: SRI INTL MNP
TWX: 910-373-1246

SRI International

CONTENTS

ILLUSTRATIONS

# I INTRODUCTION

## A. Background

Satellite-borne imaging sensors have provided the first economical means of gathering large amounts of data on the earth's resources and environment. However, there is as yet no economically feasible way of analyzing the imagery to extract the useful information. The volume of data collected is indeed so high that routine data handling requires massive supercomputers. In the face of such volume, the manual or interactive methods of image analysis currently in use are woefully inadequate.

Most research to date on automatic processing of satellite imagery has concentrated on multispectral classification of individual picture elements (pixels) using conventional pattern-recognition techniques [1]. While promising results have been obtained in selected applications, most notably crop classification [2], exhaustive pixel classification has proved either ineffective or simply too expensive for many remote sensing requirements.

In the past year, we have been exploring the feasibility of automating a variety of previously intractable remote sensing tasks using the methodology and techniques of scene analysis. A key concept in the scene-analysis approach is the use of many diverse types of knowledge to guide image interpretation [3]. In the interpretation of aerial and satellite imagery, maps provide a particularly rich source of knowledge. Map knowledge can provide important constraints on where to look in an image, what to look for, and how to interpret what is seen. Such constraints, properly exploited, permit the extraction of complex information without intensive computation.

The research, so far, has concentrated on a specific class of remote sensing tasks that entail the continuous monitoring or tracking of predefined targets. Monitoring tasks are concerned with detecting an anomalous condition at a specified geographic location or within a specified area. Examples include monitoring the effluents of a particular industrial plant for thermal or chemical pollution, oil storage facilities for spillage, forests for fires, and reservoirs for water quality. Tracking is a variant of monitoring, concerned with determining the current geographic location of a slowly moving object or boundary whose position is known approximately from a previous determination. Examples include tracking icebergs, the spreading boundaries of a known oil spill, the perimeter of reservoirs (to assess changes in water volume), coastal shorelines (to assess erosion), and the width of rivers (to assess flood threat). Ideally, a monitoring system should be able to extract updated information automatically whenever new imagery arrives and distribute it directly to interested users.

## B. Overview of Scene-Analysis Approach

The key to automating monitoring-type tasks with scene analysis lies in knowing where in an image to look and what to look for. With this information, many monitoring and tracking tasks are reduced to simple detection problems with straightforward solutions. For example, once the precise pixel location of a river passing beside a manufacturing plant is known, pollution levels in the plant's effluents can, in principle, be determined by using conventional multispectral analysis. Similarly, forest fires can be detected by looking for infrared hot spots in known forested areas. Tracking slowly changing boundaries, such as the perimeters of water bodies, is also tremendously simplified by knowledge of the boundaries' approximate prior location. Fragments of the boundary can then be detected by performing local searches with simple verification operators.

The above examples all rely heavily on knowing where to look. They thus have, as a common requirement, the location of specific ground reference points in the image. Ground locations have conventionally been determined by warping the current sensed image into correspondence with a reference image, based on a large number of local correlations [4]. The reference image serves as a map indicating locations in the sensed image that correspond to previously determined points of interest in the reference image. The process is computationally expensive and limited to cases where the reference and sensed images were obtained under similar viewing conditions.

To overcome these limitations, we abandon the use of a reference image and rely instead on a symbolic reference map containing explicit ground coordinates and elevations for all monitoring sites as well as landmarks (roads, coastlines, and so forth). The geometric correspondence between this map and the sensed image is established by calibrating an analytic camera model. The camera model makes it possible to predict precisely the image coordinates (in the original unrectified image) corresponding to any world location in the map. The need for expensive image warping is thus eliminated. (The calibration process is described in Section III and, in more detail, in Appendix A.)

Figure 1 summarizes the major steps in a typical monitoring task, such as monitoring the pollutants emitted by industrial sites. First, the parameters of the camera model are calibrated to establish map-image correspondence. Second, the camera model is used to compute the exact image coordinates of each monitoring site visible in the current image. Finally, spectral signatures are analyzed at each location to determine the current degree of pollution. The analysis provides a detailed breakdown of the pollutants detected at each plant site.

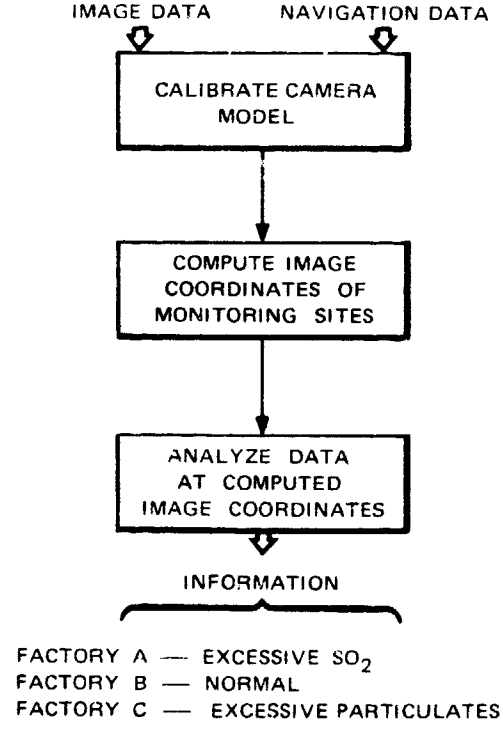



FIGURE 1 BASIC STEPS IN A MONITORING TASK

## C. Advantages of Scene-Analysis Approach

The scene analysis approach outlined above has some potentially significant advantages over conventional bulk image processing approaches of the type commonly used in such applications as crop classification. First, computational requirements are sharply reduced by avoiding intensive processing of all pixels (16 million in a typical 4000 x 4000 LANDSAT image). In particular, bulk image rectification (i.e., warping) is not needed and analysis can be restricted to specific parts of the image containing relevant information. Since processing is concentrated on a few selected image locations, sophisticated forms of analysis involving texture, spatial patterns, and the like become computationally feasible. Second, analysis routines can be simplified and made more reliable by exploiting knowledge of what to look for at each site. For example, classification criteria can be locally tuned to discriminate a few anticipated alternatives, taking into account additional local factors such as weather, season, and past appearance. Finally, the geographic specificity of the analysis yields results that are much more useful than conventional statistical summaries: Knowing that a particular factory is emitting excessive $SO_2$ is much more useful, for example, than knowing that 24 out of 16 million pixels are polluted.

## D. Overview of Report

The following sections outline an experimental scene-analysis system for performing automated monitoring tasks in aerial and satellite imagery. The core of the system is the map data base and the capability for establishing map-image correspondence described in Sections II and III, respectively. Around this core we have implemented a number of representative application programs, described in Section IV, that use map-knowledge to facilitate image analysis. Section V concludes with two possible scenarios, illustrating how the scene analysis approach to automated monitoring could help alleviate data-processing bottlenecks in NASA's present and contemplated remote sensing operations.

## II MAP DATA BASE

The map data base used in this research is essentially a compact three-dimensional description of the location and shapes of major landmarks and monitoring sites. Point features, such as road intersections, small buildings, and many monitoring sites, are represented by their three-dimensional world coordinates and (where applicable) a list of characteristics to be monitored. Linear landmarks, such as roads and coastlines, are similarly represented as curve fragments with associated ordered lists of world coordinates. Ground coordinates are expressed in a standard reference frame, the UTM grid, with elevations expressed in meters above sea level.* The data base can be accessed by location (e.g., What is at x, y, z?), by entity name (e.g., What is the location of factory x?), and by entity type (e.g., What factories are there?). Provisions also exist for associating background information with various map entities such as visual appearance and function, information normally omitted from conventional printed maps intended for human consumption. (For further details on map representation, the reader is directed to Reference [5].)

Our experimental domain throughout this project was the San Francisco Bay Area, as depicted in Figures 2 and 3. Figure 4 is a computer display of a simple map data base of this same area. The map contains a major landmark (the coastline) and number of representative monitoring sites, each designated by a cross. Longitude and latitude data for the on-line map were obtained interactively from the USGS map, using a digitizing table. Elevations were read off the map and entered manually via keyboard. Although displayed as a continuous trace, the coastline, in fact, is internally represented by just 100 discrete sample coordinates.

---

* The map data base was originally developed for another project [Ref. 5] and contains a number of sophisticated features not utilized in this project.

Several map data bases, each highlighting specific features (e.g., roads, railroad yards, piers) were used in experiments described in this report. These maps have not yet been integrated into a monolithic data base, although all software necessary to do so exists (Ref. [5]).

FIGURE 2 HIGH-ALTITUDE VERTICAL MAPPING PHOTOGRAPH OF SAN FRANCISCO BAY AREA

Taken from a U-2 at 45,000 feet

ORIGINAL PAGE IS
OF POOR QUALITY

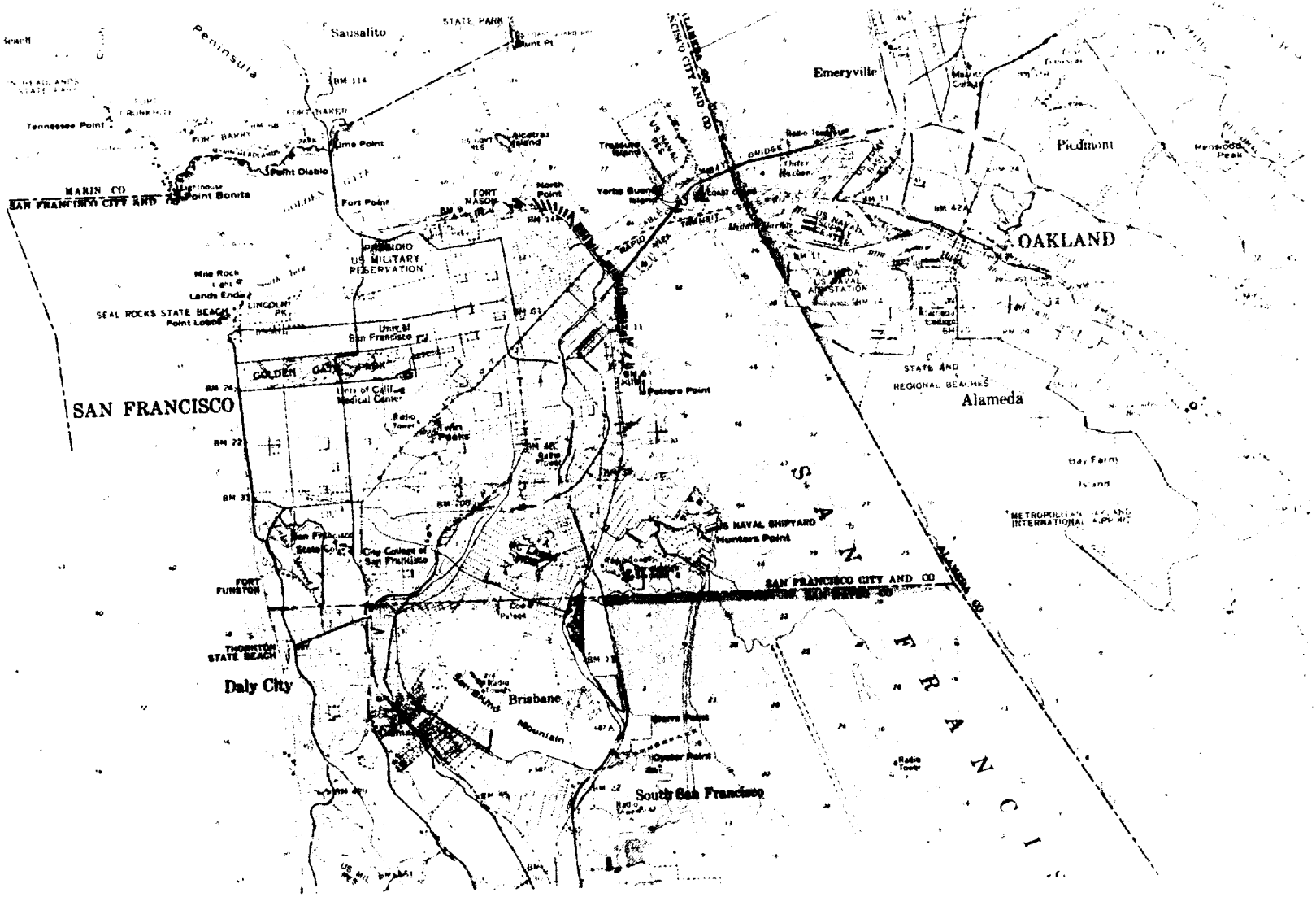



FIGURE 3 TOPOGRAPHIC MAP OF SAN FRANCISCO BAY AREA (1:125,000 SCALE)

ORIGINAL PAGE IS
OF POOR QUALITY

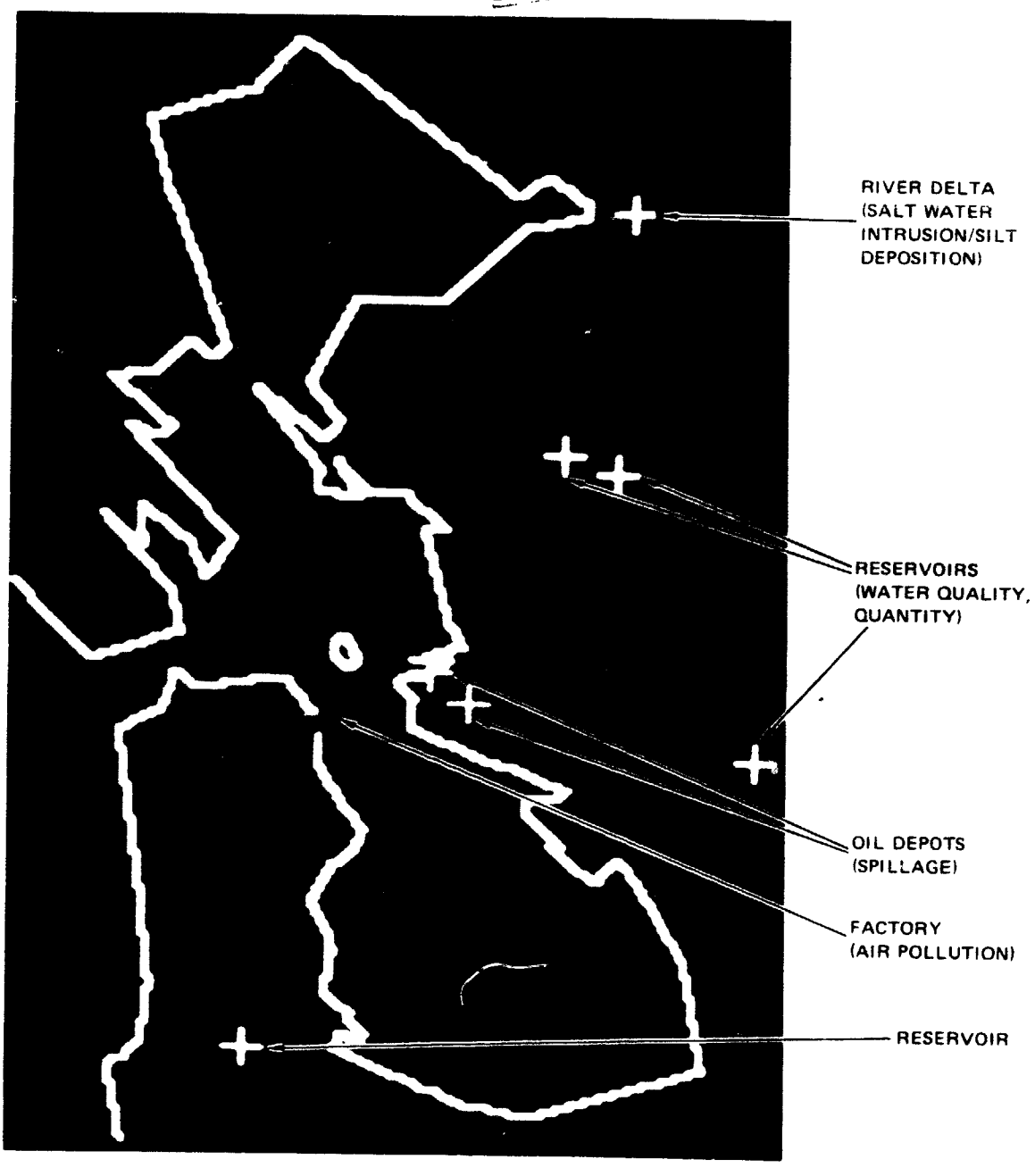



FIGURE 4 COMPUTER DISPLAY OF A SIMPLE MAP DATA BASE FOR THE SAN FRANCISCO BAY AREA, SHOWING MAJOR LANDMARK (COASTLINE) AND REPRESENTATIVE MONITORING SITES (CROSSES)

## III MAP-IMAGE CORRESPONDENCE

The geometric correspondence between map and image, for most sensors,* can be established precisely using an analytic camera model. A typical camera model, as shown in Figure 5, has between five and seven parameters that specify focal length and the location and orientation of the camera (in map coordinates) when the image was taken. Once these parameters are known, the image coordinates corresponding to any map location can be determined with straightforward trigonometry. (The camera location and map location jointly define a ray in space. The intersection of this ray with the image plane yields the desired image coordinates.)

The traditional method of calibrating a camera model requires two stages: First, a number of known landmarks are independently located in the image; and second, the camera parameters are computed from the pairs of corresponding world and image locations, by solving an over-constrained set of equations [6, 7].

The failings of the traditional method stem from the first stage: Landmarks are located in the sensed image by correlating with fragments of reference images. This requires reference images taken under the same viewing conditions as the current sensed image. Moreover, since landmarks are found individually, using only very local context (e.g., a small patch of surrounding image) and with no mutual constraints, false matches commonly occur. (The restriction to small features is mandated by the high cost of area correlation and by the fact that large image features correlate poorly over small changes in viewpoint.)

A new calibration procedure, called "Parametric Correspondence", was developed that overcomes these failings by integrating the landmark-

---

* A notable exception is the scanning multiband sensor used in LANDSAT.

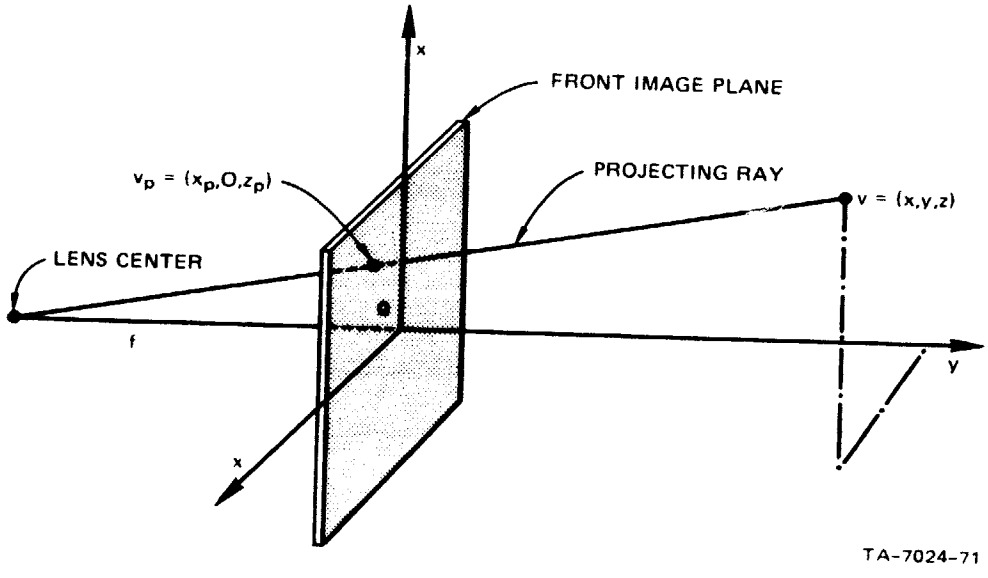



FIGURE 5 MAP/IMAGE CORRESPONDENCE

matching and parameter solving steps and by using global shape rather than tonal appearance as the basis for matching. In this procedure, initial estimates of camera location and orientation are obtained on the basis of available navigational data. The camera model is then used to predict the appearance of landmarks in an image for this assumed viewpoint. Calibration is achieved by adjusting the camera parameters (i.e., the assumed viewpoint) until the predicted appearances of the landmarks optimally match a symbolic description extracted from the image.

A detailed description of parametric correspondence is given in Appendix 1. However, the essential ideas can be quickly grasped through an example. Figure 6 illustrates the process of establishing correspondence between the symbolic map of Figure 4 and the sensed image of Figure 2, using the coastline as a landmark.

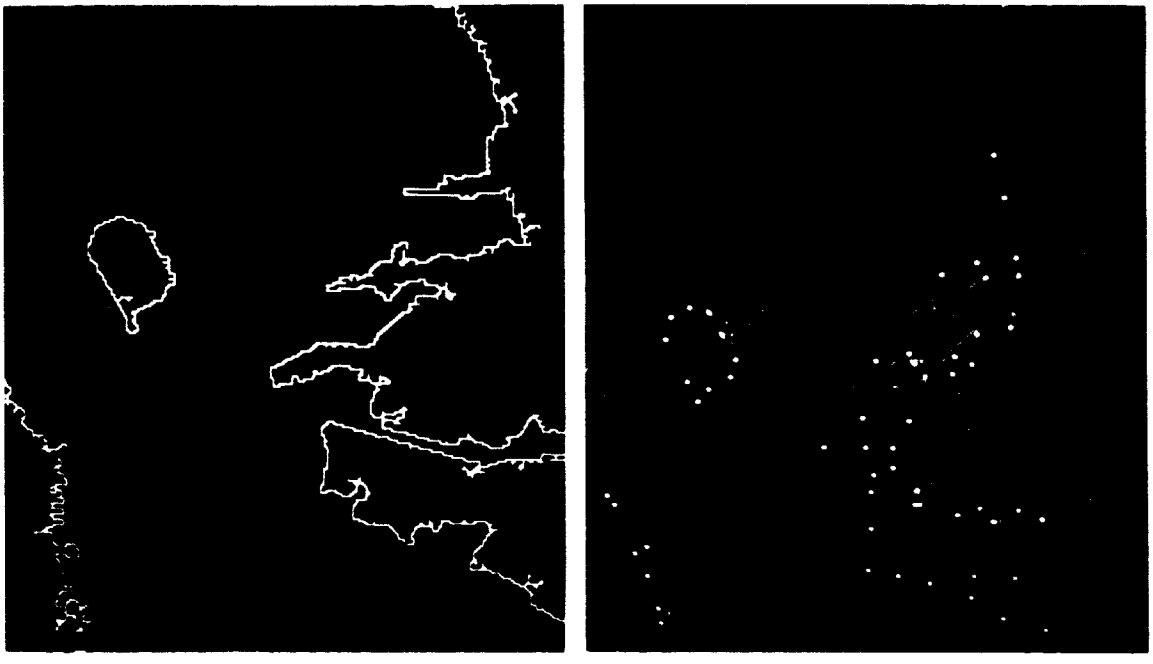
First, a simple edge follower was used to trace the high contrast coastline in Figure 2, producing the edge image shown in Figure 6(a). Next, using initial camera parameter values (estimated manually from navigational data provided with the image), the coastline coordinates in the map were transformed into corresponding image coordinates and overlaid on the extracted edge image [Figure 6(b)]. The average mean square distance between the extracted coastline and that predicted on the basis of the assumed viewpoint was seven pixels. A straightforward hill-climbing algorithm then adjusted the camera parameters to minimize this average distance. Figure 6(c) shows the final state, in which the average distance has been reduced to 0.8 pixel.

Using the final parameter values, it is now possible to determine within a pixel the precise image locations corresponding to each monitoring site in the map. Only three sites are actually visible in this image: the two oil depots and the coffee factory. These are shown in Figure 6(d), superimposed on the original image. Figures 7 and 8 provide two additional examples of the calibration process, illustrating its ability to accommodate arbitrary viewpoints. The apparent misregistrations in Figures 6(c), 7(c), and 8(c) are

actually the result of errors in contour extraction [Figures 6(a), 7(b), and 8(b), respectively]; despite such errors, the global matching criteria is still able to achieve subpixel accuracy of the projected map points.

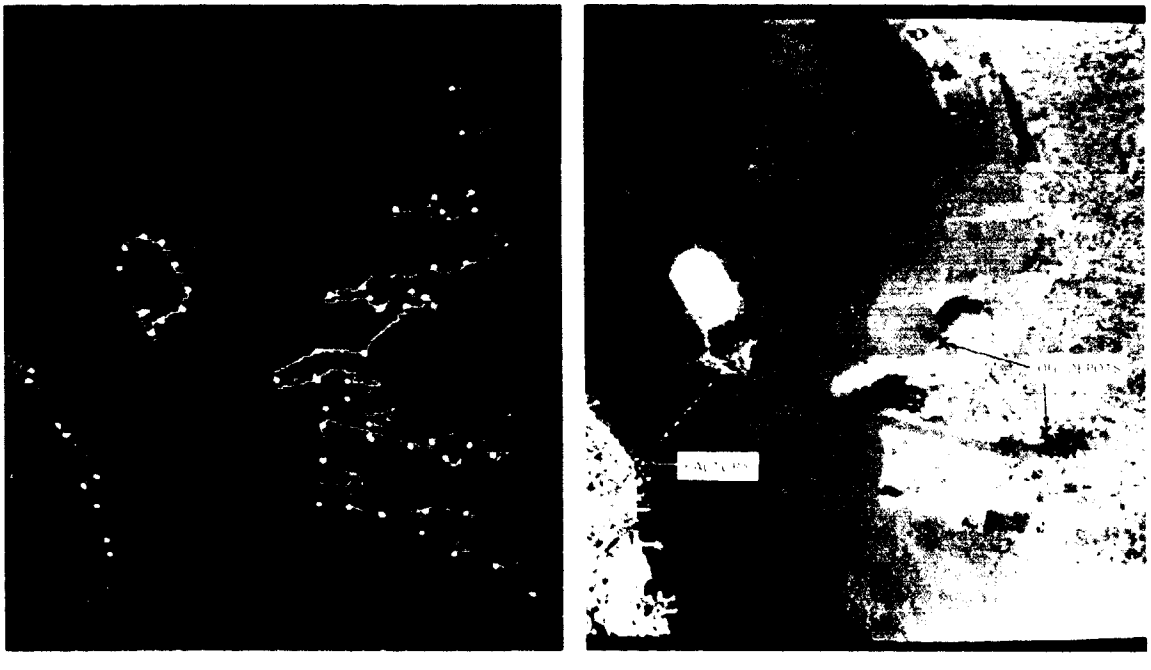
The relative merits of parametric correspondence and other approaches to map-image correspondence are discussed at length in Appendix A. However, three principal strengths (viz., computational cost, robustness, and storage economy) are worth noting briefly here:

- Computational Cost -- Determining image locations by projecting a map through a camera model entails far less computation than by warping an image into correspondence with a reference image because there are typically orders of magnitude fewer map coordinates than image pixels to transform. Warping, moreover, is only an approximation to correspondence that breaks down when reference and sensed images differ significantly in viewpoint.
- Robustness -- Parametric correspondence appears robust compared with correlation-based techniques for landmark matching because it relies on global shape features that are relatively immune to seasonal and diurnal variation and to ambiguous matches. Furthermore, because shapes are projected through the camera model before matching, distortions resulting from viewpoint are not a problem.
- Storage Economy -- Parametric correspondence permits significant storage economies--first, because a three-dimensional map typically contains much less data than a reference image; and second, because multiple reference images may be required to handle a range of viewing situations (viewpoint, sensors, sun angle, and so forth), whereas a single map will suffice. The storage factor may be critical in applications covering extensive geographic areas.



(a) COASTLINE EXTRACTED BY BOUNDARY FOLLOWER

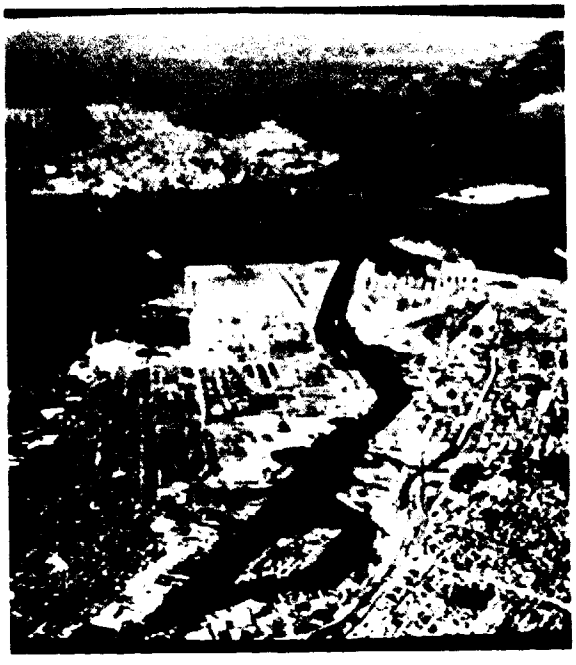
(b) PREDICTED IMAGE COORDINATES OF COASTLINE, (BASED ON NAVIGATIONAL ESTIMATES OF CAMERA LOCATION AND ORIENTATION) SUPERIMPOSED ON EXTRACTED BOUNDARY



(c) PREDICTED COASTAL COORDINATES AFTER OPTIMIZATION OF CAMERA PARAMETERS

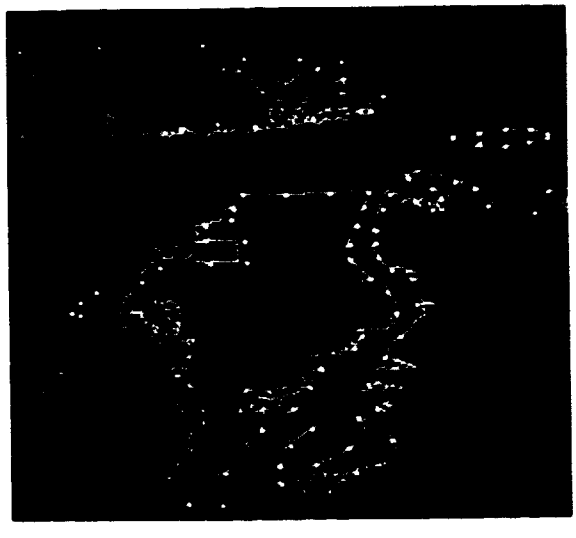
(d) PREDICTED IMAGE LOCATIONS OF VISIBLE MONITORING SITES BASED ON OPTIMIZED PARAMETERS

FIGURE 6 PARAMETRIC CORRESPONDENCE: EXAMPLE 1

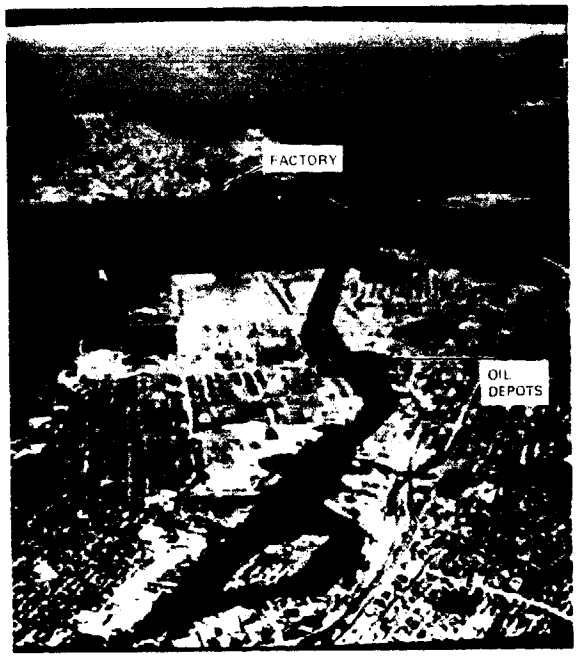


(a) OBLIQUE VIEW OF SAN FRANCISCO BAY LOOKING WEST FROM ALAMEDA (FOREGROUND)

(b) NAVIGATION-BASED PREDICTION OF COASTLINE COORDINATES SUPERIMPOSED ON EXTRACTED COASTLINE



(c) PREDICTED COASTLINE COORDINATES AFTER OPTIMIZATION OF CAMERA PARAMETERS

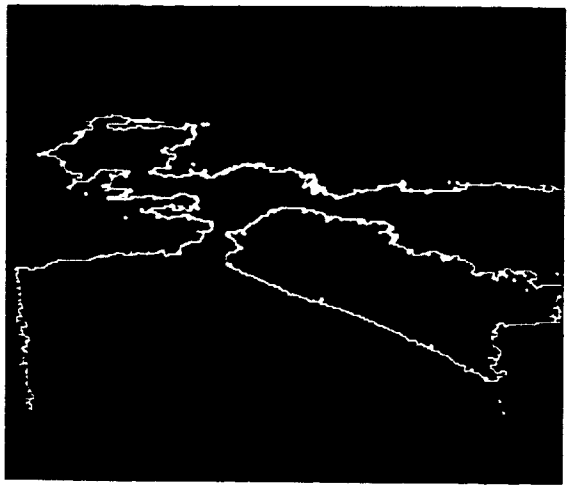



(d) PREDICTED IMAGE LOCATIONS OF VISIBLE MONITORING SITES

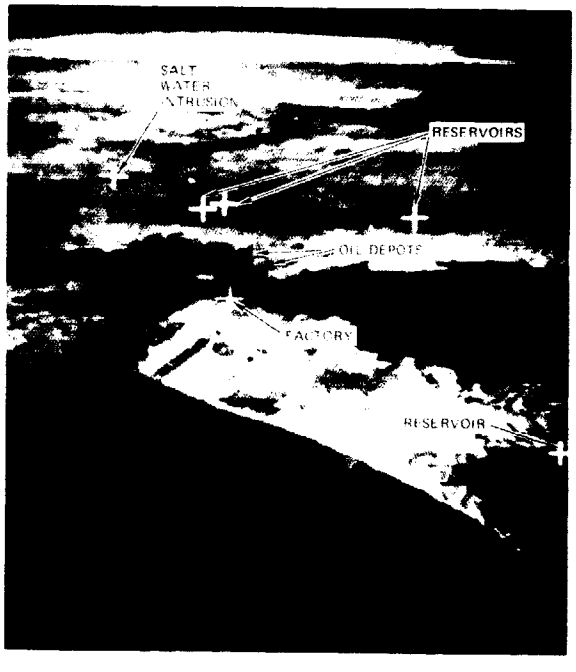
FIGURE 7 PARAMETRIC CORRESPONDENCE: EXAMPLE 2

(a) HIGH ALTITUDE OBLIQUE VIEW OF SAN FRANCISCO BAY LOOKING EAST FROM THE PACIFIC OCEAN

(b) NAVIGATION-BASED PREDICTION OF COASTLINE COORDINATES SUPERIMPOSED ON EXTRACTED COASTLINE



(c) PREDICTED COASTLINE COORDINATES AFTER OPTIMIZATION OF CAMERA PARAMETERS




(d) PREDICTED IMAGE LOCATIONS OF VISIBLE MONITORING SITES

FIGURE 8 PARAMETRIC CORRESPONDENCE: EXAMPLE 3

# IV MAP-GUIDED MONITORING

Having placed the image into parametric correspondence with the three-dimensional map, it is possible to predict the image coordinates of any feature in the map and, conversely, to predict the map features corresponding to any point in the image. Given this capability, many basic monitoring tasks of the type discussed in previous sections can be automated using straightforward image-analysis techniques. Four generic classes of tasks will be considered to illustrate the use of map-guidance in remote sensing.

## A. Pixel Classification Tasks

In pixel classification tasks, conventional multispectral analysis techniques are applied at designated image locations to detect abnormal states. In Figure 8(d), for example, one could, in principle, test the pixels located in reservoirs for water quality, the pixels located in shipping channels beside oil depots for evidence of spillage, the pixel located at the industrial plant for evidence of particulates, and the pixel located at the Sacramento River Delta for evidence of salt water intrusion.

The above examples are merely illustrative; their feasibility depends on the availability of sensory data with suitable spectral and spatial resolution. (Readers interested in such applications are referred to the vast remote sensing literature on multispectral classification, as discussed and referenced in [8]). Of interest here is the principle of using map knowledge to constrain where to look and what to look for in an image. In pixel classification tasks, the primary advantages of map guidance are the computational efficiency gained by restricting analysis to relevant pieces of the image and the utility gained by being able to associate findings with particular

geographical entities (e.g., factory X is emitting $SO_2$). However, for other, more complex tasks, the advantages are more profound.

## B. Boundary and Line Verification Tasks

An important requirement in many monitoring tasks is the need to determine the precise path through an image of a linear feature (e.g., coastline, river, road) whose location and shape are known, perhaps only approximately, from a map. Monitoring the water level of reservoirs and the traffic density on roads are two representative applications of boundary and line verification, respectively. Map knowledge can be used in such tasks to facilitate both the process of locating the boundary in the image and the subsequent interpretation of boundary characteristics in terms significant to a particular application. These uses will now be illustrated in the context of the two aforementioned tasks.

### 1. Reservoir Monitoring

Consider first the problem of determining the water level of a reservoir. Water level, of course, is not directly measurable from an aerial image; some additional information or constraint is needed. The required information can be obtained from a terrain map in registration with the image.

As the water level rises and falls, the outline of the reservoir expands and contracts in a predictable way to follow the elevation contours of the terrain (see Figure 9). Thus water level can be determined by extracting the outline of the reservoir in the image and determining its location with respect to known elevation contours. Knowing the water level, one can then integrate over the corresponding region of flooded terrain to determine the volume of stored water. (The function relating water volume and water level is monotonic and can be tabulated for each reservoir.)

Since the surface of a reservoir is flat, the water level can be determined without a complete outline; the image coordinates of even a single point on the reservoir boundary would, in principle, suffice.

In practice, elevations are determined for a number of boundary points and averaged together to compensate for statistical uncertainties in estimating the precise image coordinates of each boundary point. The distribution of elevations, which should be tightly clustered, provides a check on the quality of the map-image correspondence.

Boundary samples are concentrated where terrain slope is most gradual to maximize the sensitivity of edge location to changes in water level. [See Figure 9(b).] The image coordinates corresponding to each selected boundary site are determined to subpixel precision by analyzing the gradient of image intensity along a line perpendicular to the elevation contours at that site. The analysis can be restricted in practice to a contour interval bracketing the water level observed in a previously analyzed image. This constraint not only reduces computation but also serves as an effective contextual filter for discriminating irrelevant intensity discontinuities arising, for example, from other nearby bodies of water. (Theoretical issues that enter into determining boundary locations to subpixel accuracy using map knowledge are discussed in Appendix B.)

The terrain elevation corresponding to a detected boundary point is obtained by linearly interpolating the elevations of the terrain contours used to delimit boundary detection. If the elevation interval is large (say over ten feet), the interpolation can be iterated using progressively narrower intervals to obtain a more accurate elevation estimate. (An alternative technique for determining water level, in which extracted boundary points are matched globally to the shapes of possible contour lines, is described in Appendix B.)

In summary, the basic reservoir monitoring procedure has four steps:

(1) Establish geometric correspondence between the sensed image and a contour map of the terrain. Correspondence must be based on geographically stable landmarks unrelated to reservoir boundaries.

(2) Determine the precise image coordinates of selected points on the reservoir boundary.

(3) Determine the water level corresponding to each boundary point by interpolating the elevations of bracketing contours.

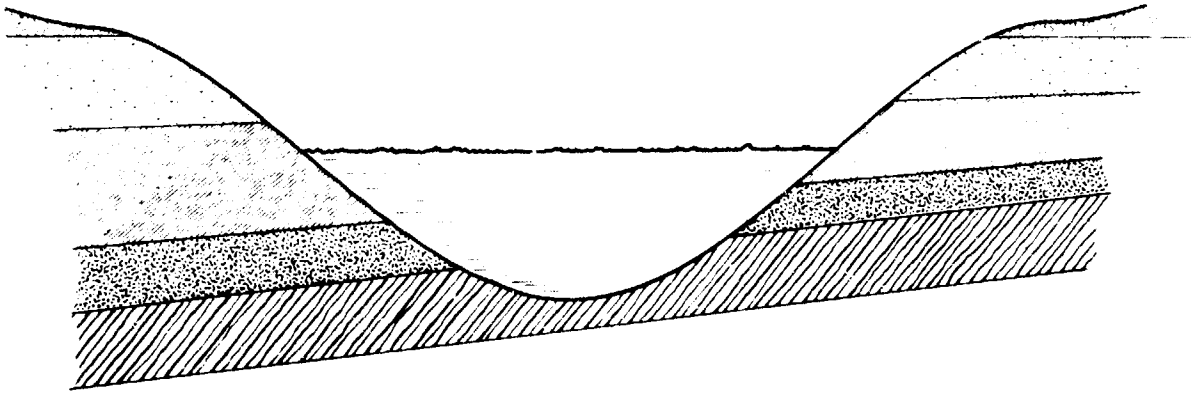
(4) Determine the water volume corresponding to the determined water level by performing a table lookup.

Steps (2)-(4) would be repeated for each reservoir in an image.

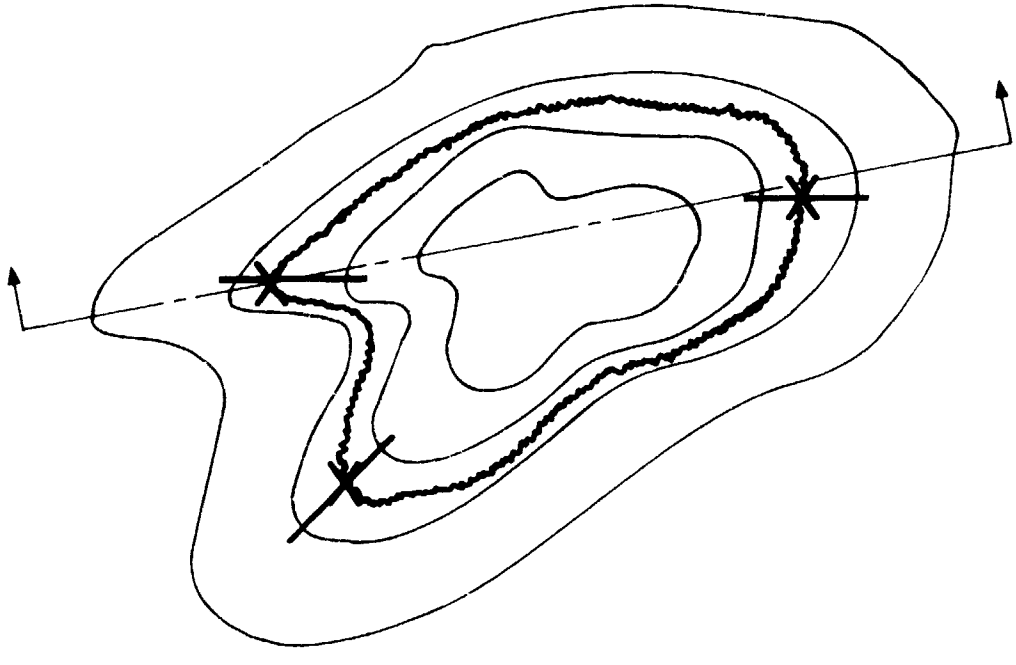
The above procedure was implemented and tested on a set of images of Briones reservoir [the rightmost of the twin reservoirs in the upper center of Figure 8(d)]. Figure 10(a) is a higher resolution image of the Briones shoreline with elevation contours superimposed. The lines in Figure 10(b) indicate selected perpendiculars between the 500 and 550 elevation contours where the terrain slope is most gradual. The location of the land/water boundary along each of these lines was assigned to the point of maximal intensity discontinuity, as shown in Figure 10(c). The water level corresponding to each boundary point was computed by interpolation resulting in the distribution of levels shown in Figure 10(d).

The mean water level in the present image of Briones, based on interpolating 170 boundary points, was determined to be 523.8 feet. This is within a foot of the ground-truth figure provided by the reservoir operator and corresponds to about a one percent error in volume. The accuracy of this approach is limited by the accuracy of the terrain map, the quality of map-image correspondence, and the precision with which the land/water interface can be located in an image. These factors are discussed further in Appendix B.

Although monitoring reservoirs by satellite may not yet be cost-competitive with conventional ground-based approaches, there are other equally important applications of map-guided boundary verification for which no practical alternatives currently exist. Some examples include the monitoring of river widths (and heights) for flood threat, the monitoring of coastlines for erosion, and the monitoring of river deltas for excessive silt deposit. Unlike reservoir monitoring, extensive manual ground-based monitoring is not economically feasible in these applications.



(a) PROFILE VIEW



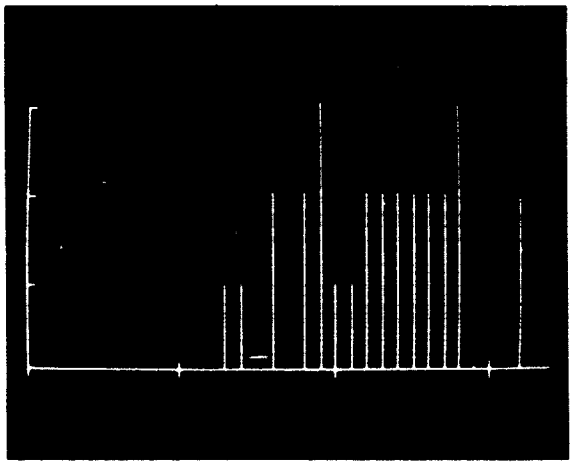
(b) TOP VIEW

FIGURE 9 RELATIONSHIP OF WATER LEVEL TO TOPOGRAPHY OF TERRAIN

(a) TERRAIN CONTOURS SUPERIMPOSED ON IMAGE OF BRIONES RESERVOIR. THE ACTUAL WATER HEIGHT IS 524 FEET ABOVE SEA LEVEL

(b) LINES DESIGNATING LOCATION FOR DETERMINATION OF LAND-WATER BOUNDARY

(c) LOCATIONS OF LAND-WATER BOUNDARY ASSIGNED TO POINTS OF HIGHEST LOCAL GRADIENT ALONG LINES SHOWN IN FIGURE 10(b)



(d) DISTRIBUTION OF WATER LEVELS CORRESPONDING TO BOUNDARY LOCATIONS IN FIGURE 19(c) AS DETERMINED BY INTERPOLATION (x-axis = elevation, 10 feet/division; y-axis = 1 sample/division)

FIGURE 10 WATER LEVEL PROFILES FOR BRIONES RESERVOIR

## 2. Road Monitoring

Periodic surveillance of road traffic is an important task in an automobile-dependent society. An obvious prerequisite for road monitoring is to determine the path of the road in the image.

Conventional sequential line-tracking algorithms are unsuited to this task because they are easily sidetracked whenever either the local evidence for a line is weak or other lines are present in close proximity. These contingencies arise frequently in aerial imagery because roads are usually clustered into networks and pass regularly through heavily textured areas where one or even both edges may be locally obscured. They can be overcome, however, by exploiting map-knowledge.

**Since maps often suppress detail, a technique is needed that can use a rough prediction of the path of the road to guide its determination of the precise path. The problem requirements differ somewhat from those encountered in extracting reservoir boundaries in that a thin linear feature is involved and a continuous path is needed.**

With these requirements in mind, we developed (under ARPA support) a line-tracing algorithm that uses map guidance to constrain the analysis to relevant parts of the image and to bridge gaps where local evidence is weak or ambiguous. The algorithm operates by applying specially developed line and edge detectors in the vicinity of an approximate path predicted by the map and then uses a parallel dynamic programming algorithm to find a globally optimal path through the local feature values. Further technical details can be found in Ref. [9].

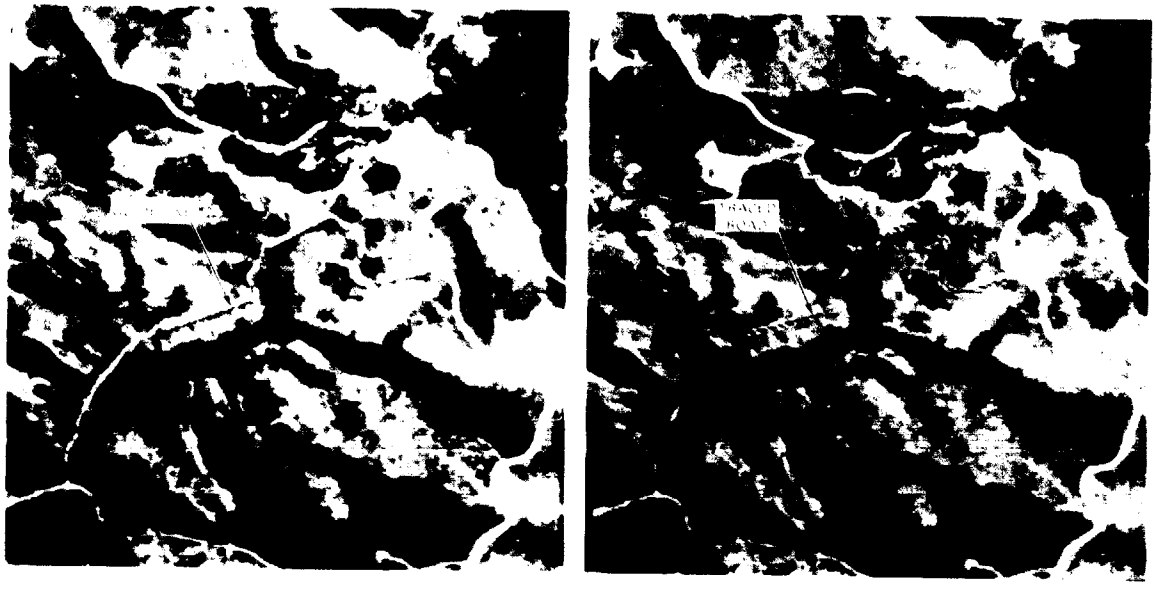
Figure 11 shows the tracing algorithm in action. Figure 11(a) is an aerial image of a rural area taken for a U. S. Geological Survey mapping project. The portion shown has been digitized into 256 x 256 pixels (representing 20-foot squares on the ground), each having one of 256 brightness levels.

Figure 11(a) shows a predicted road path one might obtain from a map with standard (50-foot) cartographic accuracy. A local line

detector was applied at all image points within a band centered on this guideline. The system then found the lowest-cost path from the start of the guideline to the finish, where the incremental path cost between adjacent image points was an inverse function of the local line detector score. The path so traced is displayed in Figure 11(b). Figure 11(c) shows the result of tracing many of the roads visible in the image. Note that the program has traced the center line of the wide road and that it has performed extremely well in areas in which the road is faint or partially obscured, such as at the lower left and the upper right of the image. Figure 11(d) shows the results of guided road tracing in an urban area containing many intersecting streets. The tracings have been fitted with straight line segments to cartographic accuracy. The results here, too, are extremely good.

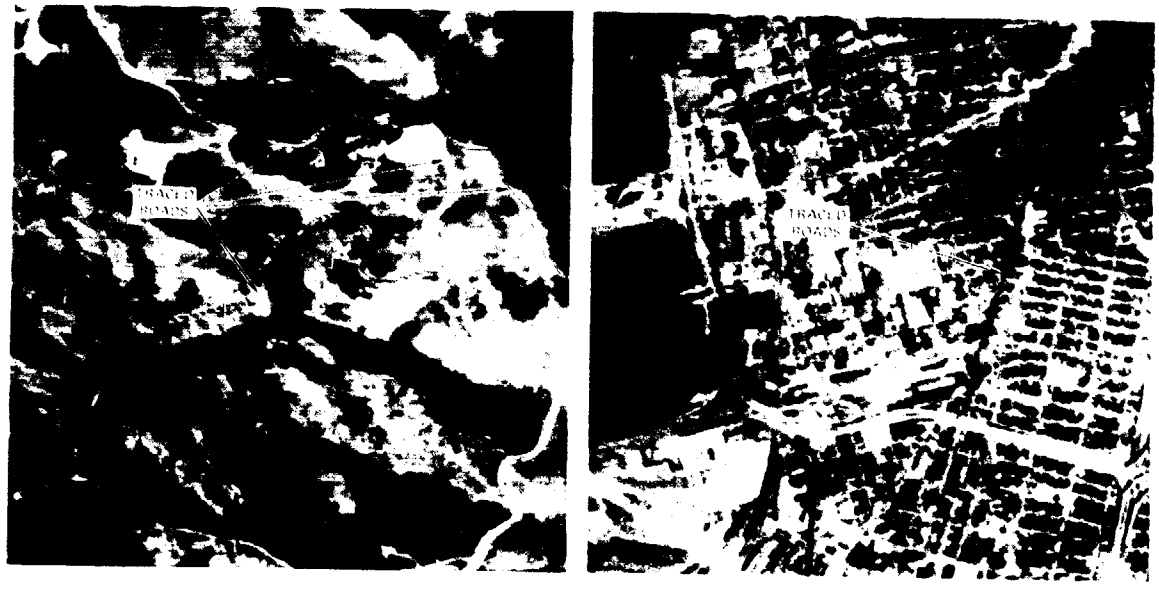
Although we have performed only a limited number of experiments with guided tracing, the results have been most encouraging. The system is capable of tracing linear features that are hard even for a human to discern through a wide range of terrain types and environments. It needs relatively little guidance; but the more guidance it is given, the more reliable and efficient is its performance. It can accept guidance interactively (via light pen), as well as from preexisting maps. Interactive guidance is helpful in map-making applications, allowing new roads to be carefully digitized, based on a crudely sketched guideline.

Guided road tracing is a key step in automating traffic monitoring. Given the road path, vehicle detection can be accomplished by analyzing the intensity variations along the road [10]. More generally, map-guided tracing of linear features is a requirement that arises in a variety of other remote sensing tasks; for example, in the monitoring of rivers and railroad lines. Given suitable operators for detecting local evidence, the global road-tracing algorithm should also work in these other line tracing applications.



(a) A RURAL ROAD WITH GUIDELINE

(b) OUTPUT OF GUIDED TRACING ALGORITHM



(c) GUIDED TRACING OF SEVERAL RURAL ROADS

(d) GUIDED TRACING OF SEVERAL URBAN STREETS

FIGURE 11 GUIDED ROAD TRACING

## C. Object Verification Tasks

Railroad and highway monitoring are two examples of a generic class of remote sensing applications we shall call object verification tasks. Such tasks entail the detection, mensuration, or counting of specified entities whose possible locations and orientations in the image are constrained by map knowledge. The general approach is to determine the image coordinates for a reference structure (such as a railroad track, ship berth, or road) and then apply special-purpose operators to detect objects of interest (such as boxcars, ships, or cars). For example, we have implemented a boxcar-counting routine that analyzes the intensity profiles along predicted paths of railroad track in an image, looking for possible ends of trains and gaps between cars. Such events usually appear as step changes in brightness and dark, transverse lines, respectively. Hypothesized gaps and ends are interpreted in the context of knowledge about trains (e.g., standard car lengths and allowed inter-car gap widths) and about the characteristics of empty track to prune artifacts and improve the overall reliability of interpretation. The program then reports the number of cars classified by length [9]. We have also implemented a ship-monitoring program that analyzes intensity patterns alongside predicted berth locations in a harbor to distinguish ships from water. (Water characteristically has a low density of edges, [11].) Railroad monitoring is illustrated in Figure 12 and ship monitoring in Figure 13.

The demonstrations of boxcar-counting and ship monitoring were implemented as part of ARPA-sponsored research on automated photointerpretation and are summarized here to emphasize the generality and power of the map-guided approach. The key to automating both tasks lies in using map-knowledge to define a highly constrained context (i.e., area of the image) in which relatively simple tests can be used to distinguish objects of interest. Knowing the locations of tracks, for example, reduces the task of boxcar counting to a one-dimensional, template-matching problem. We believe that boxcar counting and ship monitoring are representative of a broad class of object-verification

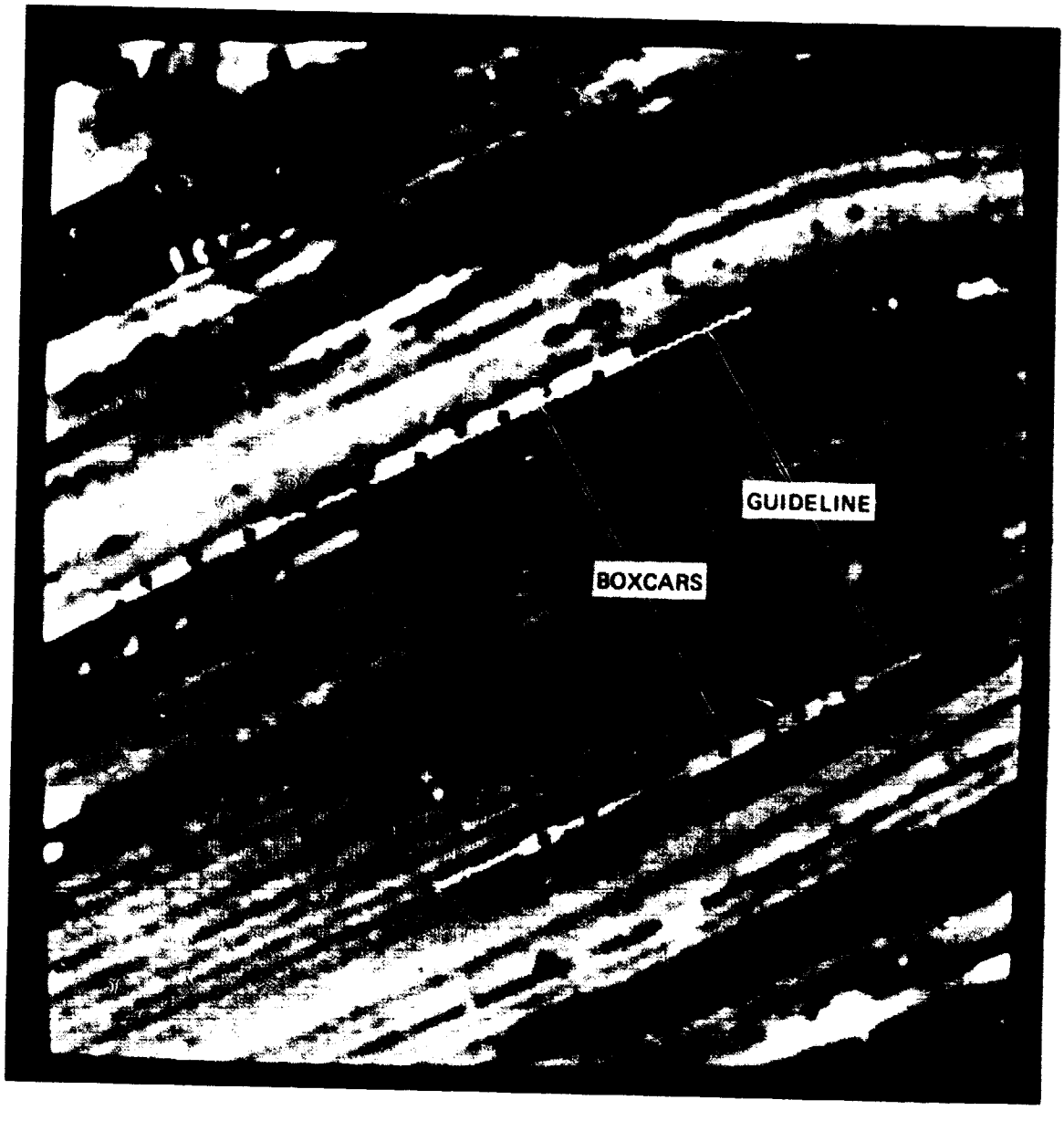



FIGURE 12 AUTOMATED BOXCAR COUNTING

Lines indicating track locations were traced interactively in this example but would, in general, be obtained by putting image in correspondence with a three-dimensional map of the railyard, as in the ship example of Figure 13. Statistical operators are flown along tracks to detect dark transverse lines that are characteristic of gaps between boxcars. Boxcars are indicated by dots whenever the spacing between hypothesized gaps is consistent with knowledge of standard car lengths.

SA-5300-42

SA-5300-43

FIGURE 13 AUTOMATIC SHIP MONITORING

The guidelines indicating known berth locations were obtained for both images from the same three-dimensional map of Oakland Harbor, based on determination of viewpoint for each image. The dark, wiggly lines beside the berths indicate regions of high edge content, characteristic of ships.

tasks that includes counting planes on runways and cars on highways, for which similar monitoring programs can be developed.

# V CONCLUDING COMMENTS

This report has described a scene-analysis approach for automating an important class of remote sensing tasks involving long-term monitoring of predefined ground sites. The key idea is the use of map knowledge to help locate the monitoring sites in an image. Knowing where to look often makes it possible to:

(1) Drastically reduce computation,

(2) Transform complex interpretation tasks into simple detection problems,

(3) Significantly enhance the utility of results by associating them with specific, geographically localized entities.

With map-guidance, many previously intractable monitoring tasks become feasible and, in some cases, even easy to automate.

The location of monitoring sites is accomplished by calibrating an analytic camera model on distinctive landmarks and then using the model to transform between reference map and image coordinates. This approach has significant advantages over the conventional one of warping the sensed image into correspondence with a reference image, most notably in terms of reduced computational requirements and increased invariance to viewing conditions.

The examples in this report should be regarded as demonstrations of concept as opposed to feasibility; further engineering studies are required to evaluate costs and performance under operational conditions. If such experiments substantiate our preliminary conclusions, the consequent payoff may be substantial.

The accelerating rate at which image data are being collected poses an increasingly difficult problem for NASA and other government agencies. Already, the volume of data collected far exceeds the

resources available for analysis. It has been estimated, for example, that only about one percent of the image data currently collected by LANDSAT is ever interpreted. Computer resources are saturated just keeping up with the routine data-processing aspects of image acquisition (rectification, storage, retrieval, distribution), and have remained so, despite years of investment in ever faster and more expensive data-processing hardware.

Buried within the masses of collected data is a relatively small amount of valuable information that justifies the whole enterprise of remote sensing. Clearly, the rich promise of remote sensing will not be realized unless an automatic means can be developed for extracting this information and distributing it in timely fashion to interested users.

Two possible scenarios for accomplishing this objective, based on concepts developed in this report, are outlined in Figure 14. In the first scenario, existing ground-based, data-processing facilities would be augmented to perform a variety of map-guided monitoring tasks automatically as new imagery arrived. The incremental computational load would be very modest; a prototype system could be in operation as early as 1980.

The second scenario is longer range (10 to 15 years) and envisions a series of application-specific satellites, with sensors and orbits optimized for particular monitoring tasks. Information could be extracted on board and relayed direct to interested users via communication satellites.

The feasibility of on-board processing rests in part on the dramatic advances anticipated in LSI technology and in part on the dramatic reductions in computational requirements made possible by the concept of map-guided image analysis. The principal advantage of on-board processing is that it completely eliminates ground-based, data-processing bottlenecks for those monitoring tasks where the actual imagery is not essential.

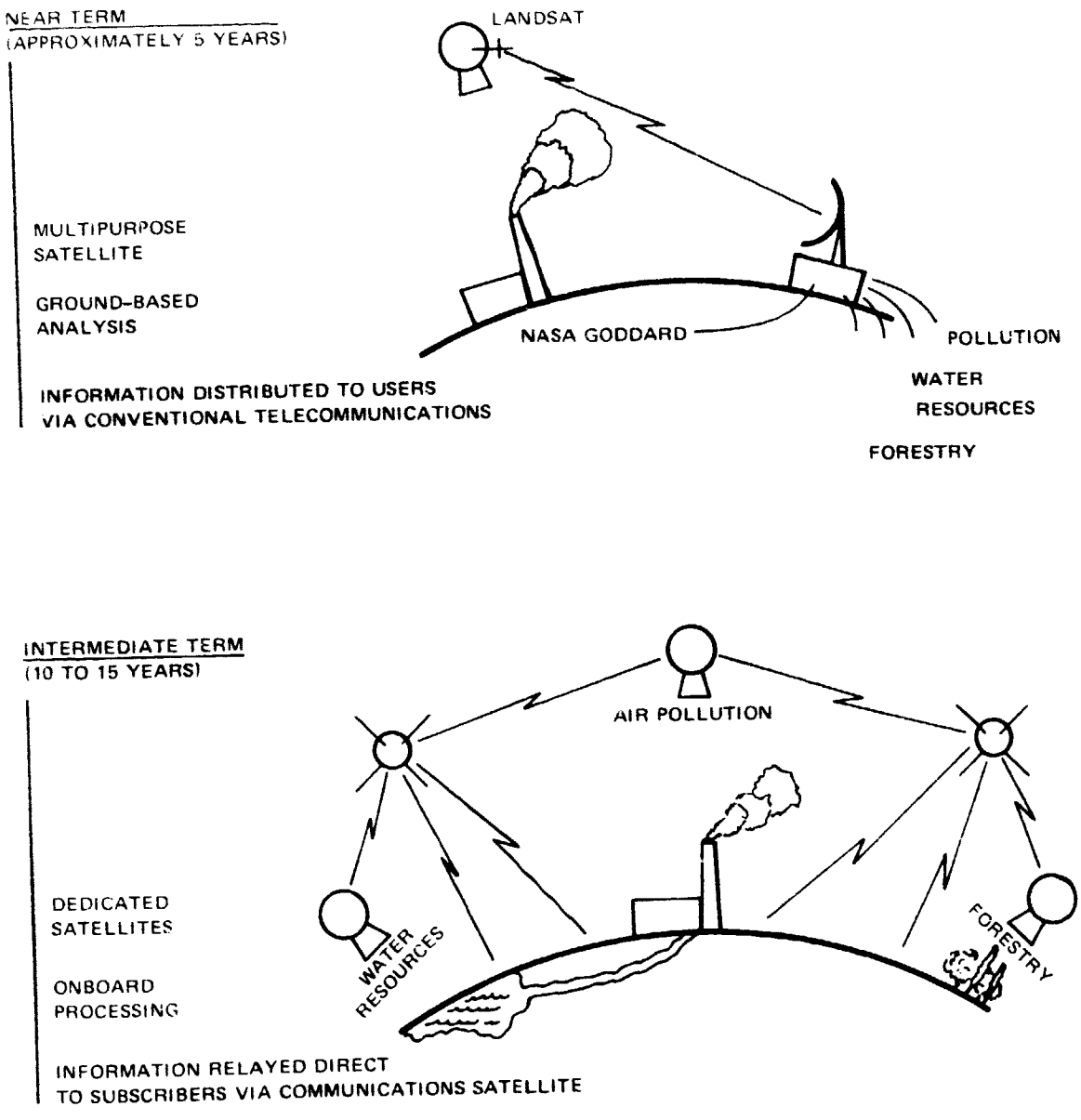



FIGURE 14 SCENARIOS FOR EXPLOITING SCENE ANALYSIS IN AUTOMATED MONITORING TASKS

Both scenarios represent an important broadening of NASA's traditional role as a supplier of data to encompass the additional responsibilities of extraction and distribution of information. For routine monitoring tasks with large user constituencies, centralized information extraction should significantly reduce the overheads of storing, retrieving, and distributing large volumes of data. Moreover, it would eliminate the need for installing image analysis facilities at many user sites. Information extraction and distribution seem destined to play increasingly important roles in NASA's future.

## REFERENCES


1. Manual of Remote Sensing, Leonard W. Bowden, ed., Chapt. 12 (American Society of Photogrammetry, Falls Church, Virginia, 1975).

2. Manual of Remote Sensing, op. cit., Chapt. 22.

3. Barrow, H. G., and J. M. Tenenbaum, "Representation and Use of Knowledge in Vision," Technical Note 108, Stanford Research Institute, Menlo Park, California (July 1975) [an expanded version of a paper that originally appeared in Sigart, No. 52 (June 1975)].

4. Bernstein, R., "Digital Image Processing of Earth Observation Sensor Data," IBM Journal of Research and Development, Vol. 20, No. 1 (January 1976).

5. Barrow, H. G., "Interactive Aids for Cartography and Photo Interpretation," Semiannual Technical Report (Appendix A), Contract DAAG29-76-C-0057, SRI Project 5300, SRI International, Menlo Park, California (December 1977).

6. Sobel, I., "On Calibrating Computer Controlled Cameras for Perceiving 3-D Scenes," Artificial Intelligence, Vol. 5, pp. 185-198 (1974).

7. Duda, R., and P. Hart, Pattern Classification and Scene Analysis (John Wiley & Sons, Inc., New York, New York, 1973).

8. Manual of Remote Sensing, op. cit., Chapts. 19, 20, and 21.

9. Barrow, H. G., "Interactive Aids for Cartography and Photo Interpretation," Semiannual Technical Report, Contract DAAG29-76-C-0057, SRI Project 5300, Stanford Research Institute, Menlo Park, California (November 1976).

10. Barrow, H. G., "Interactive Aids for Cartography and Photo Interpretation," Semiannual Technical Report, Contract DAAG29-76-C-0057, SRI Project 5300, SRI International, Menlo Park, California (June 1978).

11. Barrow, H. G., "Interactive Aids for Cartography and Photo Interpretation," Semiannual Technical Report, Contract DAAG29-76-C-0057, SRI Project 5300, Stanford Research Institute, Menlo Park, California (May 1977).

## Appendix A

# PARAMETRIC CORRESPONDENCE AND CHAMFER MATCHING: TWO NEW TECHNIQUES FOR IMAGE MATCHING

(previously published as Technical Note 153)

Appendix A

# PARAMETRIC CORRESPONDENCE AND CHAMFER MATCHING: TWO NEW TECHNIQUES FOR IMAGE MATCHING

(previously published as Technical Note 153)

## 1. Introduction

Many tasks involving pictures require the ability to put a sensed image into correspondence with a reference image or map. Examples include vehicle guidance, photo interpretation (change detection and monitoring), and cartography (map updating). The conventional approach is to determine a large number of points of correspondence by correlating small patches of the reference image with the sensed image. A polynomial interpolation is then used to estimate correspondence for arbitrary intermediate points [Bernstein]. This approach is computationally expensive and limited to cases where the reference and sensed images were obtained under similar viewing conditions. In particular, it cannot match images obtained from radically different viewpoints, sensors, or seasonal or climatic conditions; and it cannot match images against symbolic maps.

Parametric correspondence matches images to a symbolic reference map, rather than a reference image. The map contains a compact three-dimensional representation of the shape of major landmarks, such as coastlines, buildings, and roads. An analytic camera model is used to predict the location and appearance of landmarks in the image, generating a projection for an assumed viewpoint. Correspondence is achieved by adjusting the parameters of the camera model (i.e., the assumed viewpoint) until the appearances of the landmarks optimally match a symbolic description extracted from the image.

The success of this approach requires the ability to rapidly match predicted and sensed appearances after each projection. The matching of image and map features is performed by a new technique, called "chamfer

matching", that compares the shapes of two collections of curve fragments at a cost proportional to linear dimension, rather than area.

In principle, this approach should be superior, since it exploits more knowledge of the invariant three-dimensional structure of the world and of the imaging process. At a practical level, this permits matching of spatially extensive features on the basis of shape, which reduces the risk of ambiguous matches and dependence on viewing conditions.

## 2. Chamfer Matching

Point landmarks, such as intersections or promontories, are represented in the map with their associated three-dimensional world coordinates. Linear landmarks, such as roads or coastlines, are represented as curve fragments with associated ordered lists of world coordinates. Volumetric structures, such as buildings or bridges, are represented as wire-frame models.

From a knowledge of the expected viewpoint, a prediction of the image can be made by projecting world coordinates into corresponding image coordinates, suppressing hidden lines. The problem in matching is to determine how well the predicted features correspond with image features, such as edges and lines.

The first step is to extract image features by applying edge and line operators or tracing boundaries. Edge fragment linking [Nevatia, Perkins] or relaxation enhancement [Zucker, Barrow] is optional. The net result is a feature array each element of which records whether or not a line fragment passes through it. This process preserves shape information and discards greyscale information, which is less invariant.

To correlate the extracted feature array directly with the predicted feature array would encounter several problems: The correlation peak for two arrays depicting identical linear features is very sharp and, therefore, intolerant of slight misalignment or distortion (e.g., two lines, slightly rotated with respect to each other, can have at most one point of correspondence) [Andrus]; a sharply

peaked correlation surface is an inappropriate optimization criterion because it provides little indication of closeness to the true match nor of the proper direction in which to proceed; computational cost is heavy with large feature arrays.

A more robust measure of similarity between the two sets of feature points is the sum of the distances between each predicted feature point and the nearest image point. This can be computed efficiently by transforming the image feature array into an array of numbers representing distance to the nearest image feature point. The similarity measure is then easily computed by stepping through the list of predicted features and simply summing the distance array values at the predicted locations.

The distance values can be determined in two passes through the image feature array by a process known as "chamfering" [Munson, Rosenfeld]. The feature array (F[i,j], i,j=1,N) is initially two-valued: 0 for feature points and infinity otherwise. The forward pass modifies the feature array as follows:

```
FOR i _ 2 STEP 1 UNTIL N DO
  FOR j _ 2 STEP 1 UNTIL N DO
    F[i,j] _ MINIMUM(F[i,j], (F[i-1,j]+2),
               (F[i-1,j-1]+3), (F[i,j-1]+2),
               (F[i+ . -1]+3));
```

Similarly, the backward pass operates as follows:

```
FOR i _ (N-1) STEP -1 UNTIL 1 DO
  FOR j _ (N-1) STEP -1 UNTIL 1 DO
    F[i,j] _ MINIMUM(F[i,j], (F[i+1,j]+2),
               (F[i+1,j+1]+3), (F[i,j+1]+2),
               (F[i-1,j+1]+3));
```

The incremental distance values of 2 and 3 provide relative distances that approximate the Euclidean distances 1 and the square root of 2.

Chamfer matching provides an efficient way of computing the integral distance (i.e., area), or integral squared distance, between two curve fragments, two commonly used measures of shape similarity. Note that the distance array is computed only once, after image feature extraction.

## 3. Parametric Correspondence

Parametric correspondence puts an image into correspondence with a three-dimensional reference map by determining the parameters of an analytic camera model (3 position and 3 orientation parameters).

The traditional method of calibrating the camera model takes place in two stages: first, a number of known landmarks are independently located in the image; and second, the camera parameters are computed from the pairs of corresponding world and image locations by solving an over-constrained set of equations [Sobel, Quam, Hannah].

The failings of the traditional method stem from the first stage. The landmarks are found individually, using only very local context (e.g., a small patch of surrounding image) and with no mutual constraints. Thus local false matches commonly occur. The restriction to small features is mandated by the high cost of area correlation and by the fact that large image features correlate poorly over small changes in viewpoint.

Parametric correspondence overcomes these failings by integrating the landmark-matching and camera-calibration stages. It operates by hill-climbing on the camera parameters. A transformation matrix is constructed for each set of parameters considered; and it is used to project landmark descriptions from the map onto the image at a particular translation, rotation, scale and perspective. A similarity score is computed with chamfer matching and used to update parameter values. Initial parameter values are estimated from navigational data.

Integrating the two stages allows the simultaneous matching of all landmarks in their correct spatial relationships. Viewpoint problems with extended features are avoided because features are precisely projected by the camera model prior to matching. Parametric correspondence has the same advantages as rubber-sheet template matching [Fischler, Widrow] in that it obtains the best embedding of a map in an image but avoids the combinatorics of trying arbitrary distortions by only considering those corresponding to some possible viewpoint.

## 4. An Example

The following example illustrates the major concepts in chamfer matching and parametric correspondence. A sensed image (Figure A-1) was input along with manually derived initial estimates of the camera parameters. A reference map of the coastline was obtained, using a digitizing tablet to encode coordinates of a set of 51 sample points on a USGS map. Elevations for the points were entered manually. Figure A-2 is an orthographic projection of this three-dimensional map.

A simple edge follower traced the high contrast boundary of the harbor, producing the edge picture shown in Figure A-3. The chamfering algorithm was applied to this edge array to obtain a distance array. Figure A-4 depicts this distance array; distance is encoded by brightness with maximum brightness corresponding to zero distance from an edge point.

Using the initial camera parameter estimates, the map was projected onto the sensed image (Figure A-5). The average distance between projected points and the nearest edge point, as determined by chamfer matching, was 25.8 pixels.

A straightforward optimization algorithm adjusted the camera parameters, one at a time, to minimize the average distance. Figures A-6 and A-7 show an intermediate state and the final state, in which the average distance has been reduced to 0.8 pixel. This result, obtained with 51 sample points, compares favorably with a 1.1 pixel average distance for 19 sample points obtained using conventional image chip correlation followed by camera calibration. The curves in Figure A-8 characterize the local behavior of this minimum, showing how average distance varies with variation of each parameter from its optimal value. Approximately 60 iterations (each involving a parameter adjustment and reprojection) were required for this example. The number of iterations could be reduced by using a better optimization algorithm; for example, a gradient search.

## 5. Discussion

We have presented a scheme for establishing correspondence between an image and a reference map that integrates the processes of landmark matching and camera calibration. Tne potential advantages of this approach stem from 1) matching shape, rather than brightness, 2) matching spatially extensive features, rather than small patches of image, 3) matching simultaneously to all features, rather than searching the combinatorial space of alternative local matches, and 4) using a compact three-dimensional model, rather than many two-dimensional templates.

Shape has proved to be much easier to model and predict than brightness. Shape is a relatively invariant geometric property whose appearance from arbitrary viewpoints can be precisely predicted by the camera model. This eliminates the need for multiple descriptions, corresponding to different viewing conditions, and overcomes difficulties of matching large features over small changes of viewpoint.

The ability to treat the entirety of the relevant portion of the reference map as a single extensive feature reduces significantly the risk of ambiguous matches and avoids the combinatorial complexity of finding the optimal embedding of multiple local features.

A number of obstacles have been encountered in reducing the above ideas to practice. The distance metric used in chamfer matching provides a smooth, monotonic measure near the correct correspondence and nicely interpolates over gaps in curves. However, scores can be unreliable when image and reference are badly out of alignment. In particular, discrimination is poor in textured areas, aliasing can occur with parallel linear features, and a single isolated image feature can support multiple reference features.

The main problem is that edge position is not a distinguishing feature, and consequently many alternative matches receive equal weight. One way of overcoming this problem, therefore, is to use more descriptive features: brightness discontinuities can be classified, for

example, by orientation, by edge or line, and by local spatial context (texture versus isolated boundary). Each type of feature would be separately chamfered, and map features would be matched in the appropriate array. Similarly, features at a much higher level could be used, such as promontory or bay, area features having particular internal textures or structures, and even specific landmarks, such as "the top of the Transamerica pyramid". Ideally, with a few highly differentiated features distributed widely over the image, the parametric correspondence process would be able to home in directly on the solution regardless of initial conditions.

Another dimension for possible improvement is the chamfering process itself. Determining for each point of the array a weighted sum of distances to many features (e.g., a convolution with the feature array), instead of the distance to the nearest feature, would provide more immunity from isolated noise points. Alternatively, propagating the coordinates of the nearest point instead of merely the distance to it, it becomes possible to use characteristics of features, such as local slope or curvature, in evaluating the goodness of match. It also makes possible a more directed search, since corresponding pairs of points are now known; an improved set of parameter estimates can be analytically determined.

Chamfer matching and parametric correspondence are separable techniques. Conceptually, parametric correspondence can be performed by re-projecting image chips and evaluating the match with correlation. However, the cost of projection and matching grows with the square of the template size: the cost for chamfer matching grows linearly with the number of feature points. Chamfer matching is an alternative to other shape-matching techniques, such as chain-code correlation [Freeman], Fourier matching [Zahn], and graph matching [e.g., Davis]. Also, the smoothing obtained by transforming two edge arrays to distance arrays via chamfering can be used to improve the robustness of conventional area-based edge correlation.

Parametric correspondence, in its most general form, is a technique for matching two parametrically related representations of the same geometric structure. The representations can be two- or three-dimensional, iconic or symbolic; the parametric relation can be perspective projection, a simple similarity transformation, a polynomial warp, and so forth. This view is similar to rubber-sheet template matching as conceived by Fischler and Widrow [Fischler, Widrow]. The feasibility of the approach in any application, as Widrow points out, depends on efficient algorithms for "pattern stretching, hypothesis testing, and pattern memory", corresponding to our camera model, chamfer matching, and three-dimensional map.

As an illustration of its versatility, the technique can be used with a known camera location to find a known object whose position and orientation are known only approximately. In this case, the object's position and orientation are the parameters; the object is translated and rotated until its projection best matches the image data. Such an application has a more iconic flavor, as advocated by Shepard [Shepard], and is more integrated than the traditional feature extraction and graph matching approach [Roberts, Falk and Grape].

As a final consideration, the approach is amenable to efficient hardware implementation. There already exists commercially available hardware for generating parametrically specified perspective views of wire frame models at video rates, complete with hidden line suppression. The chamfering process itself requires only two passes through an array by a local operator, and match scoring requires only summing table lookups in the resulting distance array.

## 6. Conclusion

Iconic matching techniques, such as correlation, are known for efficiency and precision obtained by exploiting all available pictorial information, especially geometry. However, they are overly sensitive to changes in viewing conditions and cannot make use of non-pictorial information. Symbolic matching techniques, on the other hand, are more

robust because they rely on invariant abstractions; but they are less precise and less efficient in handling geometrical relationships. Their applicability in real scenes is limited by the difficulty of reliably extracting the invariant description. The techniques we have put forward offer a way of combining the best features of iconic and symbolic approaches.

## 7. Acknowledgments

This work was supported by ARPA under Contract DAAG29-76-C-0012. Additional support was provided by NASA under Contract NASW-2865.

## REFERENCES


Andrus, J.F., Campbell, C.W. and Jayroe, R.R., "A Digital Image Registration Method Using Boundary Maps", IEEE Trans. Comp., Sept. 1975, p. 935-939.

Barrow, H.G., "Interactive Aids for Cartography and Photo Interpretation", Interim Report of ARPA Project DAAG29-76-C-0057, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, California, Dec. 1976.

Bernstein, R., "Digital Image Processing of Earth Observation Sensor Data", IBM Journal of Research and Development, Vol. 20, No. 1, 1976.

Davis, L., "Shape Matching Using Relaxation Techniques", TR-480, Computer Science Dept., University of Maryland, Sept. 1976.

Falk, G., "Interpretation of Imperfect Line Data as a Three Dimensional Scene," Artificial Intelligence, Vol. 3, 1972, p.101-144.

Feder, J. and Freeman, H., "Segment Fitting of Curves in Pattern Analysis Using Chain Correlation", AD619525, March 1965.

Fischler, M. and Elschlager, R., "The Representation and Matching of Pictorial Structures", IEEE Trans. Comp., no. 22, 1973, p.67-92.

Grape, G.R., "Model Based (Intermediate Level) Computer Vision", Stanford AI Memo AIM-201, May 1973, Ph.D. Thesis.

Hannah, M.J., "Computer Matching of Areas in Stereo Images", Stanford AI Memo AIM-239, July 1974, Ph.D. Thesis.

Munson, J., Internal memo, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, California, Dec. 1973.

Perkins, W.P., "Multi-level Vision Recognition System", Proc. Third IJCPR, Nov. 1976, p.739-744.

Quam, L.H., "Computer Comparison of Pictures", Stanford AI Project Memo AIM-144, May 1971, Ph.D. Thesis.

Roberts, L.G., "Machine Perception of Three-dimensional Objects", in: Tippet, J.T., et al. (Eds.), "Optical and Electro-Optical Information Processing", MIT Press, 1965, p.159-197.

Rosenfeld, A. and Pflatz, J.L., "Distance Functions on Digital Pictures", Pattern Recognition, Vol. 1 No. 1, July 1968, p.33-62.

Shepard, R.N. and Metzler, J., "Mental Rotation of Three-Dimensional Objects", Science, 1971, p.701-703.

Sobel, I., "On Calibrating Computer Controlled Cameras for Perceiving 3-D Scenes", Artificial Intelligence, Vol. 5, 1974, p.185-198.

Widrow, b., "The 'Rubber-Mask' Technique ", Pattern Recognition, Vol. 5, 1953, p.175-211.

Zahn, C. and Roskies, R., "Fourier Descriptors for Plane Closed Curves", IEEE Trans. Comp., no. 21, 1972, p.269-281.

Zucker, S., Hummel, R. and Rosenfeld, A., "An Application of Relaxation Labeling to Curve and Line Enhancement", IEEE Trans. Comp., no. 25, 1976.

FIGURE A-1 AERIAL IMAGE OF A SECTION OF COASTLINE

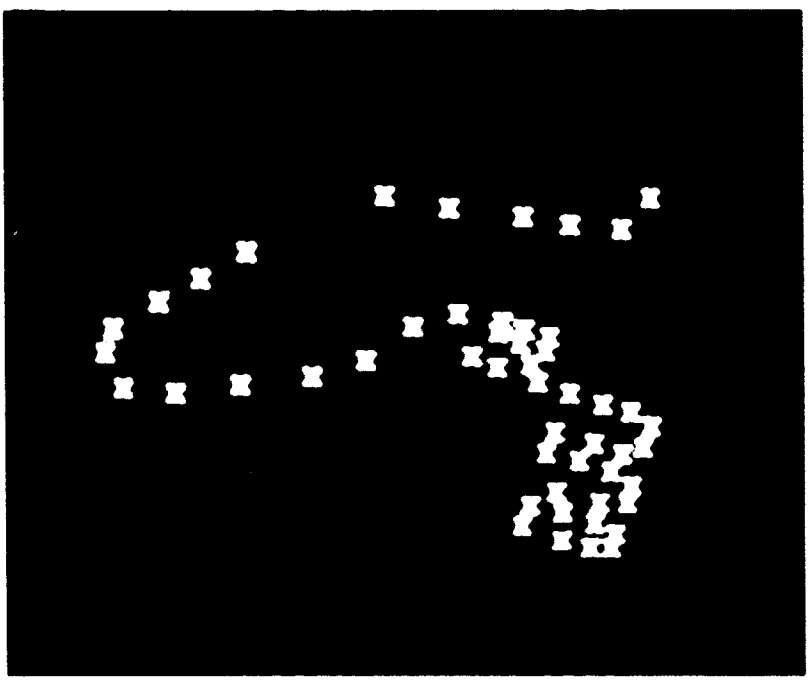


FIGURE A-2 SET OF SAMPLE POINTS TAKEN FROM A USGS MAP

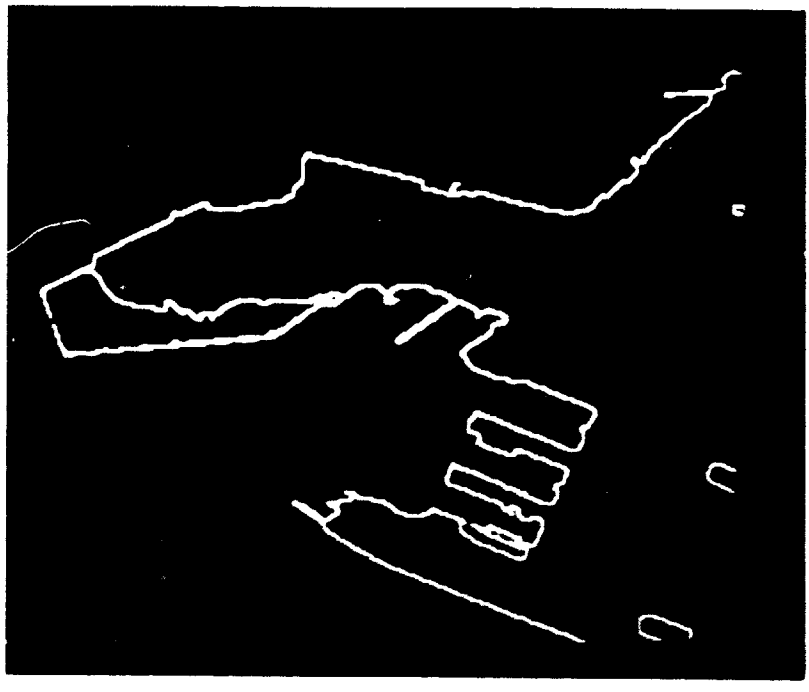


FIGURE A-3 THE TRACED BOUNDARY OF THE COASTLINE

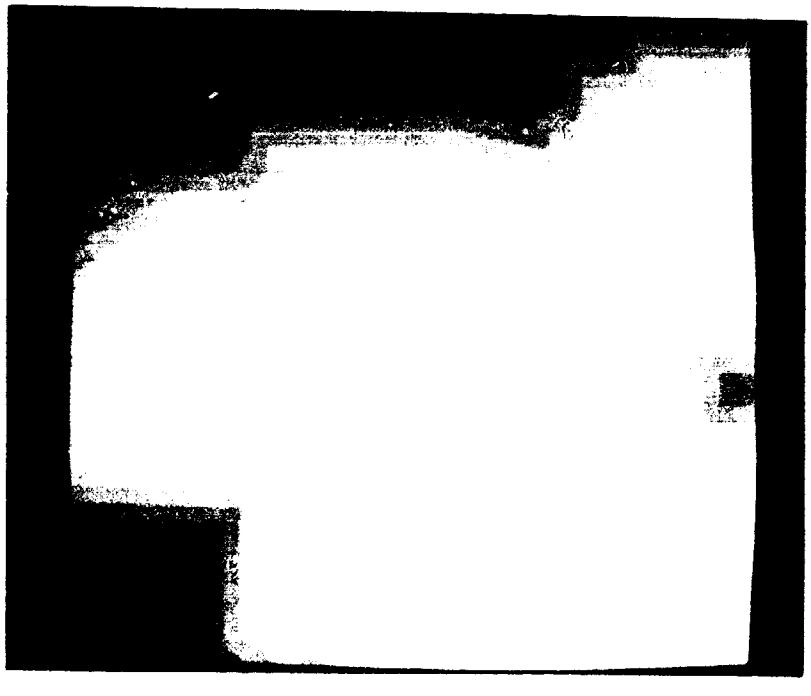


FIGURE A-4 DISTANCE ARRAY PRODUCED BY CHAMFERING THE BOUNDARY

FIGURE A-5 INITIAL PROJECTION OF MAP POINTS ONTO THE IMAGE

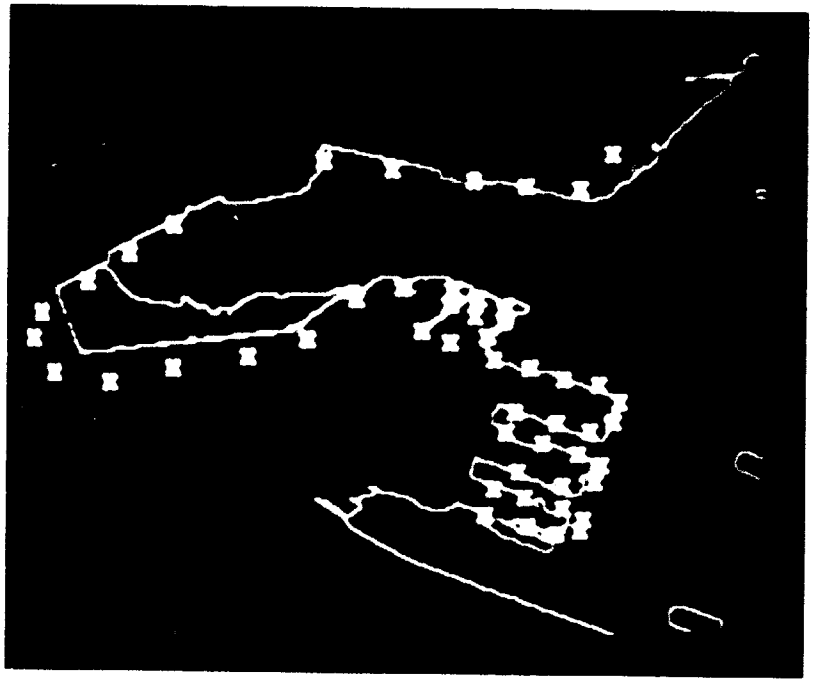




FIGURE A-6 PROJECTION OF MAP POINTS ONTO THE IMAGE AFTER SOME ADJUSTMENT OF CAMERA PARAMETERS

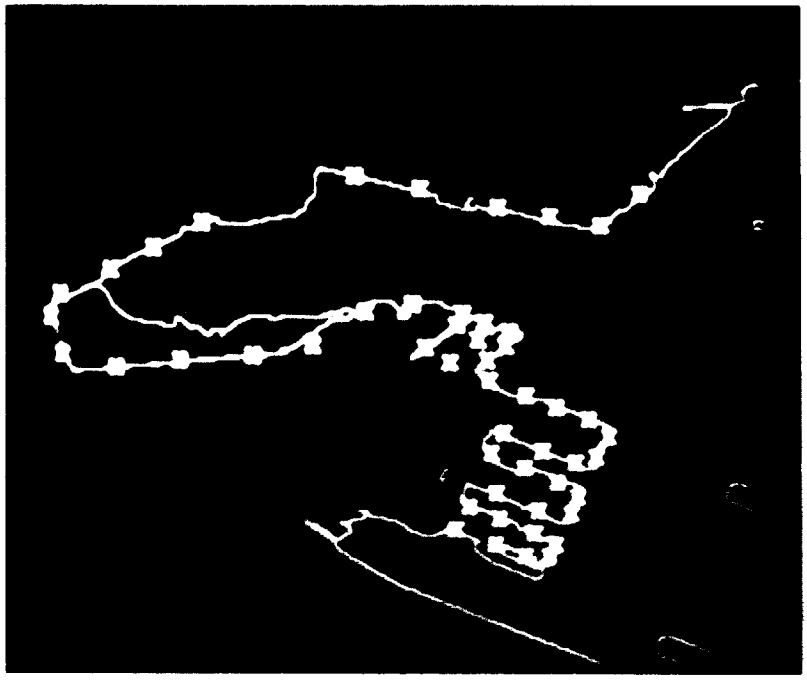


FIGURE A-7 PROJECTION OF MAP POINTS ONTO THE IMAGE AFTER OPTIMIZATION OF CAMERA PARAMETERS

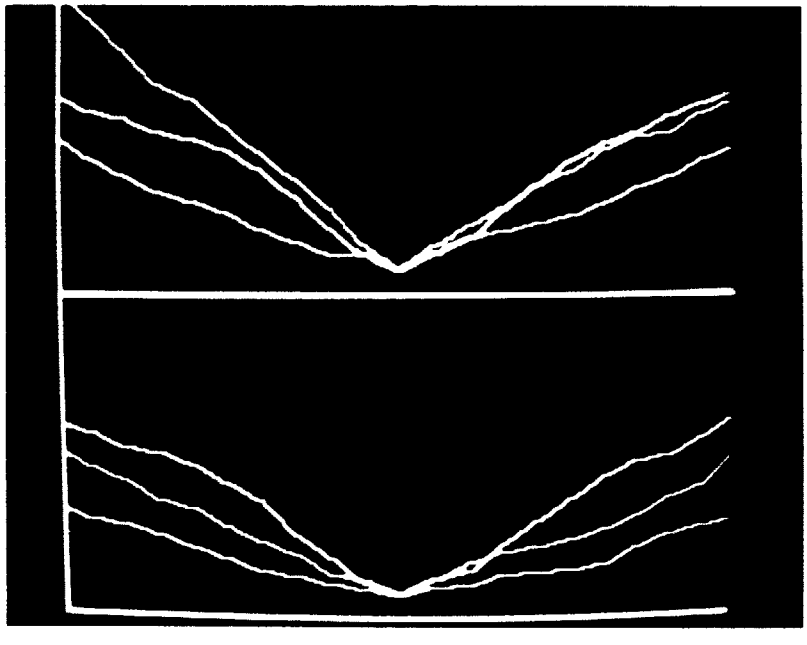


FIGURE A-8 BEHAVIOR OF AVERAGE DISTANCE SCORE WITH VARIATION OF THE SIX CAMERA PARAMETERS FROM THEIR OPTIMAL VALUES

# Appendix B

## USING MAP KNOWLEDGE TO LOCATE BOUNDARIES TO BETTER-THAN-IMAGE-RESOLUTION ACCURACY

Appendix B

# USING MAP KNOWLEDGE TO LOCATE BOUNDARIES TO BETTER-THAN-IMAGE-RESOLUTION ACCURACY

## 1. Introduction

There are many remote sensing applications in which one would like to locate the boundaries of known objects or terrain features to a precision exceeding the available image resolution. Typically, this situation occurs when there is a requirement to make precise measurements but the resolution required to make the measurements directly is excessive or impractical.

This appendix discusses an approach for automatically determining the boundary location of known objects to "subpixel" accuracy and describes an experiment in which it was applied to the problem of using aerial photography to measure water levels of reservoirs.

To determine boundary location beyond the available image resolution, some additional information or constraint must be used as the basis for making the desired determination by indirect means. In this appendix it will be assumed that this additional information is knowledge of the exact shape of the relevant features or objects, up to some unknown parameter. For example, the object of interest might be known to be a circle (or cylinder), and we wish to determine its diameter. In the example to be described later, we employed a (supposedly) very accurate contour map of a geographic area containing a reservoir, which provides the possible shapes that could be assumed by the reservoir land/water boundary as a function of water level. Global shape matching provides the means for accurately determining boundary location and thus water level. Knowing water level, the contour map can then be used to compute the volume of stored water, if this is desired.

## 2. Fundamental Issues

The problem of determining boundary location can be partitioned into two separate tasks: making local estimates of boundary location, then combining these local estimates to make a globally best selection from a given set of possible boundary shapes. As we will discuss later, the critical requirement in making the local boundary estimates is that the errors be unbiased about the "true" boundary location. This can only be ensured by having a valid model of the relation between the nature of the actual boundary and its appearance in the image.

In greater detail, the issues in local boundary estimation are:

- What is a suitable model for the appearance of an edge in an aerial image (ideal appearance perturbed by the noise and distortions of the imaging process)?
- What is the relationship between the appearance of the edge in the image and its actual location referenced to the image plane (where is the precise edge transition, given the intensity pattern indicating an edge)?
- What is a suitable algorithmic technique for locating the actual edge boundaries automatically (the description of a decision procedure for estimating edge point location in a noisy, sampled, quantized image)?
- What is the nature of the errors we can expect from our decision procedure, and how can we experimentally verify its performance?

Given unbiased local boundary estimates (and assuming that the errors in our reference shapes are also unbiased), the problem of making an optimal global decision has three aspects:

- Selecting the criteria for deciding which choice is best (we will employ a maximum likelihood criteria).
- Detecting and eliminating bad data (i.e., local boundary estimates, which are in obvious disagreement with the reference shapes).
- Matching the reference shapes to the local boundary estimates.

An effective algorithm for accomplishing the above has been developed and will be described later.

## 3. Modeling Edge Appearance

Two factors must be considered in making local boundary estimates: the appearance of boundary as determined by the characteristics of the imaging sensor (as well as by the physical nature of the boundary itself) and the alterations in boundary "appearance" introduced by the digitization process.

Physical factors affecting appearance are primarily edge sharpness (i.e., the "length" of the transition region), edge contrast (i.e., intensity variation across the edge), and edge context (i.e., the nature of the intensity variations in the image in the vicinity of the edge).

**Alterations in appearance resulting from the digitization process are primarily a blurring of edge sharpness (lengthening of the transition region) and possibly a reduction in contrast. Both of these effects result from the low-pass filtering effect introduced by the finite size of the sampling window used to accomplish the digitization step (see Figure B-5).**

### a. General Edge Model

Ideally, we might assume that the local presence of an edge in an image is indicated by a step discontinuity (see Fig. B-1) of some combination of image attributes, such as intensity or color or texture; however, this ideal edge is seldom encountered. In practice, the actual edge appearance is affected by the spatial resolution and sampling effects introduced by the sensor and digitizer, the spectral response of sensor, and the illumination and reflectance characteristics of the scene. For example, if the image were obtained by a system having limited bandwidth, then we would expect that the imaged edge of a physical object would result in an intensity ramp similar to that shown in Figure B-2. Figure B-3 shows some actual intensity profiles through a number of edge and nonedge features.

b. A Specialized Edge Model

Rather than attempting to define a completely general edge profile that is suitable for all types of edges and environments, experience has shown that it makes more sense to model the specific types of edges we are interested in. In this appendix we will consider a specialized edge model that still has a significant residue of generality; in particular, we will now define an edge model appropriate to a land/water interface or to the boundary between a man-made object (e.g., a building or road) and the surrounding natural terrain.

This model (see Figure B-4) specifies the two "aprons" adjacent to the ideal edge ramp of Figure B-2; one apron having a relatively low variance intensity profile (water, man-made object), the second apron characterized by a distinct mean value and possibly a high variance (the natural terrain). The flat (low variance) apron helps us to locate the onset of the intensity ramp and distinguish the edge from nonedge.

c. Physical Factors Affecting Edge Appearance

Now let us consider the case of a real edge representing a land/water interface. Unless we had recorded our image on infrared film, there is the possibility of some penetration of the water surface by light in the red-green bands; thus, the onset of the ramp--even assuming infinite bandwidth--could occur on the water side of the actual boundary line. Since we can easily solve this problem by a proper choice of the spectral band used to create the image, we will assume that the water is essentially opaque to light. We further note that the nonuniform intensity of the land could result in the intensity ramp extending past the actual boundary line, well into the land area.

The land/water edge model indicated in Figure B-4 shows the water as being dark as compared to the land. While this will generally be the case, the model can be trivially extended (or the image intensities inverted) to deal with the reverse situation. It should be noted, however, that to ensure good land/water contrast, we would like

to avoid specular reflection from the surface of the water. This can be accomplished by an appropriate combination of viewing and sun locations in acquiring the imagery (or by invoking a sensor that is not dependent on reflected light).

d. Effects of Digitization on Edge Appearance

For almost any reasonable set of camera, film, and digitization parameters, the greatest limitation on bandwidth will result from the sampling window used in the digitization process; the width of the ramp resulting from a step function edge will thus be at most two pixels, assuming linear interpolation between the digitized intensity values. Thus, regardless of the nature of the apron appended to the edge ramp by the land intensity profile, we can assert that the true image location of the edge denoting a land/water interface will lie within the interval (0.5 - 1.5) pixels from the water end of the intensity ramp (see Figure B-5). With reasonable land/water contrast compared to the variance in intensity over the water area, we should be able to locate the foot of the intensity ramp to within a one pixel interval of uncertainty. We can, therefore, expect to be able to locate the land/water boundary along a single intensity profile to an error of less than plus or minus one pixel.

e. Significance of Modeling the Precise Edge Location in Relation to its Imaged Appearance

The primary reason for our interest in modeling the precise edge location as part of our land/water boundary model is that, while a decision procedure based on multiple measurements can tolerate any amount of "unbiased" error given a sufficient number of measurements, the ultimate accuracy attainable (in estimating a continuous variable) is limited by the unknown bias associated with the individual measurements. We are actually less concerned with accuracy than with bias; however, to control or estimate bias, we must (at least theoretically) know the correct answer. To the extent that we must be

satisfied with asserting that the correct answer lies somewhere in an interval, a consistent bias can result in a final error of one-half of that interval.

As discussed in the preceding section, two factors can introduce a consistent bias into our model: the assumption that there is no significant penetration of the water in the spectral band of the light used to create the image and the assumption about how the edge intensity ramp is created by the digitization process acting on the actual imaged edge. If both of these assumptions are valid, then our model will permit us to obtain an unbiased estimate of the mean edge location.

## 4. An Algorithm for Model-Based Edge Detection

The algorithm we will now describe has two parts--the first part, concerned with local boundary detection, is based on the edge model described in the preceding section; the second part is a procedure for combining a-priori shape information (elevation contours) with the local land/water boundary estimates to make a globally optimal estimate of a shape parameter (i.e., the elevation associated with the contour shape that best matches the reservoir shape as detected in the image).

### a. Detecting the Land/Water Interface

This algorithm, in addition to assuming the edge model described earlier, invokes map knowledge to supply the approximate location and orientation of the edges of interest; i.e., our model is augmented by position and orientation attributes.

The algorithm operates by first placing the image to be analyzed into correspondence with a map data base using a camera calibration procedure of the type described in Appendix A. In addition to the elevation contour information, the map data base contains the coordinates for a set of lines that are approximately normal to the elevation contours at selected locations around the nominal reservoir boundary we are attempting to detect. The problem of local boundary

detection is now reduced to determining where each of these selected lines, when projected onto the image, intersects the land/water interface.

Because of edge continuity at right angles to the line along which we are searching (i.e., the projected normal), we can effectively search for a section of edge boundary, rather than a single edge point, by (bilinearly) interpolating the surrounding intensity values onto the projected normal (see Figure B-6).

We now invoke our edge model by scanning along the normal, starting from well into the water, and searching for a significant jump in intensity. Typical values of the standard deviation of intensity in the water are 5 or less, while the intensity difference moving across the land/water interface is usually on the order of at least 20 units (our images are quantized to eight bits of gray scale). Once an intensity rise of three or more standard deviations has been detected, we mark the estimated boundary point either at that location or at the succeeding pixel location, if the intensity slope is even greater there.

b. Estimating A Globally Optimal Shape Parameter

Given the set of boundary points found using the algorithm described above, we can now project the contour information from our map data base onto the normals and associate an elevation value with each of the edge points. If our assumption that the model introduces no systematic biases is valid, then a simple average of a sufficient number of edge point elevations should provide a reasonable approximation to the desired answer (i.e., reservoir water level elevation).

It is important to note here that even if the edge model provides unbiased estimates in distance from the true edge location in the image plane, this does not necessarily imply that the associated elevation values are also unbiased (i.e., the elevation gradient in the image plane is not constant). Furthermore, even if our model does permit unbiased estimates of the edge points in both image plane distance and elevation, the local boundary detection algorithm can fail

because of conditions not considered in the edge model. For example, small clouds may obscure portions of the boundary we are searching for; or vegetation changes may eliminate the anticipated contrast between water and land; or a man-made or natural alteration in the land profile may cause the contour-map data to be locally incorrect. Any of the preceding conditions can cause one or more of the edge points to be significantly displaced from its reference location and, in fact, can introduce large systematic errors into the averaging computation. Therefore, the simple average of local boundary point elevations is not a generally reliable decision procedure by itself.

To better deal with the problems caused by local deviations from the model, we choose to view the problem of contour selection as a discrimination (rather than a parameter-estimation) task. We first obtain a rough estimate of the desired parameter (elevation associated with the reservoir boundary) using the simple average. We now take two successive contours near the computed average value and determine which of the two has a shape most like the boundary detected in the image. The best fitting contour is retained and compared with the next adjacent contour. These pairwise discriminations are repeated until the best matching contour has been found.

In the experiment to be described, the contour map had elevation intervals of 5 feet and the image had a resolution of approximately 20 ft./pixel. The ability to discriminate between two contours, which can be closer together than the accuracy with which individual edge points can be located, is based on the assumption of unbiased errors in image plane distance about the contour actually corresponding to the reservoir boundary. Our model indicates that if one of the contours being tested is indeed the correct one, then almost all actual boundary points will fall within one pixel of that contour's projection onto the image plane. We can therefore eliminate from consideration all detected boundary points that do not satisfy this criteria in discriminating between a pair of candidate contours. We also eliminate from consideration those detected boundary points that

fall between the two contours being compared. These points do not offer reliable discrimination information because their location is spatially quantized and the separation between the two candidate contours is typically a fraction of a pixel.

We are thus left with those detected boundary points that lie just outside the interval between the two contours. If we label the two contours as "A" and "B", then those remaining boundary points adjacent to Contour A support the hypothesis that A is more similar to the boundary than B; the converse situation holds for the points adjacent to Contour B.

The simplest discrimination procedure is to choose that contour that has the greatest support, by comparing the number of points adjacent to each contour. A theoretically more powerful technique is to compute the likelihood ratio for the two hypotheses (A is more similar in shape to the boundary than B; B is more similar in shape to the boundary than A). Given that the errors have a normal distribution and that we measure displacements along each normal from an origin located midway between the two candidate contours, then the logorithm of the likelihood ratio is simply the sum of the signed displacements from the origin to the boundary point (x) times the absolute value of the distance separating the two contour lines along the corresponding normal, each of these terms being divided by the variance of the random variable x (see Figure B-7).

If we assume that the standard deviation of x is constant over the different normals, then we accept the hypothesis that A is more similar to the boundary shape than B if the sum of the absolute distances between the contour lines along the normals that support A is greater than the corresponding sum for B. Thus, instead of simply counting support points, we now weight the contribution at each normal with a factor that represents the significance of the corresponding support in choosing between A and B.

An additional refinement would be to drop the assumption that the variance (essentially a reliability factor) has a constant value for

all normals. It is likely that the variance associated with a given normal is inversely related to the intensity difference or gradient across the edge at that normal; however, this conjecture remains to be verified.

## 5. Experimental Results

In an attempt to investigate the feasibility of the proposed approach, we performed an experiment in which the previously described techniques were applied to an aerial photograph containing the Briones reservoir (located north of Oakland, California). We had independent knowledge of the actual water elevation as one measure to use in evaluating our results.

### a. Data Elements

(1) A USGS aerial photograph containing Briones Reservoir:

| | |
|---|---|
| ID Number: | GS-VBZJ-3-21 |
| Date of Photography: | 4-22-68 |
| Time of Day: | 14:10 |
| Altitude: | 15,000 ft. |
| Focal Length: | 6 in. |
| Image Size: | 9 x 9 in. |

(2) A digitized version of the Briones image:

| | |
|---|---|
| Resolution: | 20 ft./pixel |
| Size: | 1024 x 1024 pixels |
| Intensity Levels: | 256 (8 bits) |

(3) An East Bay Municipal Utility District (Oakland, Calif.) contour map of the region containing the Briones Reservoir. This map was produced before the reservoir was built.

| | |
|---|---|
| ID Number: | DH-4613-40 |
| Date of Photography: | 10-25-63 |
| Contour Intervals: | 5 ft. |
| Scale: | 200 ft./in. |
| (5 sheets) | |

(4) The Briones water level elevation was known to be 524.5 ft. on 4-22-68. (This information was obtained from the East Bay Municipal Utility District.)

b. Experimental Procedure

A contour map [Data Item (3)] of the Briones area was placed on a digitizing table, and a full three-dimensional projective transformation was established between it and a digital image of the same area [Data Item (2)]. A set of 136 "normals", at manually selected locations about the reservoir boundary, were then drawn on the contour map; and the coordinates of the 500- to 550-ft. elevation contour lines (11 such lines at 5-ft. elevation intervals) were marked and digitized for each of the normals.

To satisfy the needs of the algorithm, as well as to supply data for some subsidiary goals, reservoir land/water boundary points were obtained on each normal by four different methods:

* The land/water model-based edge finder, described earlier.
* The Hueckel edge detector*, using an 8-pixel diameter window. Here we ran the Hueckel operator over the region of interest and chose the best edge score on each normal.
* Manual marking of the apparent edge location by two persons involved in the experiment.
* The nominal edge locations obtained by projecting the 525-ft. contour line onto the image.

Figures B-8 and B-9 show the overall reservoir site, some representative examples of intensity profiles along the normals, and edge placements by the various techniques.

---

* Hueckel, M. H., "An Operator which Locates Edges in Digitized Pictures," Journal of the Association for Computing Machinery, Vol. 18, No. 1 (January 1971).

c. Analysis of the Experimental Data

Using edge point elevation values determined by interpolating between elevation contours [Data Item (3)], the following statistical results were obtained:

| METHOD | MEAN ELEVATION (feet) | STANDARD DEVIATION (feet) | MINIMUM ELEVATION (feet) | MAXIMUM ELEVATION (feet) |
|---|---|---|---|---|
| Model-based | 521.2 | 4.0 | 512.5 | 530.2 |
| Hueckel | 522.3 | 4.9 | 513.6 | 543.4 |
| Person 1 (HW) | 522.5 | 4.6 | 513.9 | 536.7 |
| Person 2 (MF) | 522.5 | 4.5 | 514.1 | 534.8 |

The statistics for the image plane distances between the model-based detected boundary (DB) and the 520- and 525-ft. contour line are:

| CONTOUR PAIR DIFFERENCED | MEAN SEPARATION (PIXELS) | STANDARD DEVIATION (PIXELS) | MINIMUM DIFFERENCE (PIXELS) | MAXIMUM DIFFERENCE (PIXELS) |
|---|---|---|---|---|
| 525-520 | 1.13 | 0.84 | 0.310 | 9.59 |
| DB-520 | 0.29 | 1.03 | 0.004 | 7.12 |
| 525-DB | 0.84 | 0.82 | 0.010 | 2.47 |

On 65 of the normals, the separation between the 520 and 525 contour lines was less than one pixel, while on the remaining 71 normals the separation was one pixel or greater.

A 90 percent confidence interval about the detected land/water boundary, based on an average standard deviation of 0.922 pixel and 136 sample values, is 0.14 pixel for the mean difference in distance between the detected and map-based contour lines.

The global shape-matching algorithm was applied to the model based edge points and the contours in the set (515-530), resulting in the following scores:

| CONTOUR PAIR | RESPECTIVE SUPPORT |
|---|---|
| 515 vs 520 | 4.4 vs 137.8 |
| 520 vs 525 | 55.4 vs 35.9 |
| 525 vs 530 | 165.3 vs 0.8 |

Conclusion: The reservoir land/water boundary is most similar in shape to the 520-ft. elevation contour line and lies somewhere in the interval between the 520 and 525-ft. elevations.

d. Discussion

The one key data element we did not possess at the start of our experiment was the exact location of the land/water boundary referenced to the image plane. This information was necessary to directly test the validity of our edge model, especially in regard to the question of bias. It was believed that these data could be obtained from the contour map and our knowledge of the actual water level elevation, after we had achieved precise image/map correspondence. However, after a careful manual examination of the data, we came to the conclusion that the 525-ft. elevation contour simply did not have the same shape as the reservoir land/water boundary. Figure B-8(d) shows the projection of the 525 contour line onto the image plane after our best attempt to achieve image/map correspondence. While the general match between the 525-ft. contour line and the reservoir boundary is reasonable, there are a number of obvious discrepancies; thus we could not trust the remaining data to give us the "true" boundary location.

It is apparent that what is really required to directly validate a precision edge model of the type we are concerned with here is an instrumented test site (i.e., the water line surveyed or photographed with visible markers) and test photography acquired at a

resolution well beyond that intended for the actual application. For the limited experiment we had intended to perform, it was impractical to attempt to acquire such data.

Inability to verify our edge model or to determine whether the source of the discrepancy between the 525-ft. contour line and the land/water boundary results primarily from map construction (probably random errors) or to the calibration process (probably systematic errors) are limiting factors on the conclusions we can draw from our experiment. Nevertheless, even though the projection of the map contour data is questionable in regard to providing us with the precise location of the land/water boundary on each individual normal, the distribution of errors may be reasonably compatible with our assumptions, since our final answer was within a few feet of the correct water level elevation.

Given the above considerations, if we are willing to accept the proposed edge model as being valid, then an examination of mean values of water level provided by the different edge finding techniques leads to some interesting insights.

The land/water boundary appears in the image as a spatially compact ramp edge, which is somewhat extended in a random manner on the land side of the ramp. An edge finder based on a general edge model, such as the Hueckel edge finder, will tend to place the edge location at the center of the ramp and thus be biased on average toward the land. Our evidence seems to show that humans invoke a similar strategy for this type of edge profile, thus accounting for the similar average answers provided by manual and Hueckel edge detection in this experiment.

The proposed edge model, "knowing" that the foot of the ramp is the most reliable indicator of the true edge location, provides an answer that is closer, on average, to the water end of the ramp than the other techniques. The 525-ft. contour line can be observed to project onto the edge ramps at a point that is typically fairly close to the land end of the ramp. This can be partially explained by the fact that the camera calibration used in this test was performed by a human

operator trying to match landmarks in the image with points on the contour map; many of these correspondences were points on the 525-ft. contour, with points on the perceived land/water boundary in the image that we know were biased toward the land. The point here is that even though humans with reasonably normal vision are remarkably consistent in choosing the location of boundary points in an image (in our experiment the two people agreed exactly in their choice of edge-point location on 100 of the normals and differed by one pixel in their remaining 36 choices), they use a general model that may not be in agreement with the underlying physical situation.

To complete our discussion of the experiment from the standpoint of determining the reservoir water level elevation, we note that the model-based edge finder operated well enough to provide us with an average value about as close to the correct answer as we could hope to get from the given data. In a sense, it performed too well, since the global shape matching algorithm really had nothing additional to offer. This situation resulted because we were working with a very "clean" image; the model-based edge finder made no real mistakes (i.e., selected an edge point completely off the land/water intensity ramp), and the Hueckel edge finder "refused" to make an estimate on three of the normals, while making only four mistakes elsewhere (see Figure B-10). In a continuous monitoring situation, the availability of such clean imagery cannot be counted on, and the result of a simple averaging computation will be unreliable. (There is also the additional issue of the edge model ensuring unbiased image plane edge point location estimates, but not unbiased associated elevation values.)

Given that our reference data was in the form of contour lines with 5-ft. elevation differences, the best we could expect to accomplish by selecting the best match between boundary and contour shape was to estimate the water level elevation to within a plus or minus 2.5 ft. error. This could be accomplished by either finding a single best match or by finding two contour lines that appear to bracket the detected land/water boundary. The latter situation occurred in our

experiment with the 520 and 525 contour lines being relatively good matches. The short confidence interval (0.14 pixel) about the detected boundary implies that the actual water elevation lies somewhere between a 520- and 525-ft. interval.

## 6. Conclusions

The main contribution of this effort was the development of a method for the effective use of a-priori knowledge of shape and location to infer highly accurate geometric information from coarse image measurements. The ultimate accuracy achievable by such an approach is a function of the unknown biases in the reference data (i.e., the a-priori contour/shape information) and in the models (the calibration and edge models). Assuming no such unknown biases, then the problem is to efficiently use multiple measurements to reduce random errors in both the models and the measured data and to develop efficient methods for making the needed measurements.

We have developed an effective model-based edge finder and a decision procedure for making a globally optimal decision using coarse local measurements. The effectiveness of this latter technique is enhanced by detecting and eliminating data points that have high probability of being in error and by weighting the valid data points according to their discriminative information content.

The experiment we performed must be considered illustrative, rather than an actual demonstration of the proposed methodology, since we had no reasonable way to verify a key data item (i.e., the precise location of the land/water boundary in the image data).

As a by-product of the experiment, we acquired some data on the biases to be expected when a human observer attempts to locate edges manually in terrain imagery.

FIGURE B-1 AN INTENSITY PROFILE NORMAL TO AN "IDEAL" EDGE IN AN IMAGE

FIGURE B-2 AN INTENSITY PROFILE NORMAL TO AN "IDEAL" EDGE IMAGED BY A SYSTEM WITH LIMITED BANDWIDTH

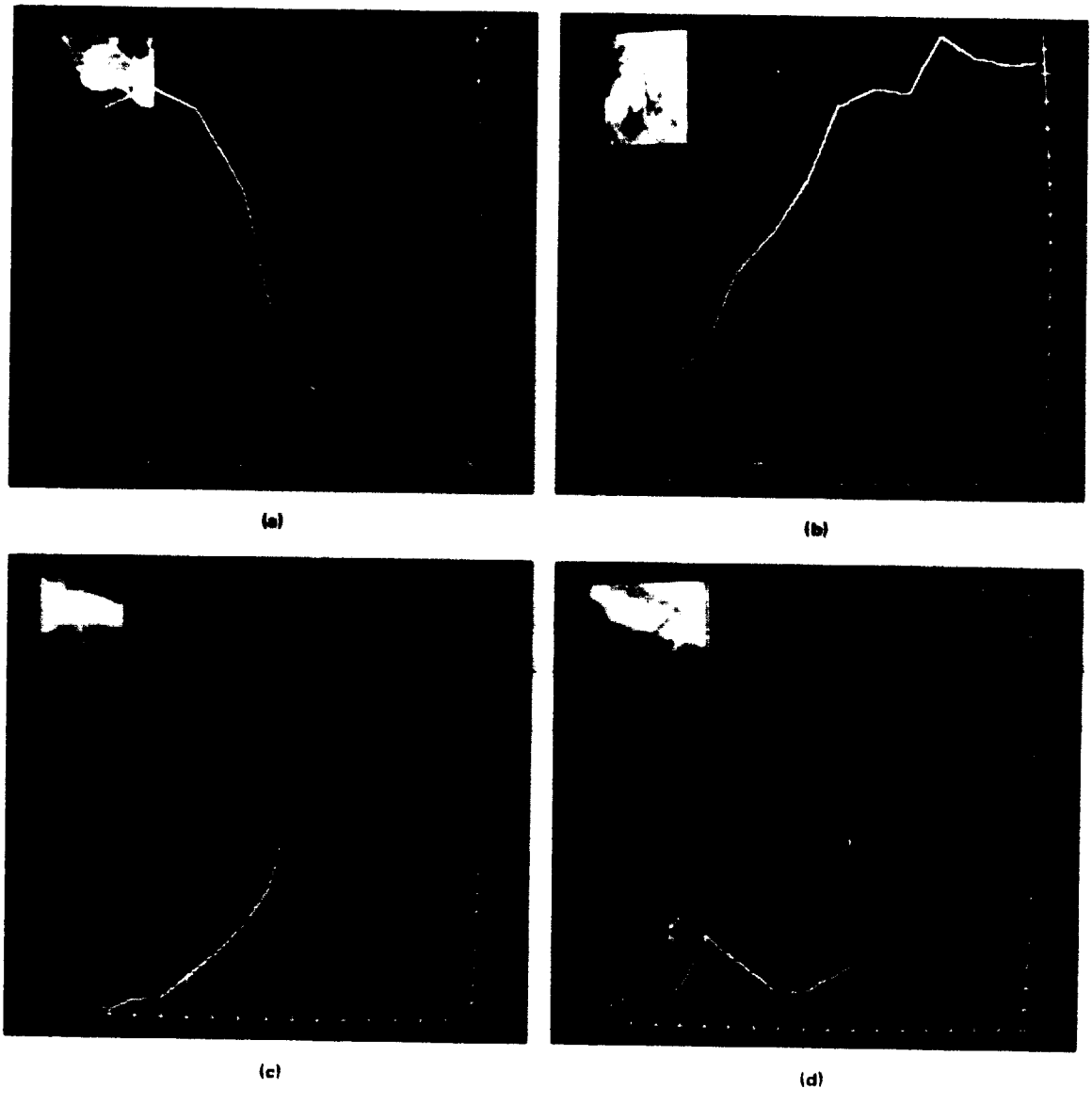



FIGURE B-3 SOME ACTUAL INTENSITY PROFILES OF EDGES AND NONEDGES

Intensity profiles of edge features; bright line = bilinear intensity, grey line = pixel intensity, X-axis = 1 pixel/unit, Y-axis = 10 intensity levels/unit. The two vertical lines show location of contours 520 and 525. (a),(b) nonedges, (c), (d), (e), (f) edges.

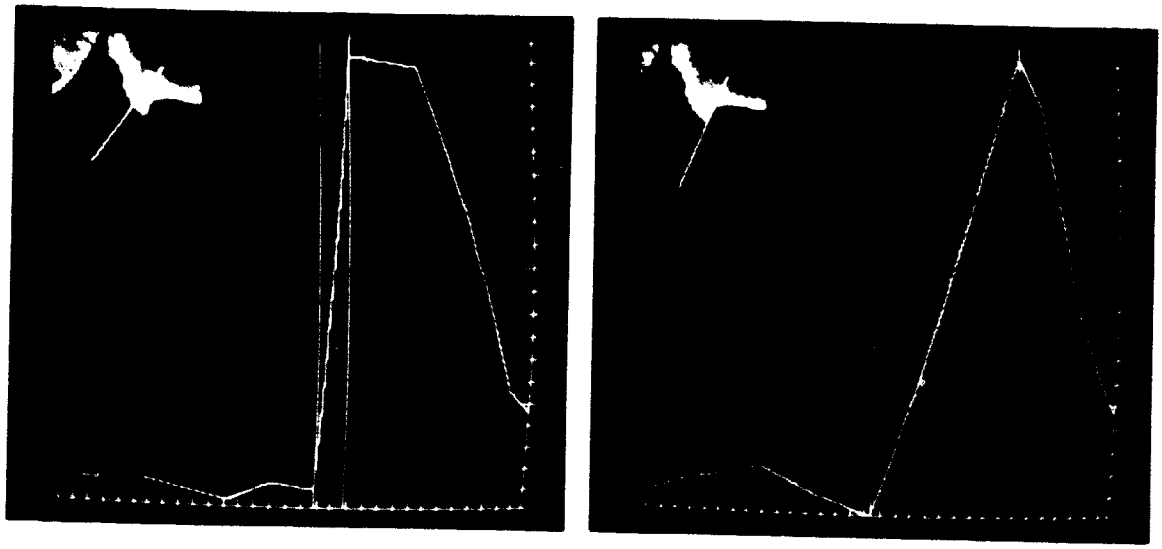



FIGURE B-3 SOME ACTUAL INTENSITY PROFILES OF EDGES AND NONEDGES (Concluded)

Intensity profiles of edge features; bright line = bilinear intensity, grey line = pixel intensity, X-axis = 1 pixel/unit, Y-axis = 10 intensity levels/unit. The two vertical lines show location of contours 520 and 525. (a), (b) nonedges, (c), (d), (e), (f) edges.



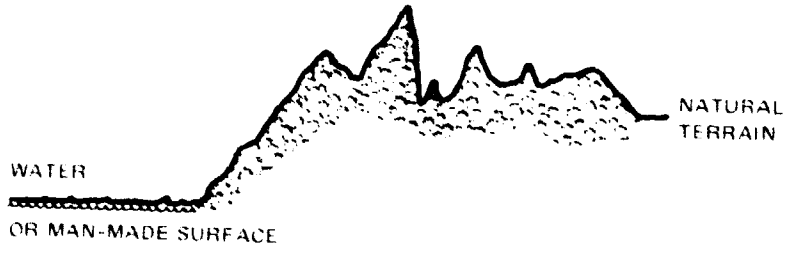



FIGURE B-4 THE INTENSITY PROFILE OF A LAND-WATER INTERFACE OR BOUNDARY BETWEEN A MAN-MADE SURFACE AND SURROUNDING NATURAL TERRAIN

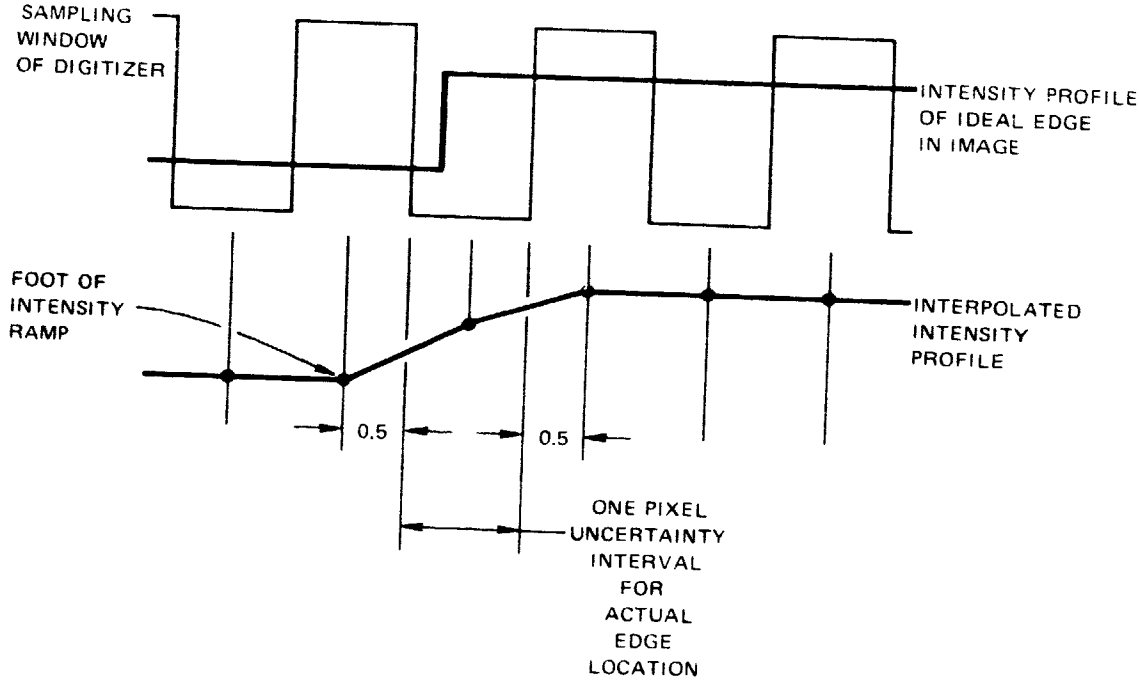
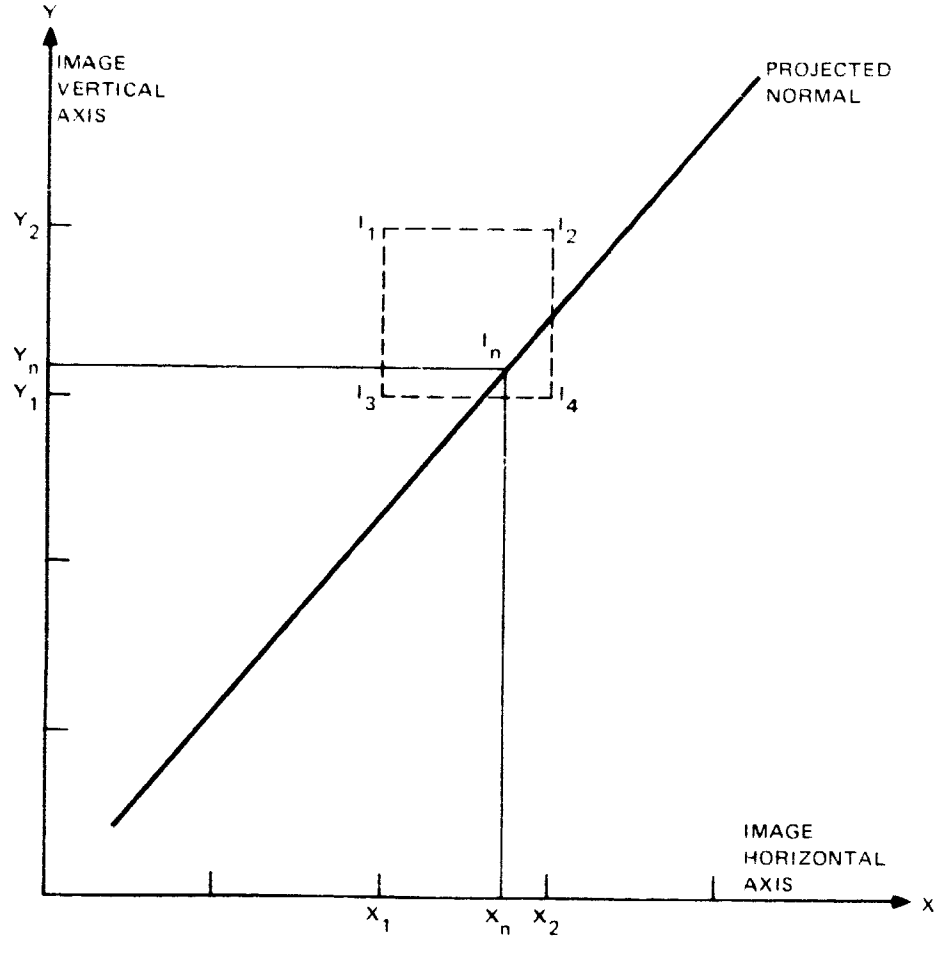
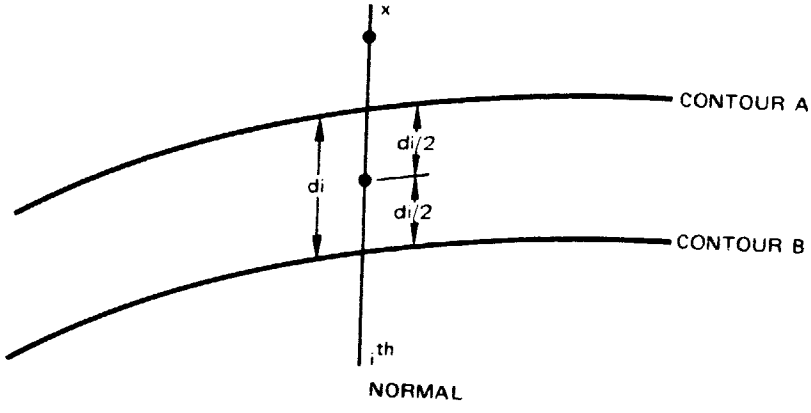



FIGURE B-5 RELATIONSHIP BETWEEN QUANTIZED EDGE INTENSITY PROFILE AND ACTUAL EDGE LOCATION IN THE IMAGE PLANE




$$I_n = \frac{1}{4} \sum_{i=1}^{4} I_i \; |X_i - X_n| \; |Y_i - Y_n|$$

FIGURE B-6 BILINEAR INTERPOLATION OF IMAGE INTENSITY VALUES ONTO A PROJECTED NORMAL




$$\text{Likelihood Ratio} = L(x) = \frac{P(x|A)}{P(x|B)}$$

Accept Hypothesis A if $L(x) \geq 1$
Accept Hypothesis B if $L(x) < 1$

$$P(x|A) = \prod_i P(x_i|A) \qquad P(x|B) = \prod_i P(x_i|B)$$

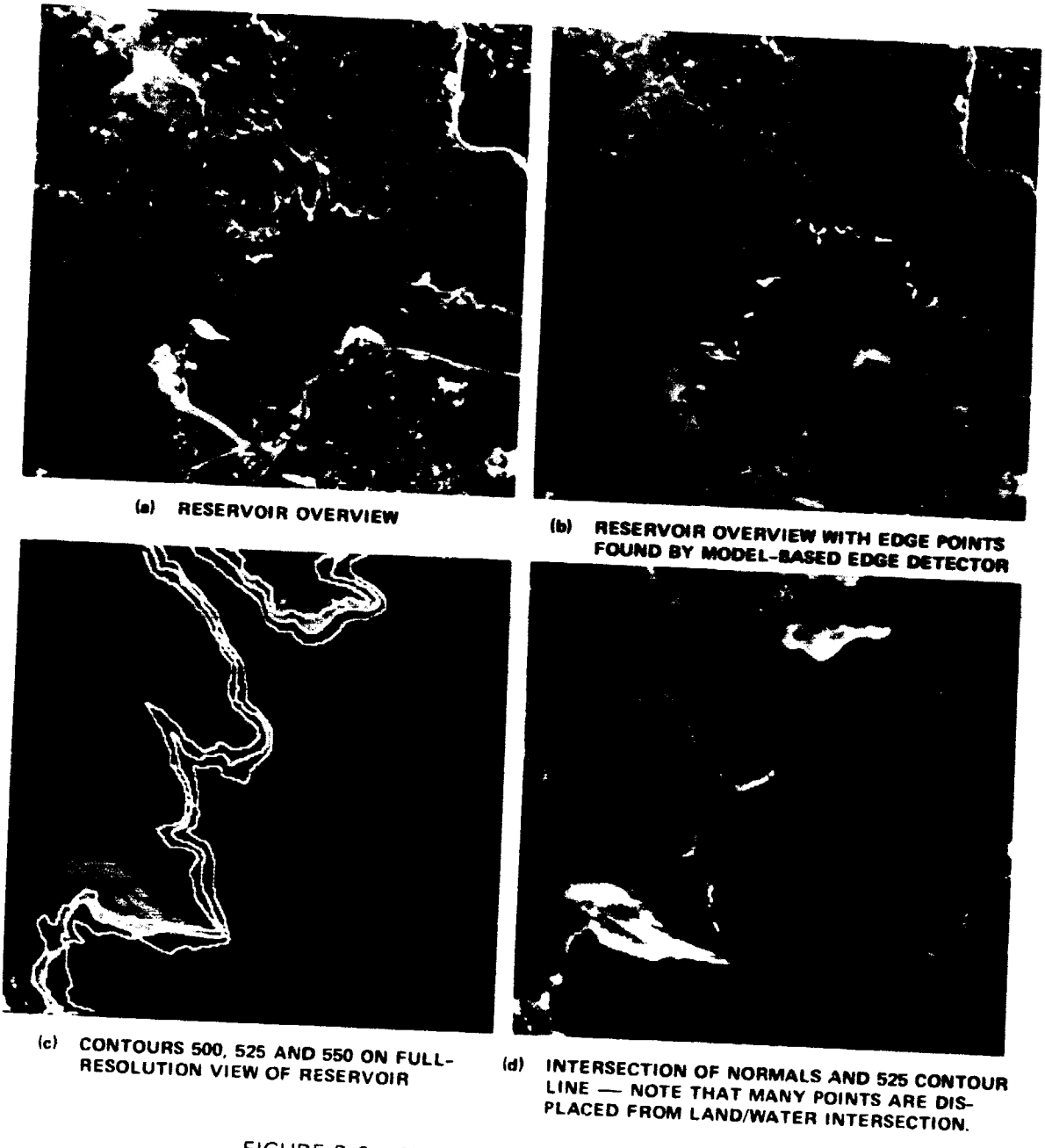
$$P(x_i|A) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left[-\frac{\left(x_i - \frac{d_i}{2}\right)^2}{2\sigma_i^2}\right]$$

$$P(x_i|B) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left[-\frac{\left(x_i + \frac{d_i}{2}\right)^2}{2\sigma_i^2}\right]$$

} assuming $x_i$ supports Hypothesis A

$$\text{Log } L(x) = \sum x_i \frac{d_i}{\sigma_i^2} \quad \text{(approximately)} \qquad \text{where } x_i = \begin{cases} 1 \text{ for } i \to A \\ -1 \text{ for } i \to B \end{cases}$$

FIGURE B-7 A DECISION PROCEDURE FOR GLOBAL SHAPE MATCHING



(a) RESERVOIR OVERVIEW

(b) RESERVOIR OVERVIEW WITH EDGE POINTS FOUND BY MODEL-BASED EDGE DETECTOR

(c) CONTOURS 500, 525 AND 550 ON FULL-RESOLUTION VIEW OF RESERVOIR

(d) INTERSECTION OF NORMALS AND 525 CONTOUR LINE — NOTE THAT MANY POINTS ARE DISPLACED FROM LAND/WATER INTERSECTION.

FIGURE 8-8 IMAGES OF BRIONES RESERVOIR

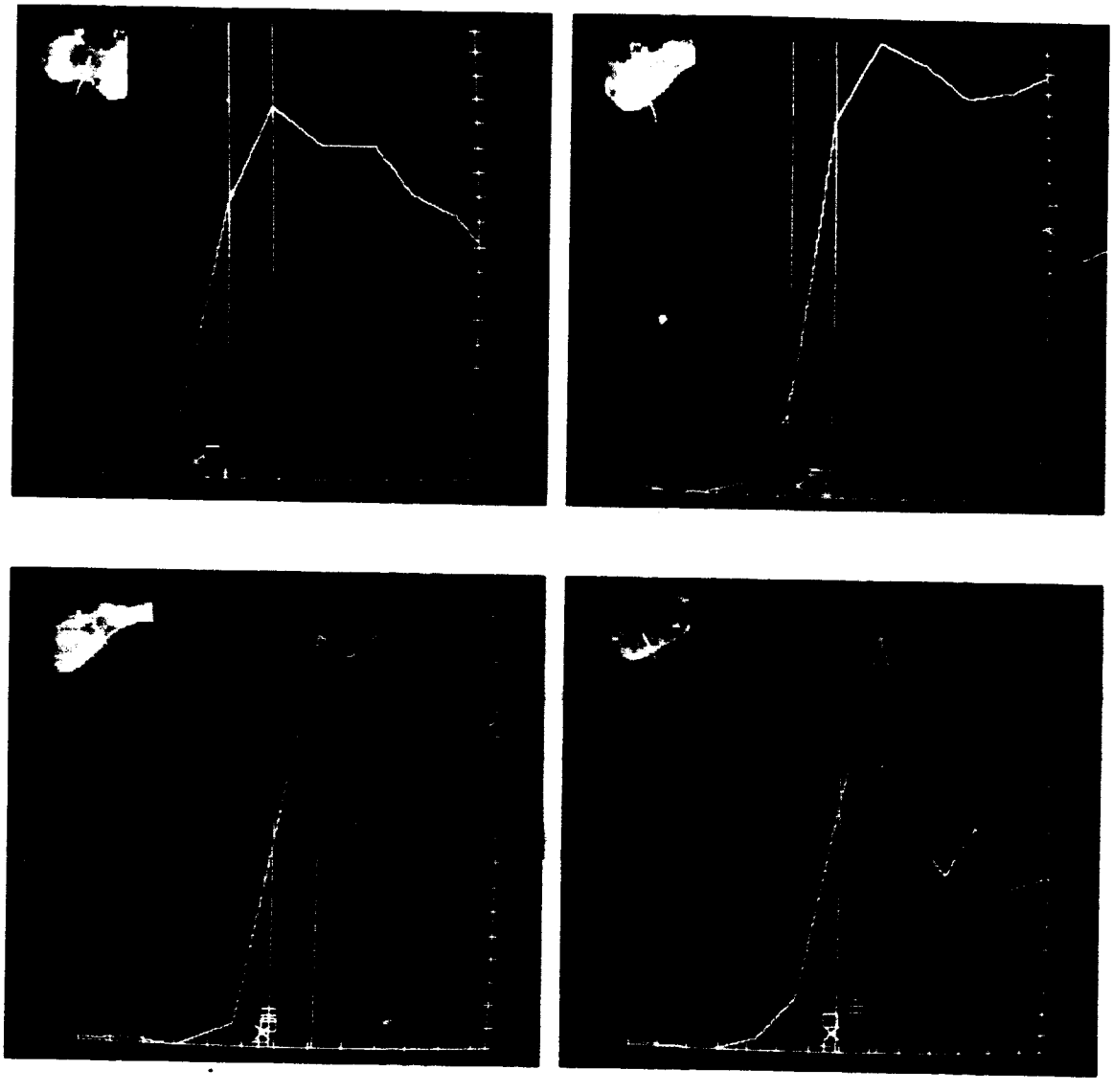


FIGURE B-9 COMPARISON OF EDGE POINTS FOUND BY AUTOMATIC AND MANUAL TECHNIQUES

Intensity profiles of edge features; bright line = bilinear intensity, X-axis = 1 pixel/unit, Y-axis = 10 intensity levels/unit. The two bright vertical lines show location contours 520 and 525. + = model-based technique, X = Hueckel edge detector, □ = person 1, ▽ = person 2

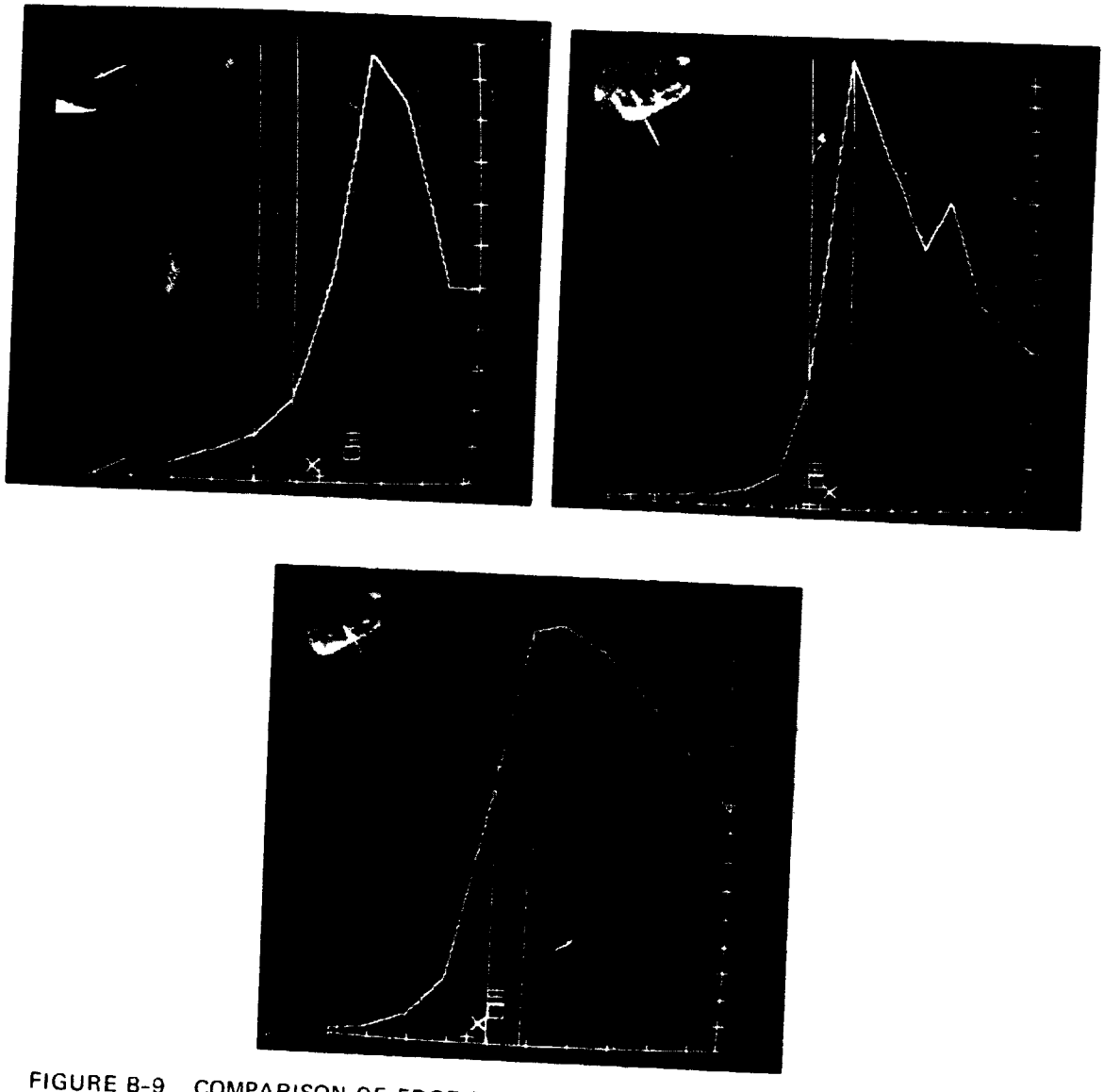


FIGURE B-9 COMPARISON OF EDGE POINTS FOUND BY AUTOMATIC AND MANUAL TECHNIQUES (Concluded)

Intensity profiles of edge features; bright line = bilinear intensity, X-axis = 1 pixel/unit, Y-axis = 10 intensity levels/unit. The two bright vertical lines show location contours 520 and 525. + = model-based technique, X = Hueckel edge detector, □ = person 1, = person 2

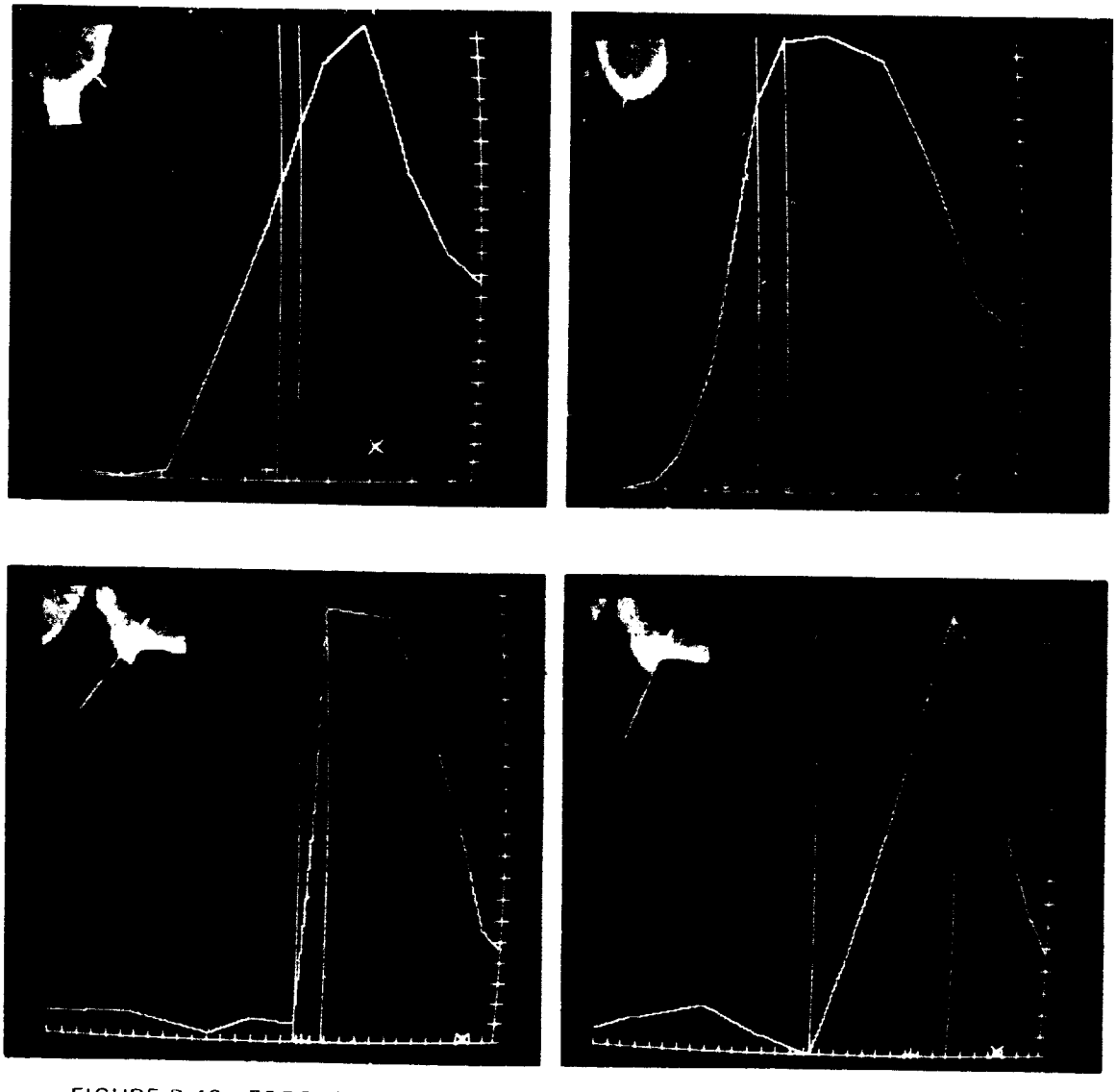


FIGURE B-10 ERRORS MADE BY HUECKEL EDGE-FINDER

Bright line = bilinear intensity, X-axis = 1 pixel/unit, Y-axis = 10 intensity levels/unit. The two bright vertical lines show contours 500 and 525. + = model-based technique, X = Hueckel edge detector

END

DATE

FILMED

FEB 12 1979