Technical Note 137

January 1977

INTERACTIVE AIDS FOR CARTOGRAPHY AND PHOTO INTERPRETATION

By: Harry G. Barrow, Thomas D. Garvey, Jan Kremers, J. Martin Tenenbaum, and Helen C. Wolf

Artificial Intelligence Center

SRI Projects 3805 and 4763

The research reported herein was supported at SRI by the Advanced Research Projects Agency of the Department of Defense, under Contract No. DAHC04-75-C-0005 and DAAG29-76-C-0012.
INTERACTIVE AIDS FOR CARTOGRAPHY AND PHOTO INTERPRETATION
by
Harry G. Barrow, Thomas D. Garvey, Jan Kremers,
J. Martin Tenenbaum, and Helen C. Wolf

A. Introduction

1. Overview

This report covers the six-month period October 1975 to April 1976. In this period, the application areas of ARPA-supported Machine Vision work at SRI were changed to Cartography and Photointerpretation. This change entailed general familiarization with the new domains, exploration of their current practices and uses, and determination of outstanding problems. In addition, some preliminary tool-building and experimentation have been performed with a view to determining feasibility of various AI approaches to the identified problems. The work of this period resulted in the production and submission to ARPA of a proposal for research into Interactive Aids for Cartography and Photointerpretation. This report will not reiterate in detail the content of the proposal, but will refer the reader to it for further information [1].

2. Selection of Domain

There were three essential criteria in selecting a domain for Image Understanding research; it must be of importance to the Department of Defense, it must present central issues of fundamental scientific importance for study, and it must present problems that are tractable.

Information in the form of aerial photographs is of prime military importance, for both strategic and tactical purposes. Many
thousands of photographs are taken each day and used for making maps, and for obtaining intelligence information. Current cartographic and photointerpretive techniques require substantial use of data processing, but computers are not yet used directly for processing the photographic data. There are still labor-intensive bottlenecks in the conversion of raw pictorial information to a symbolic form. In map-making, for example, cartographic features such as roads or lakes are traced by hand, while in photo interpretation, even basic tasks like counting or measuring are performed manually. It is therefore evident that there is considerable scope for automation and consequent increase in quality and quantity of throughput. So far, however, these tasks have resisted automation, primarily because a task that appears simple, like road tracing, involves application of considerable amounts of knowledge. Tricky cases that would defeat an elementary approach occur very often. The approaches of Artificial Intelligence, and Machine Vision in particular, are aimed directly at the embodiment of diverse sources of knowledge in computer programs which are then able to display useful expertise in selected tasks. The application of AI to cartography and photointerpretation is therefore natural and appropriate, and promises to be fruitful in the foreseeable future. We have, in consequence, selected the twin domains of Cartography and Photointerpretation for our next program of research in Machine Vision.

At first sight, Cartography and Photointerpretation may appear to be totally different fields. Cartography is concerned with locating the permanent features of an area with great accuracy, while Photointerpretation is concerned with dynamic features and events. In fact, they are simply extremes of a continuum, and overlap considerably. In order these two fields map features, one must first interpret the photograph, and in interpreting a photograph, one implicitly constructs a map. We therefore deem it inappropriate to draw an artificial distinction, and believe that there is much to be gained, in terms of both expediency and scientific interest, by taking a unified view of Cartography and Photointerpretation.
3. Scientific Objectives

The crucial issue in Image Understanding is the role of knowledge. Many important questions are still very much open. What domain-independent knowledge should be built into the lower levels of a visual system? What domain-specific knowledge is necessary for achieving specific goals? How should that knowledge be deployed, invoked, and applied? How can many diverse sources of knowledge be coordinated? The only irrefutable fact is that knowledge, both general and specific, is essential to image understanding.

The work that we have embarked upon in the present period includes an in-depth study of a particularly potent embodiment of knowledge which is of both theoretical and practical importance, namely a map. A map, which can be viewed as a data structure that preserves three-dimensional geometrical and topological properties, can be of very great assistance in interpreting a photograph.

The geometrical relationships between maps, photographs, and the real world are well understood, and it is possible to establish an exact correspondence between points in the map and points in an image. Thus, the map can be used to predict which features should appear where in the photo, thereby eliminating much general-purpose searching and processing. It appears that using the map to guide the processing and interpretation of an image may yield considerable gains in efficiency and robustness, as well as providing a coherent, modular structure for the whole system.

Within this framework, a variety of different types and levels of knowledge may be unified. To give a specific example, advice provided interactively by the user in the form of approximate manual tracings of a road may be regarded as additional information which may be stored and manipulated in map form, and which may be used along with general knowledge of the characteristic behavior of roads in locating the road accurately in the image.
The scientific objectives of the current work are thus the elucidation of what information should be stored in the map database, how it should be represented, and how it may be employed flexibly in image understanding.

4. Application Objectives

In the domains of Cartography and Photointerpretation, we are focussing our attention on those bottleneck areas upon which we expect to be able to make significant impact within a reasonable time scale. The cartographic tracing problems are currently under study, since these also underlie many photointerpretation tasks. Gradually, the emphasis will shift toward photointerpretation.

We recognize that it is not possible within the current state of the art of machine vision to fully replace human abilities. We therefore adopt as one important objective the development of a system that can accept advice supplied interactively by a human user and collaborate with him in an image understanding task.

In summary, we aim towards the development of a collaborative aid for a cartographer or photointerpreter that can employ information provided in the form of a map, or supplied interactively by the user.

5. Summary of Work Carried Out

The work of the current six-month period has been a broad exploration to determine techniques and approaches currently in use in production cartography and photointerpretation. The work has included the design of and experimentation with a basic integrated system.

The main accomplishments of the six-month period are as follows:

* We have equipped ourselves with a range of documents covering civil and military mapping and photointerpretation. We now have a reference library that should prove valuable in our current and future research.
* We have established contact and visited a number of
centers at which civil and military map-making and photointerpretation are performed, and research centers at which approaches and techniques for the future are being developed.

* We have determined bottleneck tasks in Cartography and Photointerpretation as currently practiced. We have identified the common requirements, and examined their susceptibilities to Machine Vision techniques.

* We have acquired a realistic imagery data base, representative of the important tasks and central issues.

* We are currently in the process of designing and implementing a basic interactive cartography and photointerpretation system for supporting our research work.

* We are already using the basic system in some initial experiments in map-guided tracing.

* We have written and submitted to ARPA a proposal for further research on Interactive Aids for Cartography and Photointerpretation.

The proposal contains a detailed program of research, with a description of a proposed interactive system. Therefore, we shall not dwell further on future research, but confine ourselves in the rest of this report to work completed in the present six-month period.

The next section reports in detail on the studies we have so far made into the current approaches to cartography and photointerpretation. We shall identify certain common requirements, their susceptibilities to AI techniques, and the key obstacles that must be overcome.

The succeeding section describes the experimental work performed so far, including an examination of the problem structure, acquisition of suitable experimental data, the design of an experimental system, and some initial experiments in interactive map-making.

The final section gives general conclusions concerning the research and applications upon which we have embarked, including observations relating to the guidance of the proposed future work.
B. Survey of Applications and Requirements

This section summarizes the current status of Cartography and Photointerpretation, highlighting the bottleneck areas on which our project is now focused. Our information was gathered from the reference material listed in Appendix I and from visits to the facilities listed in Appendix II.

1. Cartography
   a. Present Approaches

Despite a large body of mechanical techniques, the production of maps from aerial photos is still primarily labor-intensive. The most automated systems now in existence are the SACARTS (Semi-Automatic CARTographic System) developed for DNA-TC by USAETL and the LIS (Lineal Input System) developed for DNA-AC by RADC. A schematic of the SACARTS system is shown in Figure 1.

SACARTS includes facilities for automating all stages of map production, including compilation, editing, and reproduction. Our discussions will concentrate on the compilation and editing phases, where AI techniques can contribute most directly. Compilation consists of the acquisition in digital form of the cartographic data to be mapped. The two major types of data are elevation profiles and boundary coordinates of planimetry (lineal features, such as roads, rivers, and coastlines). Editing is concerned with verification of the internal consistency and accuracy of the information, and its correction where necessary.

The first step in compilation is performed by an automated stereo plotter, known as UNAMACE (UNiversal Automatic Map Compilation Equipment). UNAMACE takes an overlapping pair of unrectified aerial photographs and produces both an orthophoto (a rectified photograph in which all terrain points are depicted in their correct map positions) and a set of elevation profiles.
FIGURE 1  SCHEMATIC OF SACARTS OPERATION
The elevation data are smoothed by software and transformed into contours, for subsequent merging with planimetry extracted from the orthophoto.

The extraction of planimetry has always been the most labor-intensive and time-consuming step in map compilation. In SACARTS, each class of planimetry is manually traced on a separate sheet of frosted acetate that is overlaid on the orthophoto. (Planimetry can also be compiled by tracing features on a pre-existing map.) The acetate sheets are then transferred to a digitizing table where the features are retraced to produce a digital tape containing their digital coordinates. This operation concludes the compilation phase.

The digital elevation contours and planimetric data on magnetic tape are now transferred to a Univac 1108 where they are manipulated by a series of editing programs, known collectively as GISTS (Graphic Improvement Software Transformation System).

The data are first output on a high speed plotter for manual verification. Boundary and contour imperfections are currently corrected off line by retracing on the digitizing table and creating an update tape. However, an on line graphics editing system (DIODE) will soon be available.

From this point, processing becomes more automatic. Linear features of definite width, such as roads, are thinned down to a center line and smoothed. Precise intersections of features on a single overlay are computed. Selected features are checked for positional correctness against independently entered ground control coordinates. Features on different overlays are then checked for consistency. For example, the system checks whether any buildings are positioned on top of roads. Inconsistencies are resolved automatically where possible, and manually otherwise. The system is currently capable of resolving a road/building conflict by displacing the building and orienting it in the direction of the road. The system will soon be able to check for other types of local consistency, such as whether roads follow contours.
and streams cross contours, and whether bridges are indicated at road/river intersections.

b. Limitations and Needs

SACARTS, though still not fully operational, promises significant savings in time and manpower over conventional map making technology, plus a reduction in the opportunities for human error. These advantages are realized primarily in the automation of contour extraction and in the manipulation of extracted data into a form suitable for a finished map. The system still leans heavily on human inputs for the extraction of planimetry, for the extraction of contours in difficult terrain, and for error correction during editing. The following sections discuss these bottlenecks, the currently contemplated solutions, and the possible contributions of AI research.

Extraction of Planimetry

The major remaining bottleneck is the tedious manual tracing that occurs in both the extraction and the subsequent digitization of planimetry. The amount of labor required in these steps, both in hours and in dollars, is phenomenal. Typically, it takes about 75 hours to trace all the overlays (roads, streams, drainages, buildings, and the like) for a small scale (7.5′) topographic map of a rural area. In a heavily urbanized area, it can take as much as 500 hours. This tracing is performed by GS-9 level personnel, whose annual salary is about $15,000. Almost 400 of these people are employed at DMA-TC alone, and constitute at least a quarter of all personnel involved in map production generally. Tracing operations are by far the most time-consuming steps in map compilation and make it uneconomical to keep maps up to date. Clearly, automating the extraction of planimetry is a prime application for image understanding.

Several projects are currently underway at ETL, aimed at alleviating this bottleneck. The Autocartography group, under Howard Carr, which originally developed SACARTS, is now trying to eliminate the
manual retracing required for digitization. An acetate sheet containing pencilled tracings of a single feature (such as roads) is placed on a high resolution drum scanner and digitized in a raster format. The digitized data are thresholded to extract the lines. Each line is then thinned down to a center line and smoothed to eliminate gaps. Finally, interconnections of the lines are computed to recover the road network. Some of these features are found in the competing LIS system developed at RADC.

Preliminary tests verify the feasibility of this approach but at the same time, have underlined the practical issues that arise in processing high resolution cartographic data. Processing times of about one and a half hours per overlay on a CDC 6400 have prompted examination of more appropriate computer architecture, such as the Goodyear STARAN. Speedups on the order of 25 times have already been observed for selected operations.

The Computer Sciences group, under Larry Cambino, is setting up an interactive image processing facility that will be used to study the much harder problem of extracting planimetry directly from photographs. The group is also investigating a variety of image enhancement techniques that should facilitate extraction by both man and machine.

The Technology Development branch, under Bernie Scheps, has developed an interactive image processing system based primarily on analog techniques. Density slices from up to four bands of imagery can be logically combined to produce a binary overlay. Although density slicing is a very limited form of feature extraction, with multi-spectral imagery, it is frequently possible to find thresholds that will extract most instances of a given feature, say roads, in a particular image. The real time nature of analog processing facilitates the empirical selection of suitable thresholds; the operator turns potentiometers while observing the displayed overlay, and stops when the best one is achieved. The resulting overlay could be digitally
processed in a manner similar to raster scanned pencil tracings to obtain an approximate road network.

None of the research to date has attacked the problem of tracing planimetry in black and white aerial imagery. This process, in our opinion, is too difficult to automate completely at this time. We are therefore advocating an interactive approach wherein features traced crudely at free hand speeds are used to guide the automatic extraction of precise boundaries. In updating old maps, the map itself is used to guide the tracing of pre-existing features. The man can interactively refine the machine's output if necessary, eliminating the subsequent need for an explicit editing step.

The same techniques can also be employed in updating old maps from new imagery. Here, the old map itself is used to guide the tracing of pre-existing features.

Elevation Contour Extraction

The UNAMACE, and all other automatic stereoplotters produced to date, use local area correlation to extract disparity information from a stereo pair of imagery. Thus they do not function reliably in steep terrain, where appearances can vary significantly with slight changes of viewpoint, or in flat featureless terrain, such as open water. They are also unable to recognize and adjust to planimetry. Another problem is the strict demands made on the quality of photography used in automatic stereocompilation. Currently, about 40% of the images are rejected because the equipment will not correlate continuously. Some of these images can be handled after reprocessing in the photo lab. Because of all these difficulties, a human's judgement and global viewpoint are needed to correct errors and to rescue the machine when it gets lost.

It should be possible to incorporate some of the needed judgement in a smart correlator based on AI techniques. For example, the knowledge that contours must close and cannot cross can be used to
constrain image correspondence in featureless areas. It should also be possible to combine the extraction of contours and planimetry so that the depth disparity can be used as a feature in extracting planimetry, and the resulting planimetry can constrain the placement of contours.

Editing

Improving the sophistication of automatic map evaluation and error correction is the subject of ongoing research at ETL. So far, attention has focused on the detection of local inconsistencies in the map. The automatic correction of such errors is, in general, a much harder task, which can require knowledge of the overall landform, road networks, and the like. Additional knowledge about the relations of roads, rivers, contours, overpasses, and other features, and about their corresponding appearance in images, is needed to ensure that a map conforms to all the standards set forth in [2]. As the level and the global scale required for error detection and correction increase, the need for AI techniques becomes more apparent.

Map Using

Cartographic data bases of the type produced by systems like SACARTS can be used for many purposes other than just for printing new topographic maps. ETL is currently investigating a variety of such applications. They include the generation of special-purpose thematic maps, the efficient updating of previously compiled maps, and a variety of military geographic intelligence applications, such as computing the cross-country travel times of a vehicle or the construction time for an airfield. The latter capabilities could be combined with AI problem solving techniques to augment manual decision making in many tasks. Some examples are: finding a best route between two destinations, subject to constraints; determining the best location for an airfield; balancing such factors as tactical surprise against construction time; and determining regions of critical terrain.
2. Photointerpretation
   a. Present Approaches

   To learn about photointerpretation as currently practiced, we visited the 544th reconnaissance wing of the Strategic Air Command at Omaha. There we received an unclassified briefing on PACER, an operational computerized information handling facility for photointerpretation and analysis.

   The PACER system, shown schematically in Figure 2, consists of a number of photointerpretation stations and analysis stations, all on line to two Honeywell 6080 computers. The computers maintain a multi-source data base of intelligence information that serves as the principal means of communication between the photointerpreters and the analysts. The data base contains known target descriptions and related intelligence, expressed by both formal descriptors and free text elaboration. The data base also contains notes from analysts requesting further interpretation or monitoring of selected targets. There is a limited amount of on-line cartography, but no on-line imagery.

   The analyst stations are conventional CRT text terminals. Analysts enter their photointerpretation requests which are communicated through the data base to the appropriate PI stations assigned to the geographic areas in question. Interpretations made by the PI are then entered in the data base where subsequently the analyst can examine them on line.

   Each PI station has three components: a light table, a film chip viewer, and a Bunker-Ramo BR-90 display. The light table is equipped with two film transports and a conventional zoom microstereoscope. The chip viewer is augmented by a mechanical storage and retrieval facility. The display is a graphics console with light gun. Slides can also be rear projected on the display surface.
FIGURE 2 SCHEMATIC OF THE PACER SYSTEM
A pair of PIs are assigned to each station. One views the imagery on the light table, while the other interrogates the database and retrieves chips as needed. Both are specialists in a particular geographic area and they switch roles halfway through their shift.

PIs work with PACER in two distinct modes, a priority mode, concerned with monitoring high interest areas in support of a particular mission, and a more systematic mode, concerned with detecting, identifying, locating, and reporting in detail on installations and activities whose existence was not previously known. In priority mode, the targets selected for monitoring are flagged on the display with respect to a rear projected map. This location is verbally relayed to the interpreter at the light table along with the analyst’s instructions on what to look for. This active interpreter must search through rolls of film to find frames containing the designated target. He then responds to the analyst’s request (often with a terse phrase such as "no change"). The response is entered into PACER by the other interpreter at the console who is, in effect, serving as a natural language interface. In the second mode, the interpreter at the light table searches systematically for new areas of interest in each frame. When a target is encountered, the interpreter verbally describes some distinguishing features with which his partner can determine whether an up-to-date description already exists in PACER. Target location can be designated on the map, using the light gun, and the computer will transform the display coordinates to obtain the actual map coordinates. Additional features may be needed to distinguish among similar targets clustered in an area. If PACER has such a target, it will output a description which can then be checked against the image for verification. The resulting dialog leads to a new or updated target description.
b. Limitations and Needs

PACER has gone a long way toward facilitating the timely exchange of intelligence information among PIs and analysts, and it is well regarded as having improved the overall effectiveness of SAC's reconnaissance operations. Nevertheless, its performance is still limited by the fact that all actual operations on imagery are performed off line and completely manually.

It is impractical to automate completely human perceptual and decision making skills. However, there appear to be many interactive aids based on Artificial Intelligence technology that could improve the speed, cost, and reliability of existing PACER capabilities. We can describe briefly some obvious opportunities for aiding image archiving, image enhancement, and graphical communication, as well as pure image understanding.

Archiving

Photointerpretation requires a substantial amount of comparison with previous imagery on film chips. There is thus the need for cataloging and storing images in a way that permits efficient retrieval of those relevant to some area or operation. The relevance of an image is not simply defined by the area it covers, but may involve the interpretation of what it depicts. For example, if one desires to determine where a ship unloaded her deck cargo, it is necessary to find out whether there is any coverage of the actual unloading, and failing that, to retrieve all recent photos of the ship at sea to determine when the cargo disappeared.

Images should thus be accessible by content as well as location and time, and if possible, should be available for on line viewing and processing. The current ARPA-sponsored work on very large memories and data bases may find application here.

Enhancement
A great variety of enhancement techniques have been experimentally investigated, ranging from false color displays to mathematically sophisticated filtering. Appropriately used, these techniques can sometimes trivialize the detection of targets that were indistinct in the raw image. In practice, enhancement techniques are seldom used, except for special studies of high priority targets. One reason, according to a CIA source, is that the average photointerpreter lacks the background needed for selecting appropriate enhancements for particular tasks. This shortcoming suggests the need for a knowledge-based system to serve as an enhancement expert. Given a description (or pictorial examples) of the desired target and the expected background, such a system would select the appropriate enhancement. With on-line imagery and electronic viewing, the enhancement can be optimized for each area of the image.

Graphical Communication

The person manning the PACER terminal is serving essentially as a natural language interface between the person doing the actual photointerpretation and the computerized data base. This interface is both wasteful of skilled manpower and awkward because it forces the verbal communication of concepts that are primarily visual. Consider as an alternative, a single user system with both on-line cartography and imagery. Image and map can be superimposed optically or electronically and brought into correspondence by manually designating a few corresponding points in each. The system could then task the PI by encircling the target to be monitored directly in the image. New target detection could be simplified by encircling all previously known ones. New finds could then be reported by pointing at them with the light pen and having the computer extract directly the geographic coordinates and possibly other descriptive parameters (see below). The PI might also enter a verbal description in a restricted vocabulary, using one of the commercially available speech input devices (such as Threshold Technology's VIP).
Image Understanding

Aids for image processing tasks such as counting, measuring, and change detection offer potentially the highest payoff, but also involve the most technological risk. Image processing research, supported primarily by the military, has been underway for almost 20 years. Almost universally, the results have been lacking either generality or reliability. The reason, we believe, is because the unreliable techniques did not use adequate knowledge of the domain and the nongeneral techniques were inflexible in the use of what knowledge they had. ARPA's Image Understanding program was indeed founded on the belief that knowledge, appropriately used, could dramatically simplify many image processing tasks of practical importance.

The PACER data base already contains a substantial amount of information about the location and appearance of targets. This information provides powerful constraints on where to look in the imagery and what to look for. Here are examples of how such knowledge might be exploited in two classes of PI tasks.

Change Monitoring

The conventional approach to automatic change detection entails the warping of intensity normalized images into geometric correspondence, and the subsequent subtraction and thresholding of corresponding pixel intensities [3]. Such an approach is both computationally expensive and limited in effectiveness to images taken with the same sensor from approximately the same viewpoint, at the same season and time of day. The approach also is incapable of distinguishing between change that is significant and change that is insignificant in a military context. To paraphrase one contact "Change detection is easy in vertical photographs taken at noon from the same position and orientation. The challenge comes in dealing with low angle obliques taken from different positions with low sun angles."
The general change detection problem is admittedly beyond the state of the art. However, automating the more constrained problem of change monitoring, that is, checking a particular location for changes of a particular type, may be feasible. We have concluded that the use of knowledge-based search techniques [4] are appropriate to locate precisely the specific targets tasked for monitoring in each frame of imagery. A search for significant change can then proceed in a verification mode, guided by the previous target description and independent knowledge of how changes in viewing conditions transform appearances. Indeed, simple techniques like subtraction can often be used quite effectively in suitably constrained contexts. For example, it should be very easy to determine reliably whether a given ship is still in port, once the pier at which it was previously docked has been located.

Counting and Measuring

Counting tasks, such as determining the number of box cars in a railyard or oil wells in an oil field, and measuring tasks, such as determining exact runway lengths and orientations or the capacity of a reservoir, are among the more tedious and time-consuming of a photointerpreter’s duties. Conventional mechanical aids now used in practice include sampling grids and image marking devices for counting, rulers, proportional dividers, planimeters, and the like for measuring. As a consequence, the counting and measuring required in reporting, for example, a new airfield and its associated equipment can take up to 3 hours.

Some work is underway at RADC and similar establishments to alleviate this condition. A developmental system, known as Compass Preview, intended as a successor to PACER, features several interactive measurement aids. For example, a section of an image can be moved across a cursor to compute true ground lengths independent of the current optical magnification. At a research level, a variety of analog and digital electronic systems have been built that detect target areas
(by density slicing, template matching, or feature extraction and classification) and then perform simple mensuration, such as counting the number of distinct connected components, or determining their areas and perimeters. Such techniques have proved quite effective in remote sensing applications, such as estimating crop acreage, especially when the classification criterion or slicing threshold is chosen interactively for particular images. However, military targets are not usually as homogeneous as crops, and it is unlikely that any single classification criterion or threshold will suffice to extract all targets (or all of a single, extended target).

We have several possibilities in mind for applying image understanding research to problems of counting and measuring. As with change monitoring, knowing where to look can substantially simplify the processing needed to discriminate relevant targets. For example, by first delimiting the runway area (either interactively or automatically with the aid of a map), it may indeed be possible to extract airplanes by some relatively trivial operation such as thresholding. Similarly, if the exact location of the rail lines in the image is known, box car counting becomes a one-dimensional template matching problem.

Having delineated an area of interest, the PI might be able to use interactive scene analysis techniques like those developed at SRI [5] to rapidly develop a target finding strategy that takes maximal advantage of that specific context.

3. Research Requirements

Cartography and Photointerpretation have a sufficient number of techniques and requirements in common to suggest a unified technical approach for their automation. Three components appear fundamental:

(1) A database for storing map information.

(2) Techniques for using the knowledge contained in a map to guide image analysis.

(3) Techniques for establishing correspondence between an image and a map.
Maps play primary roles in photo interpretation and cartography, both as formats for output and as sources of knowledge for guiding image analysis. Maps will be represented by a data structure that is indexed by geography and content. This data base will contain dynamic intelligence information in addition to static cartographic and cultural features. It will also contain auxiliary knowledge needed for image understanding, such as the pictorial attributes of objects.

Techniques will be developed for using map knowledge to guide image analysis. For cartography, it is necessary to trace linear features, such as roads, rivers, railways, and coastlines, using as a guide an approximate trace that has been manually entered. Similar techniques can be used for verifying the existence of features whose presence is indicated in a pre-existing map, for such purposes as map updating or change monitoring.

The techniques for map guided image analysis presume a reliable means of establishing geometric correspondence between map and image coordinates. A general transformation can be determined, given sufficient pairs of corresponding points in the image and map. Initially, these pairs will be designated interactively. Eventually, they will be found automatically by using crude map/image correspondence based on navigational data to constrain the search for known landmarks.

The three components outlined above will facilitate development of special-purpose counting, mensuration, and change monitoring aids for photointerpreters, which rely on a map/image correspondence to constrain where to look and what to look for. The aids capabilities are described at length in the proposal.

C. Experimental Studies

1. Domain

It was decided that the task domain for the initial experimental studies should be cartography, because the ability to construct and augment maps seems an essential requirement for advanced
photointerpretation. In addition, we already had some experience with applying AI techniques to cartographic problems [6].

The approach adopted was therefore to construct a simplified system for interactive map-making and updating, which may later also support map-guided interpretation.

2. Problem Structure

The paradigm task addressed runs as follows:

A new image of an area of interest is digitized and displayed. The user indicates a few features of known location by pointing at them and giving their coordinates. The system then computes the coordinate transform between picture and world, enabling it to display a map of the area (if one already exists) accurately superimposed on the photo. The user can make additions or modifications to the map by tracing on the photograph: the system can use crude manual tracing or an existing map to locate and trace accurately features in the picture, and hence automatically update the map.

This paradigm requires all the three fundamental components described above, a digital map, a way of establishing map/image correspondence, and a way of using the map to guide image analysis. It therefore provides a good testbed for both the underlying concepts and their implementation.

* The map must be represented in a form that allows inherent geometrical and topological properties to be retrieved readily. It must also permit storage of symbolic information that goes beyond what is contained in conventional maps.

* Determining and using the correspondence among photo, map, and scene is a crucial part of our approach. Photogrammetry and AI research have together given us most of the necessary computational tools. The one remaining gap is determining which map and photo features correspond. For the time being, we rely on the user to bridge that gap, and we employ established algorithms to do the rest.

* Guided tracing of a linear feature, such as a road or coastline, is a particularly clear way in which knowledge in the form of a map may be used to guide analysis of an image.
3. Data

In order to provide a coherent problem area for our research, we have selected a single geographical area and obtained comprehensive photo coverage of it for use as our primary set of data. The area we have selected is centered upon Oakland, California. Included in this area are Alameda Naval Air Station, Treasure Island Naval Reservation, Oakland Army Terminal, Oakland Naval Supply Center, the cities of Oakland and Berkeley, and the mountains behind them. Also in this area are railyards, harbors, airfields, bridges (including the Bay Bridge), freeways, urban areas, open hillsides, lakes, and streams.

We have obtained photo coverage from the US Geological Survey, including:

* NASA Skylab coverage of the entire San Francisco Bay area, taken from an altitude of 250 miles.
* U2 coverage of the Oakland area - stereo pairs of vertical and oblique views, taken from 60,000 feet.
* High altitude mapping photography, taken from 40,000 feet.
* Medium altitude mapping photography, taken from 15,000 feet, giving comprehensive overlapping coverage of the Oakland area.

In addition, we have obtained detailed USGS maps of the entire Oakland and San Francisco Bay areas. We have also obtained Digital Terrain Data tapes which give an array of ground elevations for this area; this terrain data was originally compiled by the Defense Mapping Agency.

From 544th Squadron, Strategic Air Command, we have received a set of (unclassified) reconnaissance photographs of various locations within the United States. These photographs provide representative examples of several common photointerpretive problems, including change detection, box-car classification and counting, airfield measuring, and storage container capacity measurement.
4. System Building

We are constructing a basic system to provide a framework for experimentation.

a. Requirements

The system must provide the following capabilities:

* Basic facilities for handling large digitized images.
* Basic image processing routines.
* Storage of maps and symbolic information in an appropriate data structure.
* Interactive facilities for adding to or modifying the map.
* Coordinate transformation routines for establishing and using correspondences among picture, map, and scene.
* Techniques for using map information to guide image processing.

b. System Structure

The structure of the basic system is shown in block diagram form in Figure 3. It is currently implemented as two communicating forks under the TENEX operating system. One fork is written in SAIL, for efficient numerical and array operations, and contains image handling, processing, and displaying routines. The other fork is written in INTERLISP, for symbol manipulation, and contains the map data structure, the routines for manipulating it, the interactive interface, and the main control routines.

c. Map Representation and Use

Maps are represented by a network data structure, which we call an Association Net. It is similar in organization to a Semantic Net, but differs in two ways: there is only one type of object -- an
association -- instead of two -- objects and relations: it is composed of direct pointers, instead of names that stand for objects. Associations are represented by lists, each of which has an indicator and a value. For example, (TYPE . ROAD) has TYPE as indicator and ROAD as value. The value may be a list, and in particular, a list of associations. For example,

(ARC (NAME . MAINSTREET) (TYPE . ROAD) (WIDTH . 50) ...)

Networks of roads are represented as networks in the data base. Significant points, such as intersections or bends, are represented by data objects with indicator NODE. Significant linear segments, such as portions of coastline or roads, are represented by objects with indicator ARC. Each NODE has an association POINT which gives it real world location, and as many ARCs as line segments
radiating from it (frequently two, for bends). For convenience, a linear segment is actually represented by two ARCs, each associated with one end point and with the other arc, as in Figure 4. This interlinked structure makes it particularly easy to trace a road by stepping through a sequence of ARCs and NODEs. It also facilitates editing by allowing the insertion of new nodes without radically disturbing the existing data structure.

![Diagram of a line segment representation](image)

**FIGURE 4 REPRESENTATION OF A LINE SEGMENT**

Each node and arc has an associated name and type. The name allows rapid indexing into the map structure, and following of named features, such as following roads through intersections. A type is an association list, like a node or an arc, which has an associated mode (line or point), a function for displaying objects of this type, and associated display color. The types currently in the system are: City, Bridge, River, Lake, Coast, Tank, Intersection, Building, Boundary, Cpoint, Tpoint and Invis. The latter three types are not displayed but are used for crossings of linear features, through points and conceptual linking of entities, respectively.

27
The nodes and arcs are contained in a "World" association list that represents the entire map. The world contains association lists of nodes and arcs according to type, thus giving a means of access to the entire association net for purposes such as creating a thematic overlay displaying all rivers.

A simple example of a fragment of a map is shown in Figure 5.

d. Image Representation and Use

The photographs we have collected are of very high quality, having an effective resolution of about 20000x20000 pixels. This poses something of a data conversion and storage problem: If we were to completely digitize every picture at the resolution necessary for detailed analysis of small features, the process would take months and would require hundreds of reels of magnetic tape. Our solution was to digitize* the pictures completely at the manageable resolution of 1000x1000 pixels of 8 bits of (logarithmic) density, and then to digitize selected picture fragments at higher resolution as required. This simulates a future operational system in which the original photographs are used as primary storage and fragments are digitized on-line, on demand.

A single picture at 1000x1000 resolution fills 256k of PDP-10 core, so that we are processing one 256x256 subimage at a time. We have written the necessary software to sample, reformat, file, and display pictures in this manner. We can, for example, sample the entire picture and display it on the 256x256 Ramtek display. We can then indicate an area interactively with the cursor and have the picture resampled and redisplayed so that the selected area now fills the entire screen.

* High quality image digitization was performed on an Optronics scanner at the Image Processing Institute, University of Southern California. We are grateful to Professor W.K. Pratt and his staff for their cooperation.
FIGURE 5  A SIMPLIFIED MAP FRAGMENT
Interactive routines have been written for using the display trackball to indicate points or draw lines overlaid on the displayed gray-scale picture. They incorporate the necessary coordinate transformation routines so that locations on the display screen may be measured and referred to in terms of full picture coordinates (rather than subimage coordinates).

We have also provided ourselves with basic picture processing primitives which operate on 256x256 arrays (of arbitrary byte size). Available functions include logical primitives that operate on 1-bit arrays, and operations for level slicing, scaling, and determining statistics, such as minimum and maximum values or histogram. More sophisticated functions include a routine for applying an arbitrary operator to points within a specified window and indicated by a 1-bit mask array, an edge-tracker that traces the boundary of a region designated by a predicate, and a line-fitting routine that approximates a traced boundary by a sequence of straight line segments.

All of the picture processing functions are encoded in SAIL and operate in a lower TENEX fork. Appropriate routines have also been written in LISP which interface to the SAIL routines so that they may be called by high level application-specific functions, or interactively from the user's console.

5. Initial Experiments

We are working toward the demonstration paradigm described earlier by constructing a crude, but complete, system and then refining expertise of the various components.

a. Completed Facilities

The current state of the implementation is best indicated by an example.

The user can retrieve an aerial photograph, and have it sampled and displayed at coarse resolution (Figure 6). He can then use
the cursor to indicate points and draw lines superimposed on the picture. For example, he can trace a road with the cursor, or indicate locations of intersections or buildings. Linear features are traced as a sequence of line segments and can be backed up and redrawn until the user is satisfied as to accuracy and appearance (Figure 7).

After the user has designated a name and type, the traced feature is automatically entered into the map database. For example, in drawing a minor road linking two previously traced major roads, the end points of the new road will be linked to nodes lying on the existing roads, if suitable ones exist. If no existing nodes are sufficiently close to the end points, new nodes will be automatically inserted into the existing roads. The updated map is then redisplayed over the image.

The user may now select an area for closer scrutiny by indicating it with the cursor on the display (Figure 8). The picture file is sampled appropriately and the selected area is displayed to fill the screen. A clipping routine determines which line segments in the map cross the displayed area, and displays those that are visible. At the same time, an index to the displayed nodes and arcs is created so that it is possible to identify rapidly which is being pointed at with the cursor (Figure 9).

The detailed display may now reveal that the earlier tracing on the coarse display was inaccurate. The user may then point at a node and have it moved to a new location, to which he also points. Nodes may also be inserted into existing arcs, to better approximate the shape of a curve, for example, and nodes or arcs may be deleted. Figure 10 shows the result of deleting an erroneous fragment of the map, while Figure 11 shows the insertion of a better fragment. Finally, Figure 12 shows the result of a sequence of editing operations in the selected area, and Figure 13 shows the entire map.

The editing routines maintain the integrity of the data base in these operations. It is also possible for the user to undo his actions, by using the LISPX command "UNDO". In this way he not only can
FIGURE 6  FULL PICTURE AT
256 x 256 RESOLUTION

FIGURE 7  COARSE MAP TRACED
ON PICTURE

FIGURE 8  AREA SELECTED FOR
DETAILED EXAMINATION

FIGURE 9  AREA DISPLAYED AT
FULL RESOLUTION
FIGURE 10  DELETION OF AN ERRONEOUS PIECE OF MAP

FIGURE 11  INSERTION OF A BETTER TRACE

FIGURE 12  RESULT OF EDITING MAP

FIGURE 13  THE MAP AT COARSE RESOLUTION
undo mistakes but also can freely construct temporary data structures for use in specialized processing later without worrying about difficulties in cleaning them up.

The different areas displayed by the user are remembered by the system, together with names that he gives them. He may thus return to a previously displayed area, such as the initial overall view, for reexamination or modification of the map.

In situations in which the feature being traced has good contrast, the user can call upon a simple automatic tracing routine. This routine tracks round the boundary of a region defined by any arbitrary predicate. The current default predicate is comparison with a threshold brightness, such as "Darker than threshold". The tracing routine must be provided by the user with a starting location within the region (usually via the cursor) and an initial direction (Figure 14). It steps in the specified direction until the predicate is no longer satisfied, and then tracks along the boundary between satisfaction and non-satisfaction.

The result of the edge tracer is a chain-encoded boundary of the region (Figure 15). This boundary is then passed to a line-fitting routine which approximates the boundary by straight line segments. For cartography we wish to represent the boundary by line segments that are nowhere more than a specified distance from the real boundary: in a conventional map, no feature may be more than 50 feet in error. This approximation is achieved by building up the line segments incrementally. Fixing one end point of the segment, the other end is stepped along the boundary one point at a time. For each position, the straight line joining the end points is computed, and the deviations of the intervening points from the line are also computed. If the maximum of these deviations is less than the permitted error, we step the end one point further and repeat. If the maximum deviation is greater than the permitted error, we have stepped too far. In this case, the previous point, or alternatively the point that possesses the maximum
FIGURE 14  START OF BOUNDARY TRACING

FIGURE 15  AUTOMATIC TRACE OF BOUNDARY

FIGURE 16  AUTOMATIC LINE FITTING AND INSERTION INTO MAP

FIGURE 17  RESULT OF EDITING THE TRACED BOUNDARY
deviation, is taken as the end point of this segment, and also as the
start of the next segment. The stepping and testing are repeated until
the entire boundary has been fitted to the desired accuracy; the result
of the fitting routine is a list of the line segments found (Figure 16).

The line segments found by the tracing and fitting
routines are then automatically incorporated into the map data
structure. Because of the elementary nature of the tracing routine, the
boundary may be in error in several places. The user can, however, edit
the map and redisplay the improved boundary (Figure 17). Despite the
very early stage of this work, the semi-automatic tracing and editing
process is much quicker than manually tracing the boundary with the
cursor.

b. Work in Progress

The current work is concerned with two problems:
coordinate transformations and smarter tracing.

At present, the coordinates of points in the map are
simply the coordinates of the feature in the full 1000x1000 picture.
The only transformation we currently handle is the simple one between
full picture and displayed subimage. We are thus simply tracing on the
picture without considering the true relation between the picture and
the ground, and without possessing the ability to use multiple pictures.
We are now in the process of testing the necessary routines to correct
these defects.

When the coordinate transformation routines are
integrated into the system, the user will be able to point at several
landmarks in the picture, giving their world locations if they are not
already in the map. The system will then compute the coordinate
transformation between picture and world by performing a numerical
optimization of the parameters of a camera model, which includes
location, heading, altitude, pitch, and roll. Once this optimization
has been done, it will be possible to represent world coordinates in the
map, to employ multiple pictures and correctly display a windowed portion of the map overlaid on the picture, to point at features in the picture and obtain their world locations, and to measure distances between two indicated points.

The current work on automatic tracing is directed at two problems: devising techniques for detecting and locating different kinds of linear features, and using approximate tracings to guide the detector.

The linear features with which we are concerned may be one-sided, such as a lake boundary, or two-sided, such as a road, and the two cases require slightly different treatment. For example, in road tracing, if one edge becomes hard to detect, the other may still reveal the course of the road. Generally speaking, knowledge about the type of feature being traced and statistics of what has so far been found may be used to tailor the detector and greatly improve its performance. We are currently studying the general characteristics of roads -- which are far from being the ideal of a uniform line on a uniform background. Existing edge detectors do not appear to be optimally matched to real roads.

We are considering several ways of guiding the tracing. One possibility is to apply the detector generally, over the region where the road may be, and to weight its output according to distance from the approximate tracing. The problem then becomes one of finding the path for which the sum of the weighted outputs is greatest. Relatively little work has been done previously on guided tracing, especially in the complex and variable domain of aerial photographs.

When guided tracing is an integral part of our system, it should be possible for the user to sketch features roughly and for the system to trace them accurately with reasonable reliability. Features that are already in the map may be found and traced entirely automatically. A natural extension of this technique will allow determination of the coordinate transformation more automatically.
D. Conclusions

1. Applications Requirements

The most important findings of the work reported here are that many problems in Cartography and Photointerpretation are bottlenecks in production and require the approaches of Machine Vision to automate them.

In Cartography, the primary need is for automatic tracing of features in aerial photographs, both of new features and of features on existing maps. This need in turn requires use of knowledge of the characteristics of features traced and accurate registration of map and picture for photogrammetric purposes.

In Photointerpretation, specialized aids, such as for counting, measuring, and change monitoring, are required to alleviate the photointerpreter’s burden. Such capabilities must draw upon a great amount of knowledge from the map data base to perform reliably.

The common requirements of the two domains thus include establishment of the correspondences among map, picture, and scene, and the exploitation of diverse types of knowledge to guide image analysis.

The current techniques of Machine Vision appear adequate for dealing with some of these tasks at a basic level. We recognize that complete automation is some years away; the optimal approach is via an interactive system that can employ advice offered by the user to guide its activities.

2. Practical Issues

We have already encountered some practical impediments which are not within the immediate scope of the current and proposed research, but which must eventually be removed to accomplish applications objectives. In the meantime, we circumvent them to minimize their drain on our resources.
The quantity of information inherent in a high resolution aerial photograph is too great to handle entirely digitally with ease. The solution is to use photographs as primary storage and digitize pieces of them on demand. We are simulating this solution.

The low level operations on pictures are simple but time consuming. Many of them could be readily accomplished by special-purpose hardware, yielding a speed increase of several orders of magnitude. In the absence of such hardware, we perform all operations sequentially on the computer.

The programming languages currently available do not support application-oriented image understanding: both compilation of efficient code and handling of symbolic information are necessary, and no single language provides both. In addition, very large addressing spaces are needed. For the time being, we employ several communicating forks, written in different languages. This compromise enables us to exploit the best features of several different languages to some extent, but costs considerable overhead in programming, redundant representation, and data communication.

3. Scientific Issues

The central scientific issue in the above applications, and in Machine Vision in general, involves the role of knowledge. It is evident that the more knowledge that is brought to bear on a problem, such as road tracing, the better the performance can be. The nature of the knowledge employed and the manner of its employment are open questions which must be answered before high performance image understanding systems can be constructed. Resolution of these issues is crucial to the attainment of more fully automatic cartography and photointerpretation.
REFERENCES


Appendix I

CARTOGRAPHIC AND PHOTO INTERPRETIVE REFERENCES
Appendix I

CARTOGRAPHIC AND PHOTO INTERPRETIVE REFERENCES

I. General Reference Material

A. Reference Books


B. Major Journals

1. Photogrammetric Engineering and Remote Sensing

C. Major Conference Proceedings

1. American Society of Photogrammetry Falls Church, Virginia (Annual)


II. Department of the Army Technical Manuals

A. GEODETIC AND TOPOGRAPHIC SURVEYING, TM 5-441, February 1970.

B. ENGINEER INTELLIGENCE, FM 5-34.


D. ENGINEERS' REFERENCE & LOGISTIC DATA, FM 5-35.

E. FOREIGN MAPS, TM 5-248, October 1963.
F. FIELD ARTILLERY CANNON GUNNERY, FM 6-40.

G. MILITARY GEOGRAPHIC INTELLIGENCE (TERRAIN), FM 30-10, March 1972.

H. MILITARY SYMBOLS, FM 21-30.


J. STAFF OFFICER FIELD MANUAL, FM 101-10.


L. ELEMENTS OF SURVEYING, TM 5-232.


N. MAPPING FUNCTIONS OF THE CORPS OF ENGINEERS, TM 5-231.


P. TERRAIN MODELS & RELIEF MAP MAKING, TM 5-249.

Q. SPECIFICATIONS FOR MILITARY MAPS, TM 5-1, September 1974, revised December 31, 1975.

R. TACTICAL INTERPRETATION NOTEBOOK, TM 30-246.

S. PHOTOGRAPHIC, TM 30-245.

T. MAP READING, FM 21-26.

III. DMA Reports

A. COMMONWEALTH SURVEY OFFICERS CONFERENCE 1975, DMA AUTOMATED CARTOGRAPHIC SYSTEMS by Robert A. Penney.

B. OFF-LINE ORTHOPHOTO PRINTING AT THE DEFENSE MAPPING AGENCY TOPOGRAPHIC CENTER by Hayne B. Dominick.

C. COLOR SEPARATION SYMBOLIZATION IN SEMIAUTOMATED MAP PRODUCTION by William H. Burdette.

D. DIGITAL TOPOGRAPHIC INFORMATION BANK by Henry R. Cook.
E. MAKING A MAP WITH DIGITAL DATA by Robert L. Struck.

F. AUTOMATED DELINEATION OF GROUND SLOPE by Merle J. Biggin.

G. PROBLEMS, SHORTFALLS, AND NEEDS OF TOPOGRAPHIC MAPPING by Reuben Cook.

IV. US Army Engineer Topographic Laboratories' Reports


B. PARALLEL OPTICAL PROCESSING TO CONVERT ELEVATION DATA TO SLOPE MAPS, PHASE II: PRACTICAL CONSIDERATIONS, February 1975.


D. SURFACE MATERIALS AND TERRAIN FEATURES OF YUMA PROVING GROUND, May 1975.

E. A TERRAIN EFFECTS ANALYSIS ROUTINE FOR AN MGI SYSTEM, April 1975.

F. TEXTURE TONE STUDY: SUBJUARY AND EVALUATION, March 1975.

G. A SYSTEM FOR TOPOGRAPHIC INQUIRY NO. 3 ALPHANUMERIC SUBSYSTEM DATA BASE LISTING, March 1975.


I. PARALLEL OPTICAL PROCESSING TO CONVERT ELEVATION DATA TO SLOPE MAPS, PHASE I, THEORETICAL ANALYSIS, October 1974.


L. ANALOG GRAPHIC PROCESSING FOR 3-D TERRAIN DISPLAYS, PROFILES, AND ELEVATION LAYER TINTS, October 1975.
Appendix II

CONTACTS WITH CARTOGRAPHIC AND PHOTOINTERPRETIVE CENTERS

I Defense Mapping Agency Topographic Center, Washington, D.C.

D. Meier, Office of the Director

II U.S. Army Engineering Topographic Labs., Fort Belvoir, Virginia

H. Carr, Chief, Autocartography Group
L. Gambino, Chief, Computer Science Lab
B. Scheps, Chief, Technology Development Branch

III U.S. Geological Survey, Reston, Virginia

M. McKenzie, Chief of Photogrammetry

IV U.S. Geological Survey, Menlo Park, California

A. Stein, Photogrammetry
D. Edson, Autocartography

V Strategic Air Command, Offutt Air Force Base, Nebraska

Col. J.L. Passauer, Vice Commander, 544th Aerospace Reconnaissance Technical Wing
Maj. T. Profett, Air Force Global Weather Center

VI Rome Air Development Corp., Griffis Air Force Base, New York
R. Hoffman
Maj. J. Broglie

VII Central Intelligence Agency, Washington, D.C.
L.F. Wise

VIII Aeronutronics Ford, Palo Alto, California
R. Asendorf
S. Fraelick
R. Widergren

IX. Lockheed Missiles and Space Co., Palo Alto Research Labs, Palo Alto, California
M.A. Fischler