

AD-A197 950

RADC-TR-88-11, Vol IV (of eight)
Interim Technical Report
June 1988



4

NORTHEAST ARTIFICIAL INTELLIGENCE CONSORTIUM ANNUAL REPORT 1986

Part A: Hierarchical Region – Based Approach to Automatic Photointerpretation; Part B: Application of AI Techniques to Image Segmentation and Region Identification

Syracuse University

James W. Modestino and George Nagy



This effort was funded partially by the Laboratory Director's fund.

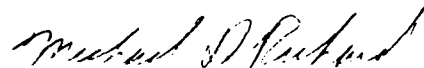
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss AFB, NY 13441-5700

This report has been reviewed by the RADC Public Affairs Division (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

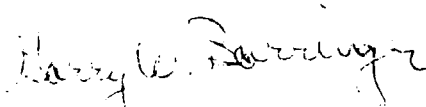
RADC-TR-88-11, Vol IV (of eight) has been reviewed and is approved for publication.

APPROVED:



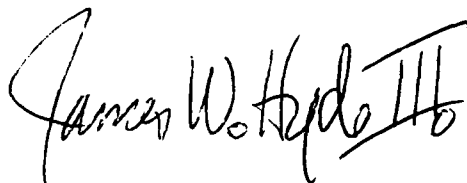
MICHAEL D. RICHARD, Capt, USAF
Project Engineer

APPROVED:



GARRY W. BARRINGER
Technical Director
Directorate of Intelligence & Reconnaissance

FOR THE COMMANDER:



JAMES W. HYDE, III
Directorate of Plans & Programs

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (IRRE) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notice on a specific document requires that it be returned.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

ADA197958

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b RESTRICTIVE MARKINGS N/A	
2a SECURITY CLASSIFICATION AUTHORITY N/A		3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b DECLASSIFICATION/DOWNGRADING SCHEDULE N/A			
4 PERFORMING ORGANIZATION REPORT NUMBER(S) N/A		5 MONITORING ORGANIZATION REPORT NUMBER(S) RADC-TR-88-11, Vol IV (of eight)	
6a. NAME OF PERFORMING ORGANIZATION Northeast Artificial Intelligence Consortium (NAIC)	6b OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION Rome Air Development Center (COES)	
6c. ADDRESS (City, State, and ZIP Code) 409 Link Hall Syracuse University Syracuse NY 13244-1240		7b. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Rome Air Development Center	8b OFFICE SYMBOL (if applicable) (COES)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F30602-85-C-0008	
8c. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700		10. SOURCE OF FUNDING NUMBERS	
		PROGRAM ELEMENT NO 62702F (over)	PROJECT NO 5581
		TASK NO 27	WORK UNIT ACCESSION NO 13
11 TITLE (Include Security Classification) NORTHEAST ARTIFICIAL INTELLIGENCE CONSORTIUM ANNUAL REPORT 1986 Part A: Hierarchical Region - Based Approach to Automatic Photointerpretation; Part B: Application of AI Techniques to Image Segmentation and Region Identification			
12 PERSONAL AUTHOR(S) James W. Modestino and George Nagy			
13a. TYPE OF REPORT Interim	13b TIME COVERED FROM Jan 86 to Dec 86	14. DATE OF REPORT (Year, Month, Day) June 1988	15 PAGE COUNT 64
16 SUPPLEMENTARY NOTATION This effort was performed as a subcontract by the Rensselaer Polytechnic Institute to Syracuse University, Office of Sponsored Programs. This effort was funded partially by the Laboratory Director's Fund.			
17. COSATI CODES		18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	
12	05		
		Artificial Intelligence; Photointerpretation	
		Expert Systems; Image Analysis	
		Machine Vision	
19 ABSTRACT (Continue on reverse if necessary and identify by block number) The Northeast Artificial Intelligence Consortium (NAIC) was created by the Air Force Systems Command, Rome Air Development Center, and the Office of Scientific Research. Its purpose is to conduct pertinent research in artificial intelligence and to perform activities ancillary to this research. This report describes progress that has been made in the second year of the existence of the NAIC on the technical research tasks undertaken at the member universities. The topics covered in general are: versatile expert system for equipment maintenance, distributed AI for communications system control, automatic photointerpretation, time-oriented problem solving, speech understanding systems, knowledge base maintenance, hardware architectures for very large systems, knowledge-based reasoning and planning, and a knowledge acquisition, assistance, and explanation system. The specific topics for this volume are the use of expert systems for automated photointerpretation and other AI techniques to image segmentation and region identification.			
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS		21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a NAME OF RESPONSIBLE INDIVIDUAL Michael D. Richard, Capt, USAF		22b TELEPHONE (Include Area Code) (315) 330-7788	22c OFFICE SYMBOL RADC (IRRE)

UNCLASSIFIED

10. SOURCE OF FUNDING NUMBERS (Continued).

Program Element Number	Project Number	Task Number	Work Unit Number
62702F	4594	18	E2
61101F	LDFP	15	C4
61102F	2304	J5	01
33126F	2155	02	10



Administrative Form

NAME: []
DATE: []
BY: []
Justification: []

By: []
Distribution: []
Availability Codes: []

Doc: []
A-1

UNCLASSIFIED

A.1 A Hierarchical Region-Based Approach to Automated Photointerpretation

Report Submitted by: J. W. Modestino
Electrical, Computer & Systems Engineering
Rensselaer Polytechnic Institute
Troy, New York 12180
(518)276-6823

Table of Contents	Page
I Introduction.	1
II Background.	2
III Technical Discussion.	3
IV Summary and Conclusions	17
References.	19
Table 1..Distinct Cliques Associated with the Adjacency Graph of Figure 10.	20
Table 2..Distinct Cliques Associated with Adjacency Graph of Figure 12b	21
Table 3..Summary of Region Knowledge and Associated Clique Function $V_1(\cdot)$	22
Table 4..Summary of Mutual Knowledge and Associated Clique Function $V_2(\cdot, \cdot)$	23
Table 5..Summary of Higher-Order Knowledge and Associated Clique Function $V_3(\cdot, \cdot, \cdot)$	24
Fig. 1..Automated Photointerpretation Testbed.	25
Fig. 2..Large 1024x1024 Aerial Image.	26
Fig. 3..Typical 256x256 Test Subimages Obtained from Original Image in Figure 2	27

Fig. 4..Segmentation Extracted from Test Image 1.	28
Fig. 5..Segmentation Extracted from Test Image 2.	29
Fig. 6..Segmentation Extracted from Test Image 3.	30
Fig. 7..Illustration of Iterative Region Segmentation Procedure.	31
Fig. 8..Illustration of Additional Results of Iterative Region Segmentation Procedure..	32
Fig. 9..An Initial Segmentation of an Image. . . .	33
Fig. 10..Adjacency Graph for Segmented Image in Figure 9.	34
Fig. 11..Illustration of Spatial Subdivision Process.	35
Fig. 12..A Schematic Image and Its Corresponding Adjacency Graph.	36

PART A

A HIERARCHICAL REGION-BASED APPROACH TO AUTOMATED PHOTOINTERPRETATION[†]

J. W. Modestino
Electrical, Computer and Systems Engineering Department
Rensselaer Polytechnic Institute
Troy, New York 12180

I. Introduction:

For the past year we have been evolving an approach to the development of an expert system for automated photointerpretation. This has included an extensive literature review, the development of some new and improved low-level image processing concepts, consideration of appropriate data and control structures and the evaluation of promising inferencing mechanisms. A major part of our work has been directed toward the development of a testbed which will serve the role of allowing demonstration of well-defined and developed concepts while at the same time serving as a development tool in exploring and testing new concepts. This testbed is being developed on the RPI Image Processing Laboratory (IPL) PRIME-750 System. It is our intent to gradually transition this testbed to one of the recently acquired IPL TI-Explorer Systems. However, this must await the incorporation of an interactive image processing and display capability into the Explorer.

In order to facilitate future development efforts and, in particular, to help guide evolution of the testbed, it's important at this point that a clear statement of present technical directions be provided. The purpose of the present note then is to provide a summary of the technical approach being considered at this time and to indicate future directions. This note can then be considered a working paper which can be amended or modified as work progresses.

[†] This work was supported in part by RADC under Contract No. F30602-85-C-0008.

II. Background:

There have been a number of attempts to develop limited-domain vision systems which provide semantic interpretations of raw image data. A good survey of some of the more promising techniques can be found in [1]. In most cases there are vast differences in the domain (i.e., aerial images, outdoor scenes, mechanical parts, etc.), the nature of the raw image data (i.e., resolution, monochrome or color, depth information, etc.), the purpose (i.e., industrial inspection, robot vision, aerial photointerpretation, etc.) and the use of world knowledge (i.e., simple constraint relations, 3-D geometrical models, 2-D template models, etc.).

It's important then to define the precise nature of the problem at hand, describe what we hope to accomplish, indicate the nature of the raw data and world knowledge we expect to have available and, finally, indicate potential future developments. We will attempt to accomplish this in the present section.

We expect to be working with medium to high altitude monochrome aerial imagery data. This imagery will include a variety of industrial, agricultural, military, residential, commercial, natural and man-made objects. We desire to be able to consistently segment the raw image data into distinct regions and provide a semantic description of these regions. This semantic description will specifically designate regions corresponding to a relatively small number of relevant objects together with a number of more general categories corresponding to objects which are either irrelevant, or for which no unambiguous interpretation can be provided. The relevant objects will include: roads, rivers, bridges, oil tanks, houses, aircraft, cars, runways, fields, forests, etc. For each of these relevant objects we will maintain an evolving knowledge database which not only

contains pertinent information on each relevant object, but also the spatial relationships between them. As new relevant objects are added to our list the knowledge database will have to be appropriately updated.

While initial development efforts will include only relatively primitive world knowledge, we hope to provide some flexibility for future expansion. For example, our present raw image database does not include any ground truth. In future work we might want to include map data to help in the photointerpretation process. Another possibility might be to use a previously interpreted image of the same scene as a guide in interpreting changes from one image to the next. Finally, we would not like to rule out the future possibility of using models, either 2-D or 3-D, of relevant objects to aid the photointerpretation process.

III. Technical Discussion:

In this section we will describe the current status of our automated photointerpretation system, review the pertinent details of the evolving testbed which will support it and illustrate some typical results obtained so far.

A block diagram of the overall testbed structure is illustrated in Fig. 1. The main function of the preprocessor is to provide a segmentation of the image into disjoint regions which are homogeneous within a region but differ in some sense from adjacent regions. We will be more specific on how this is accomplished later. It's important to note, however, that in order to be effective this segmentation does not make use of raw image data alone, but makes use of feedback from the interpretation process. In this sense we are implementing an interpretation-aided segmentation process.

Once a segmentation is obtained, however preliminary, the regions are labeled and region maps are stored in the image database. That is, the

actual pixel values associated with a region are stored separately for each region. In addition, various attributes associated with each region are stored. This includes such parameters as area, perimeter, boundary, elongation, etc. In addition, the spatial relationships between the various regions are maintained. This is most easily done by using an adjacency graph where the nodes correspond to regions and the connectivity indicating spatial relationships. In particular, two nodes are connected by an arc or edge if they are in some sense spatial neighbors. We will be more specific in defining what we mean by neighbors as we proceed. At any rate, the values associated with arcs can include mutual information corresponding to the connected nodes. This information might include: mutual boundaries, spatial distances, strength of mutual edges, etc. Image interpretations are provided by the inferencing mechanism which has access to the region information stored in the image database, as well as the world knowledge stored in the knowledge database. Feedback to the image preprocessor is through the inferencing mechanism.

It should be noted from Fig. 1 that the testbed allows operator intervention through an interactive image processing and display terminal. More specifically, the operator can manually extract regions using a joystick or trackball and, if desired, actually provide interpretation of the various extracted regions. Once the disjoint regions are outlined by the operator, the various region attributes are automatically extracted and stored in the image database in exactly the same format as if they were automatically extracted by the image preprocessor. Furthermore, in cases where the operator provides region interpretations the relevant spatial relationships are provided to the knowledge database allowing updating of our world knowledge.

The use of operator intervention then serves several purposes:

- a.) It can be used to isolate the image preprocessing from subsequent semantic interpretation by providing good segmentations.
- b.) It can be used as an aid in a partially automated system by resolving ambiguous segmentations or interpretations.
- c.) It can be used as a performance benchmark in assessing the efficacy of a fully automated photointerpretation system.
- d.) Finally it can be useful in developing and updating our knowledge database by providing correct interpretations of images.

Now let's describe how the interpretation-based segmentation scheme works. First we must recognize that very large images generally contain too much detail to be appropriate for automated photointerpretation, at least in early development efforts. This is illustrated in the reasonably large 1024x1024 image illustrated in Fig. 2 which contains much detail and many different types of distinguishable objects. In Fig. 3 we illustrate three 256x256 subimages extracted from the original image in Fig. 2. Each of these three subimages contains many more localized features and/or objects and are thus more suited to our early development efforts since we can maintain a much smaller knowledge base for each image and its associated relevant objects.

Suppose now that we obtain an initial segmentation of each of these test images. This segmentation can be effected on the basis of tonal or texture properties, or a combination of the two. As an example, we consider the tonal segmentation approach described in [2]. This scheme is based upon a clustering approach and requires a priori specification of the number of distinct region types, or classes. In Fig.'s 4-6 we illustrate the results of this initial segmentation for our three test images and for both three and

six classes. For comparison purposes we also include corresponding manually extracted segmentations.

Note that using six classes we tend to get reasonably good segmentations except in textured regions where a large number of very small regions are generated in each case. This could be improved somewhat by employing texture measurements in the segmentation or, alternatively, by attempting to merge these small regions with surrounding regions. Using three classes, on the other hand, gives a much coarser segmentation although the information provided is still useful. Unfortunately, it's not powerful enough to distinguish major objects from surrounding areas. For example, in Fig. 6 with three classes, we do not get good segmentation of the top of the oil tank from the surrounding ground area. Using 6 classes, on the other hand, we do get good segmentation of the oil tank from the surrounding ground, but now the vegetation area at the top of the figure produces a large number of somewhat irrelevant small areas.

Our approach has been to provide a crude initial segmentation employing three classes as a way to focus attention on large meaningful regions. The segmentation procedure is then repeated on individual regions and this process is continued until meaningful segmentations no longer are obtained. With individual regions silhouetted against dark backgrounds, this procedure can result in an individual region segmented into, at most, two regions. Since the scheme is based upon a clustering approach, we continue until either the ratio of the distance between cluster centers normalized to the geometric mean of the intraclass standard deviation is less than some prescribed threshold T_c , or the area of a region is below some threshold, T_s .

An illustration of this procedure is provided in Fig. 7. Here we begin with test image 3 in Fig. 7a which is segmented into the three classes in Fig. 7b and is identical to Fig. 6c. A connected region resulting from this segmentation which includes the tops of two oil tanks, as well as some of the ground area between them, is illustrated in Fig. 7c. This region is further segmented as indicated in Fig. 7d. Now the tops of the two oil tanks are separated from the ground area. This procedure should result in reasonably good initial segmentations. Additional segmentation results are illustrated in Fig. 8. More work needs to be done to determine appropriate threshold levels, T_c and T_s , in order to implement this stopping criterion.

While this hierarchical region-based segmentation procedure is quite simple, there are several areas where it can be improved considerably. Texture information should help with clustering now performed in a multi-dimensional feature space which includes texture as well as tonal features. Also, edge information should be useful in splitting two regions along a strong mutual edge. Furthermore, there are a number of possibilities which include feedback of interpretation information to help in splitting or merging regions to form visually meaningful segmentations.

Now suppose that an appropriate initial segmentation is obtained. Let the distinct regions be labeled $R_1, R_2, \dots, R_N$ as, for example, in Fig. 9 where $N=7$. The corresponding first-order adjacency graph associated with this segmented image then appears as indicated in Fig. 10. By first-order adjacency we mean here that regions are adjacent, or are neighbors, if and only if they are spatially contiguous. This concept of first-order adjacency should suffice for initial efforts although we should note that there are more general concepts of a neighborhood system that could be applied here. At any rate, the problem is now: given an initial

segmentation, to provide a global interpretation for each of the nodes given measurement attributes associated with each node, context information associated with the mutual relationships specified in the adjacency graph and world knowledge as prescribed in the knowledge database.

Before proceeding with a description of how this global interpretation is to be accomplished, one more comment is in order concerning the appropriate spatial size of our subimages. Consider, for example, the test image 2 as illustrated in Fig. 11. In Fig. 11a we illustrate the original image with a large manually extracted region, representing a road network, illustrated in Fig. 11b. One of the important characteristics of roads, at least locally, is that they are elongated and for this reason one of our important region measurement attributes is elongation. Unfortunately, the road network illustrated in Fig. 11b does not exhibit any elongatedness properties; the problem lies in the fact that the spatial scale is too large to observe this basically local property. In such cases it may make sense to further subdivide the extracted region until the elongatedness lies within certain ranges, or the area of the subdivided regions falls below some threshold. More specifically, we first extract the region possessing the largest area. If the elongation of this area, e_i , satisfies $e_i > T_{eu}$ or $e_i < T_{el}$ then we consider the spatial scale as appropriate; otherwise we divide the image in four quadrants and split the original region into at most, four parts corresponding to the quadrant in which the subregions fall, as illustrated in Fig. 11c. At this point a new adjacency graph is created for each of the resulting new regions. The process is then repeated until the elongation criterion is satisfied or the area of a subdivided region falls below some threshold, T_s . A typical result of this spatial subdivision process is illustrated in Fig. 11d where, in addition, we

illustrate a final subregion which meets the elongatedness criterion together with its surrounding or neighboring regions. This spatial subdivision process is then continued for all subsequent regions whose area is above some threshold, T_a .

Once regions have been spatially subdivided in this fashion, we then proceed to provide global interpretations. However, rather than provide a global interpretation over the entire image, we attempt this interpretation only over individual subquadrants which have resulted from the spatial subdivision process, e.g., for typical regions as illustrated in Fig. 11d. More specifically, we begin with the largest area region and initiate the spatial subdivision process. Take the region corresponding to the largest-sized subquadrant which results from the spatial subdivision process. We will initially focus attention upon this subquadrant in making a global interpretation. Any unambiguous interpretations that can be made in this subquadrant will then be propagated to neighboring subquadrants as initial conditions. We then proceed to the next largest subquadrant and repeat the global interpretation process with appropriate backtracking to previously explored regions to resolve inconsistencies. Much work needs to be done in defining how this is to be accomplished. Nevertheless, assuming that a global interpretation has been completed in the vicinity of the largest area region, we then proceed to do the same for the next larger region, etc., until this procedure has been completed for each region whose area is larger than the threshold, T_a . In this process we will propagate previous interpretations as initial conditions to newly explored neighboring regions. Also we must implement a backtracking scheme to insure that new interpretations do not result in inconsistencies with previous interpretations. At the conclusion of the process there may be some

uninterpreted regions. These can either be merged with neighboring regions for which unambiguous interpretations have been found or, at this point, an overall global interpretation can be attempted using as initial conditions the available interpretations of the large area regions and their immediate neighborhoods.

Suppose that within some subquadrant the regions are labeled $R_1, R_2, \dots, R_N$ and let $I_1, I_2, \dots, I_N$ be the corresponding global interpretations given to each of these regions where $I_i \in \{\phi, 1, 2, \dots, K\}$. Here we have K specific object types whose labels are to be assigned to each of the regions plus the ambiguous or irrelevant object type represented by the label or symbol ϕ . Suppose we define the region information as $\underline{R} = (R_1, R_2, \dots, R_N)$ and the interpretation vector $\underline{I} = (I_1, I_2, \dots, I_N)$. Note there are at most $(K+1)^N$ possible interpretation vectors although, in reality, there are many fewer than this since a valid global interpretation should not allow neighboring, or adjacent, regions to carry identical labels except for the uncertain symbol, ϕ . The exact number of interpretation vectors will then depend specifically upon the spatial arrangements of regions and is thus a random variable.

At any rate, our criterion will be to choose the estimated global interpretation $\hat{\underline{I}} = \underline{I}_0$ iff

$$p(\underline{I}_0 | \underline{R}, K, X) = \max_{\underline{I}} p(\underline{I} | \underline{R}, K, X) \quad (1)$$

Here, $\underline{R}$ represents information describing the partitioning into regions, K represents information in the knowledge database and X represents the corresponding adjacency graph which includes all measurement information, both for each region separately as well as mutual measurement information between regions. The quantity $p(\underline{I} | \underline{R}, K, X)$ represents the conditional

probability of $\underline{I}$ given R , K and X . This quantity may be difficult to specify theoretically, but it should be possible to approximate it empirically. The maximization in (1) is then over all legitimate interpretation vectors; the resulting estimate is called the maximum a posteriori (MAP) estimate and is well-founded in statistical decision and estimation theory [3].

At this point we will make the assumption that, conditioned on R, K and X , the interpretation vector $\underline{I}$ is a Markov random field (MRF) defined on the corresponding adjacency graph. The concept of a MRF defined on a 2-D lattice has provided a useful model for images and, for example, has been used as a texture model in [4]-[7]. However, as pointed out in [8], the concept of a MRF need not be restricted to lattices but can be defined on more general structures such as graphs. Thus, it appears quite natural to define the interpretation vector, $\underline{I}$, as an MRF defined on the associated adjacency graph.

Under the assumption that $\underline{I}$ is then a conditional MRF it's well-known, through the equivalence of MRF's with Gibbs random fields (GRF's), that the conditional probability must be of the form

$$p(\underline{I} | R, K, X) = \frac{e^{-U(\underline{I}; R, K, X)}}{Z}, \quad (2)$$

where $U(\underline{I}; R, K, X)$ is the associated Gibbs energy function and Z is the corresponding partition function which serves the role of a normalization constant. More specifically, we have

$$Z = \sum_{\underline{I}} e^{-U(\underline{I}; R, K, X)}, \quad (3)$$

where the summation is over all legitimate interpretation vectors. The energy function must then be designed to take into account the information represented by R, K and X . Before defining the precise nature of the energy

function, $U(\underline{I}; R, K, X)$, we will describe how the maximization in (1) is to be achieved.

As can be seen from (1) and (2), the MAP estimate is obtained by minimizing the energy function. This is a difficult combinatorial problem since, as we have noted previously, there are as many as $(K+1)^N$ possible interpretation vectors, $\underline{I}$. For example, with just 9 object types, we have as many as 10^N possibilities which can become impractically large for exhaustive search when we consider N can be as large as several hundred, even by employing the previously discussed spatial subdivision scheme which tends to keep N small by focusing attention in specific regions. Fortunately, there exist good, although heuristic, combinatorial optimization procedures which are ideally suited to this problem. In particular, we propose to use simulated annealing as first applied in [9] to combinatorial optimization problems.

Initially we choose an interpretation vector at random and a sufficiently high temperature parameter, T , which serves as a control parameter of the algorithm. We then perturb this initial interpretation vector in some well-defined way and measure the resulting energy difference, ΔU . If the energy has decreased (i.e., $\Delta U < 0$) we adopt the new configuration; otherwise we adopt the new interpretation vector with probability $\exp(-\Delta U/T)$. After sufficiently many iterations, the process tends to stabilize at one of a number of possible interpretation vectors which may represent only locally optimal solutions. The temperature is then lowered according to a prespecified so-called annealing schedule. Note that at high initial temperatures, $\exp(-\Delta U/T)$ is close to one for all positive ΔU so we tend to adopt the new interpretation with near certainty. For low temperatures, on the other hand, $\exp(-\Delta U/T)$ is close to zero for positive ΔU with the result

that we are not likely to adopt a new interpretation vector which increases the energy. Thus, we tend to make frequent changes for high initial temperatures and tend to be much more selective as the temperature decreases. The iterated sequence of solutions as T decreases tends to a global optimum in a number of steps much less than required by exhaustive search.

Now consider the choice of a Gibbs energy function. Again, it's well-known (cf. [8]) that this must be of the form

$$U(\underline{I}; R, K, X) = \sum_c V_c(\underline{I}_c; R, K, X), \quad (4)$$

where $V_c(\underline{I}_c; R, K, X)$ is called a clique function and the summation in (4) is over all possible cliques with $\underline{I}_c$ the restriction of $\underline{I}$ to the clique c . A clique is basically a set of nodes all of which are neighbors of each other. Unlike the case of a MRF defined on a lattice, where each node has identical connectivity (except possibly on the boundary), the connectivity of each node on a graph may be different. In particular, since the adjacency graph is determined by segmenting the image, the connectivity of each node representing a region can be highly variable. As a result, the cliques associated with each node may be quite different.

As an example, the cliques associated with region R_1 in Fig. 10 consists of the singleton $\{R_1\}$, the couples $\{R_1, R_2\}$, $\{R_1, R_5\}$ and the triple $\{R_1, R_2, R_5\}$. Similarly, the cliques corresponding to region R_2 are $\{R_2\}$, $\{R_2, R_1\}$, $\{R_2, R_3\}$, $\{R_2, R_4\}$, $\{R_2, R_5\}$, $\{R_2, R_1, R_5\}$, $\{R_2, R_3, R_4\}$, $\{R_2, R_4, R_5\}$. A summary of the distinct cliques associated with each node, or region, in the adjacency graph of Fig. 10 is illustrated in Table 1. The convention employed here has been to associate the first appearance of a given clique with the lowest indexed region to avoid double counting. For example, the

clique $\{R_1, R_2\}$ is associated with region R_1 and does not appear associated with node R_2 since it carries a larger index. In essence, we are exploiting the fact that $\{R_1, R_2\}$ and $\{R_2, R_1\}$ are identical cliques and to include both of them would result in double counting. In this manner we can partition the distinct cliques into disjoint sets associated with each region. Thus, the summation in (4) can be rewritten as

$$U(I; R, K, X) = \sum_{i=1}^N \sum_{c \in C_i} V_c(I_c; R, K, X) . \quad (5)$$

Here, the outer sum is over the individual nodes while the inner sum is over the set of distinct cliques, C_i , associated with node $i=1, 2, \dots, N$. The outstanding problem at this point then is in the determination and specification of an appropriate set of clique functions.

As an extremely simple illustration of how it's possible to assign clique functions, consider the simple schematic image in Fig. 12a, which is intended to illustrate a car on a road bordered on each side by fields and all under a clear sky. Our semantic object set is then $I = \{\text{sky, road, car, field}\}$, or alternatively $I = \{1, 2, 3, 4\}$ where now the object types are identified with the first four ordinate integers. The corresponding adjacency graph is illustrated in Fig. 12b with the associated distinct cliques provided in Table 2. There are relatively few distinct cliques in this case and regions R_3 - R_5 only require consideration of singletons. Region R_2 requires consideration of cliques composed of singletons and couples while region R_1 requires consideration of triples as well. Clearly, the correct interpretation vector in this case is $I_0 = (1, 2, 3, 4, 4)$.

Suppose that the measurement information X and knowledge K are very simple in order to illustrate this approach. More specifically, assume X consists of

A.) Region Attributes:

- 1.) Area A_i , $i=1,2,\dots,N$
- 2.) Average Gray Level G_i , $i=1,2,\dots,N$

B.) Mutual Attributes:

- 1.) Common Boundaries $B_{i,j}$ for neighboring R_i, R_j .
- 2.) Contrast $C_{ij}=|G_i-G_j|$ for neighboring R_i, R_j .

Furthermore, suppose that the knowledge database information, consists of the following:

A.) Region Knowledge:

- 1.) Cars generally have area less than A_c and average gray level equal to G_c .
- 2.) Sky generally has area greater than A_s and average gray level less than G_s .
- 3.) Roads generally have area equal to A_r and average gray level equal to G_r .
- 4.) Fields generally have area equal to A_f and average gray level greater than G_f .
- 5.) These quantities are related by $A_c < A_r < A_f < A_s$ and $G_s < G_r < G_c < G_f$.

B.) Mutual Knowledge:

- 1.) Sky and car do not share a common boundary.
- 2.) Field and car do not share a common boundary.
- 3.) Sky and road generally share a small common boundary of length less than B_{sr} and typically possess contrast equal to C_{sr} .
- 4.) Car and road typically have a common boundary equal to B_{cr} and small contrast less than C_{cr} .
- 5.) Sky and field typically have a common boundary equal to B_{sf} and a large contrast greater than C_{sf} .
- 6.) The road and field share a large common boundary greater than B_{rf} and a contrast equal to C_{rf} .
- 7.) These quantities are related by $B_{sr} < B_{cr} < B_{sf} < B_{rf}$ and $C_{cr} < C_{rf} < C_{sr} < C_{sf}$.

C.) Higher-Order Knowledge:

- 1.) The only valid set of three adjoining regions is sky, road and field.

It's easy to begin by dispensing with the choice of a clique function $V_{(R_i, R_j, R_k)}(I_i, I_j, I_k; R, K, X)$ for each triple so let's start here. However, for notational convenience we will write this as $V_3(I_1, I_2, I_3)$ where we have dropped the functional dependence upon R , K and X and do not specify the particular clique (R_i, R_j, R_k) but assume this is implicitly understood. The quantities I_1, I_2 , and I_3 are then the interpretations to be given to regions R_i, R_j , and R_k , respectively, and the subscript 3 is a reminder that this is the clique function defined for triples. We will employ a similar notation for $V_{(R_i, R_j)}(I_i, I_j; R, K, X)$ and $V_{(R_i)}(I_i; R, K, X)$ replacing them by $V_2(I_1, I_2)$ and $V_1(I_1)$, respectively. At any rate, using the higher-order knowledge in K , which is the only information available for triples, a possible choice for $V_3(I_1, I_2, I_3)$ is

$$V_3(I_1, I_2, I_3) = \begin{cases} 0 & ; (I_1, I_2, I_3) = P(\text{sky, road, field}) \\ \infty & ; \text{otherwise} \end{cases} \quad (6)$$

Here $P(\text{sky, road, field})$ means any permutation of the enclosed interpretations.

For $V_1(\bullet)$ and $V_2(\bullet, \bullet)$ we must make use not only of the knowledge available in K but the corresponding measurement information in X . The region knowledge given previously is summarized in Table 3 together with an appropriate choice for $V_1(\bullet)$. Here, the clique function evaluated for a particular object type is defined on the row of Table 3 corresponding to that object type. For example,

$$V_1(I) = \alpha \left[\frac{1}{1 + (A_c/A_i)^2} \right] + \beta [G_i - G_c]^2 \quad ; \quad I = \text{car}, \quad (7)$$

where α and β are appropriately chosen scale parameters. The quantities A_i and G_i are the measured area and average gray level, respectively, of the underlying regions R_i . Note that the value of $V_1(\text{car})$ is small only for

areas, A_i , smaller than A_c and for gray levels, G_i , close to G_c ; characteristics associated with cars. Analogous comments apply to the values of $V_1(I)$ for $I \in I = \{\text{sky, road, car, field}\}$.

In Table 4 we summarize the mutual information available in K and also illustrate a possible choice for $V_2(\bullet, \bullet)$ in the same format as provided for $V_1(\bullet)$ in Table 3. Here, α' and β' are scale parameters and B_{ij} and C_{ij} are the mutual boundary length and contrast, respectively, corresponding to the underlying regions. Again, it can be seen that the contributions to the overall Gibbs energy function, are minimized only under the correct interpretation. Finally, in Table 5, we summarize the available higher-level knowledge and the corresponding clique function $V_3(\bullet, \bullet, \bullet)$.

In no way are we suggesting that the choice of clique functions described here are optimum in any sense. Rather, we have attempted to illustrate at least one way of choosing them in a consistent fashion for an admittedly simple, contrived problem.

IV. Summary and Conclusions:

We have attempted to describe in some detail a hierarchical region-based approach to automated photointerpretation. This approach has evolved from the past year's effort to develop and implement an expert system for automated photointerpretation. Much more work remains to complete the development of this system and to provide a complete evaluation of its performance in realistic photointerpretation tasks. Nevertheless, the work described here has provided a useful focus for our efforts and should provide a meaningful context for future investigations.

Among some of the research issues that we will be investigating in the future include the following:

1. Additional and more powerful features have to be incorporated into the segmentation procedure.
2. More effective stopping criterion for the iterative region segmentation procedure needs to be employed.
3. Object detection and boundary extraction procedures need to be incorporated.
4. More comprehensive region and mutual attributes need to be employed.
5. The manual segmentation procedure needs to be improved and interfaces with knowledge database worked out.
6. Our raw image database needs to be expanded.
7. More effective procedures for localizing search in the spatial subdivision process need to be developed.
8. General procedures for designing the clique functions need to be worked out.
9. Annealing schedules for effecting the simulated annealing search procedure need to be developed.
10. Propagation of interpretations from one region to the next needs to be investigated.
11. We need to provide feedback from the interpretation process to the segmentation process to improve its performance.
12. Flexible and effective data and control structures need to be developed.
13. We have to investigate how map data and/or archival, previously interpreted, image data can be utilized to improve the photointerpretation process or to implement change detection/interpretation procedures.

References

1. T. O. Binford, "Survey of Model-Based Image Analysis Systems", The Int. Jour. of Robotics Research, vol. 1, pp. 587-633, Spring 1982.
2. J. Zhang and J. W. Modestino, "Image Segmentation Using a Gaussian Model", unpublished RPI report, March, 1987.
3. H. L. Van Trees, DETECTION, ESTIMATION AND MODULATION THEORY; PART I, Wiley, New York, 1968.
4. G. C. Cross and A. K. Jain, "Markov Random Field Texture Models", IEEE Trans. Pattern Anal. Mach. Intel., vol. PAMI-5, pp. 25-39, Jan. 1983.
5. S. German and D. German, "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images", IEEE Trans. Pattern Anal. Mach. Intel. vol. PAMI-6, pp. 721-741, Nov. 1984.
6. J. Zhang and J. W. Modestino, "Markov Random Field Model with Applications to Texture Classification and Discrimination", Proc. of 20'th Annual Conf. on Information Sciences and Systems, Princeton Univ., Princeton, NJ, pp. 230-237, March 1986.
7. J. Zhang and J. W. Modestino, "Texture Classification and Discrimination Using the Markov Random Field Model", unpublished RPI report, March, 1987.
8. R. Kinderman and J. L. Snell, MARKOV RANDOM FIELDS AND THEIR APPLICATIONS, American Mathematical Society, Providence, RI, 1980.
9. S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, "Optimization by Simulated Annealing", Science, vol. 220, pp. 671-680, May, 1983.

Node or Region	Associated Cliques
R_1	$\{R_1\}, \{R_1, R_2\}, \{R_1, R_5\}, \{R_1, R_2, R_5\}$
R_2	$\{R_2\}, \{R_2, R_3\}, \{R_2, R_4\}$ $\{R_2, R_5\}, \{R_2, R_1, R_5\}, \{R_2, R_3, R_4\},$ $\{R_2, R_4, R_5\}$
R_3	$\{R_3\}, \{R_3, R_4\}, \{R_3, R_2, R_4\}$
R_4	$\{R_4\}, \{R_4, R_5\}, \{R_4, R_6\}, \{R_4, R_7\},$ $\{R_4, R_5, R_6\}, \{R_4, R_5, R_7\},$ $\{R_4, R_6, R_7\}, \{R_4, R_5, R_6, R_7\}$
R_5	$\{R_5\}, \{R_5, R_6\}, \{R_5, R_7\}, \{R_5, R_6, R_7\}$
R_6	$\{R_6\}, \{R_6, R_7\}$
R_7	$\{R_7\}$

Table 1

Distinct Cliques Associated with the
Adjacency Graph of Figure 10

Node or Region	Associated Cliques
R_1	$\{R_1\}, \{R_1, R_2\}, \{R_1, R_4\}, \{R_1, R_5\},$ $\{R_1, R_2, R_4\}, \{R_1, R_2, R_5\}$
R_2	$\{R_2\}, \{R_2, R_3\}, \{R_2, R_4\}, \{R_2, R_5\}$
R_3	$\{R_3\}$
R_4	$\{R_4\}$
R_5	$\{R_5\}$

Table 2

Distinct Cliques Associated with
Adjacency Graph of Figure 12b.

Object Type	Region Knowledge		Clique Function $V_1(\cdot)$
	Area	Average Gray Level	
Car	$\leq A_c$	$= G_c$	$\alpha \left[\frac{1}{1+(A_i/A_c)^2} \right] + \beta [G_i - G_c]^2$
Sky	$\geq A_s$	$\leq G_s$	$\alpha \left[\frac{1}{1+(A_i/A_s)^2} \right] + \beta \left[\frac{1}{1+(G_i/G_s)^2} \right]$
Road	$= A_r$	$= G_r$	$\alpha [A_i - A_r]^2 + \beta [G_i - G_r]^2$
Field	$= A_f$	$\geq G_f$	$\alpha [A_i - A_f]^2 + \beta \left[\frac{1}{1+(G_i/G_f)^2} \right]$

Assumptions: $A_c \leq A_r \leq A_f \leq A_s$
 $G_s \leq G_r \leq G_c \leq G_f$

Table 3

Summary of Region Knowledge and
Associated Clique Function $V_1(\cdot)$.

Object Type	Mutual Knowledge		Clique Function $V_2(\cdot, \cdot)$
	Boundary Length	Contrast	
Sky, Car	0	-	∞
Field, Car	0	-	∞
Sky, Road	$\leq B_{sr}$	$= C_{sr}$	$\alpha' \left[\frac{1}{1+(B_{sr}/B_{ij})^2} \right] + \beta' [C_{ij} - C_{sr}]^2$
Car, Road	$= B_{cr}$	$\leq C_{cr}$	$\alpha' [B_{ij} - B_{cr}]^2 + \beta' \left[\frac{1}{1+(C_{cr}/C_{ij})^2} \right]$
Sky, Field	$= B_{sf}$	$\geq C_{sf}$	$\alpha' [B_{ij} - B_{sf}]^2 + \beta' \left[\frac{1}{1+(C_{ij}/C_{sf})^2} \right]$
Road, Field	$\geq B_{rf}$	$= C_{rf}$	$\alpha' \left[\frac{1}{1+(B_{ij}/B_{rf})^2} \right] + \beta' [C_{ij} - C_{rf}]^2$

Assumption: $B_{sr} < B_{cr} < B_{sf} < B_{rf}$
 $C_{cr} < C_{rf} < C_{sr} < C_{sf}$

Table 4
 Summary of Mutual Knowledge and
 Associated Clique Function $V_2(\cdot, \cdot)$.

Object Types	High - Order Knowledge	Clique Function $V_3(.,.,.)$
Sky,car,field	impossible combination	∞
Sky,car,road	impossible combination	∞
Sky,road,field	valid combination	0
Road,car,field	impossible combination	∞

Assumption: Not all combinations of triples possible.

Table 5
Summary of Higher-Order Knowledge and
Associated Clique Function $V_3(.,.,.)$.

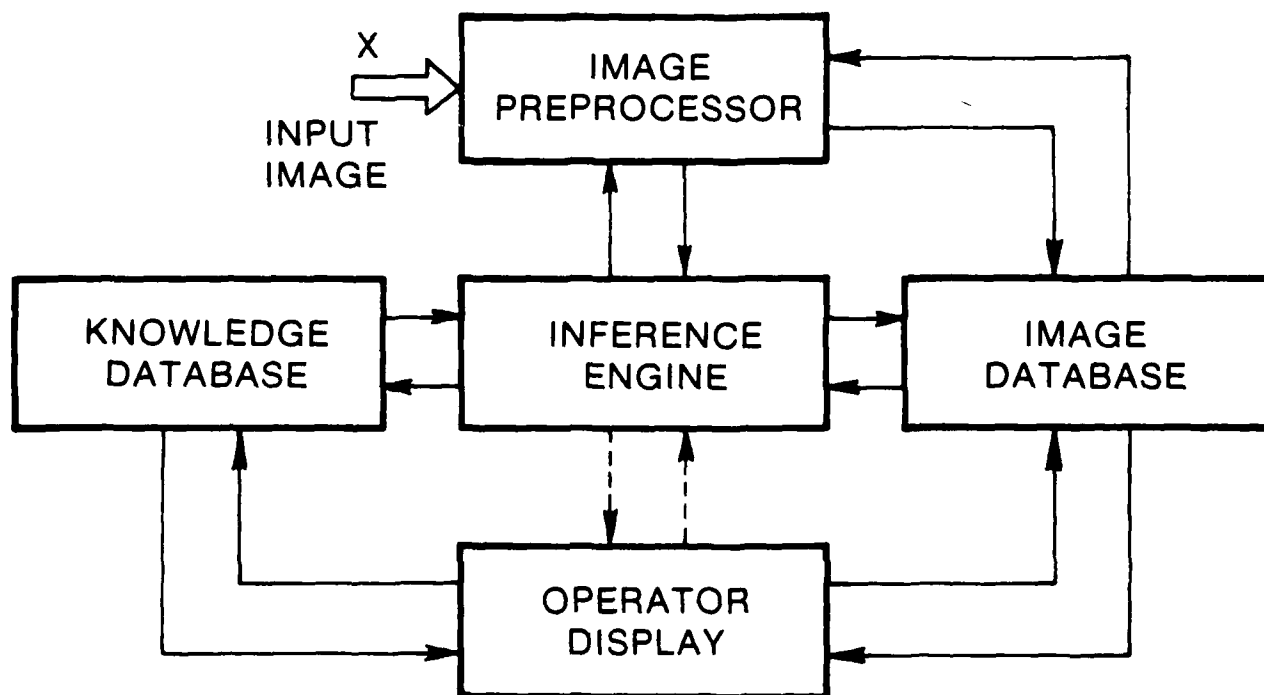


Figure 1
Automated Photointerpretation Testbed.



Figure 2

Large 1024x1024 Aerial Image.



a.) Subimage 1



b.) Subimage 1



c.) Subimage 3

Figure 3

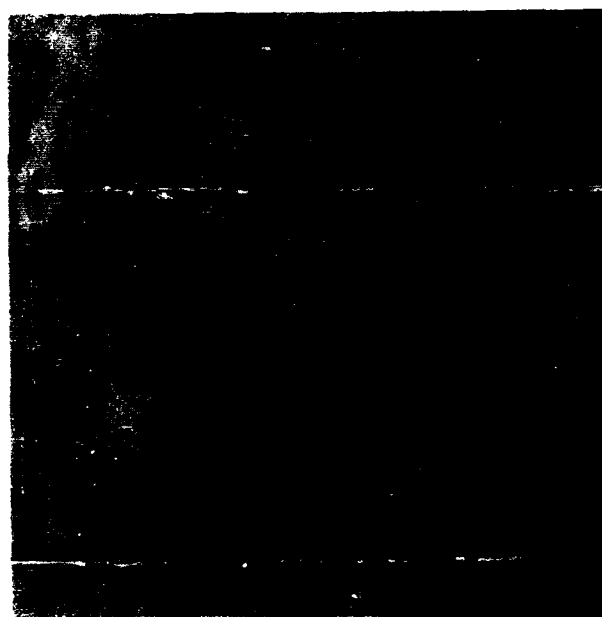
Typical 256x256 Test Subimages Obtained from
Original Image in Figure 2.



a.) Original Test Image 1



b.) Manually Extracted Segmentation



c.) Three-Class Segmentation



d.) Six-Class Segmentation

Figure 4

Segmentation Extracted from Test Image 1.



a.) Original Test Image 2



b.) Manually Extracted Segmentation



c.) Three-Class Segmentation



d.) Six-Class Segmentation

Figure 5

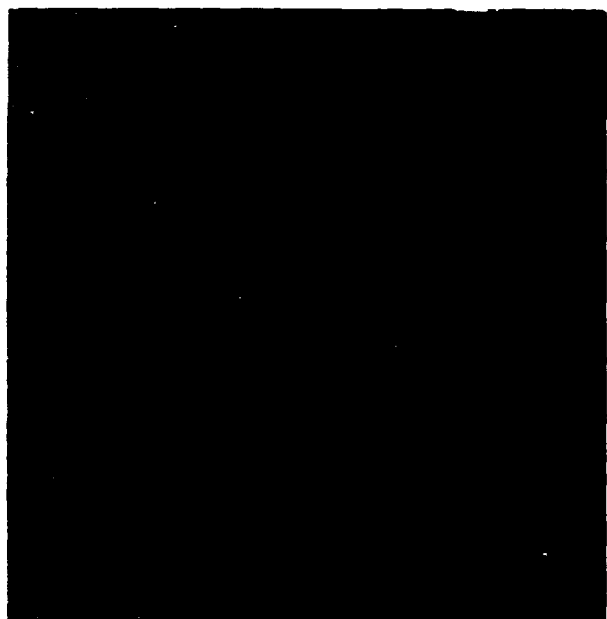
Segmentation Extracted from Test Image 2.



a.) Original Test Image 3



b.) Manually Extracted Segmentation



c.) Three-Class Segmentation



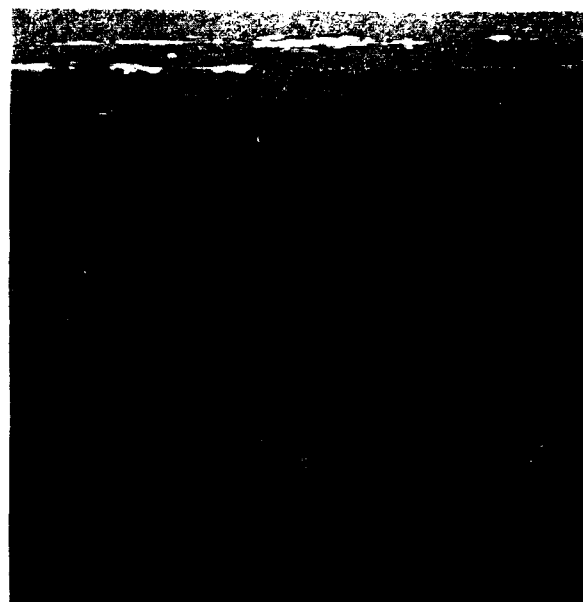
d.) Six-Class Segmentation

Figure 6

Segmentation Extracted from Test Image 3.



a.) Original Test Image



b.) Initial 3-Class Segmentation



c.) Resulting Region



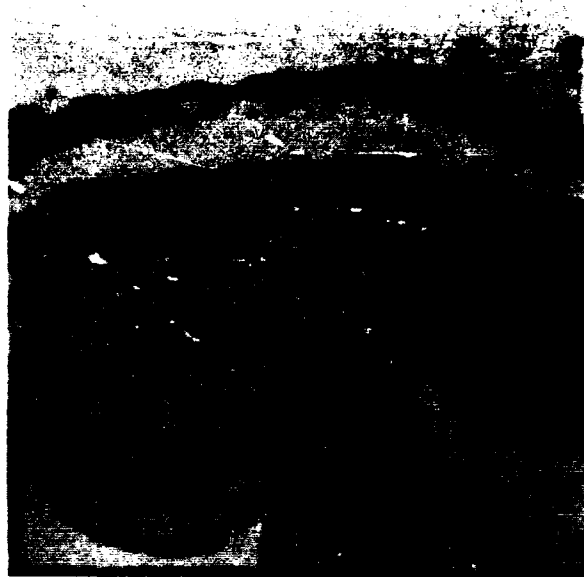
d.) Subsequent 3-Class Segmentation

Figure 7

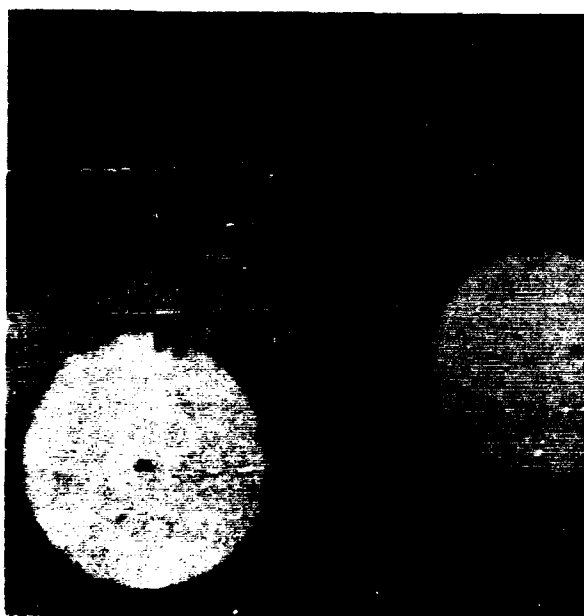
Illustration of Iterative Region Segmentation Procedure.



a.) Original Test Image 3



b.) Initial 3-Class Segmentation



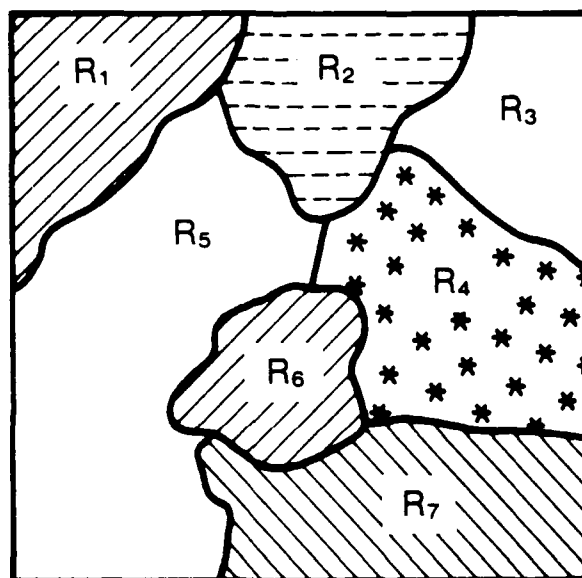
c.) First Iteration



d.) Second Iteration

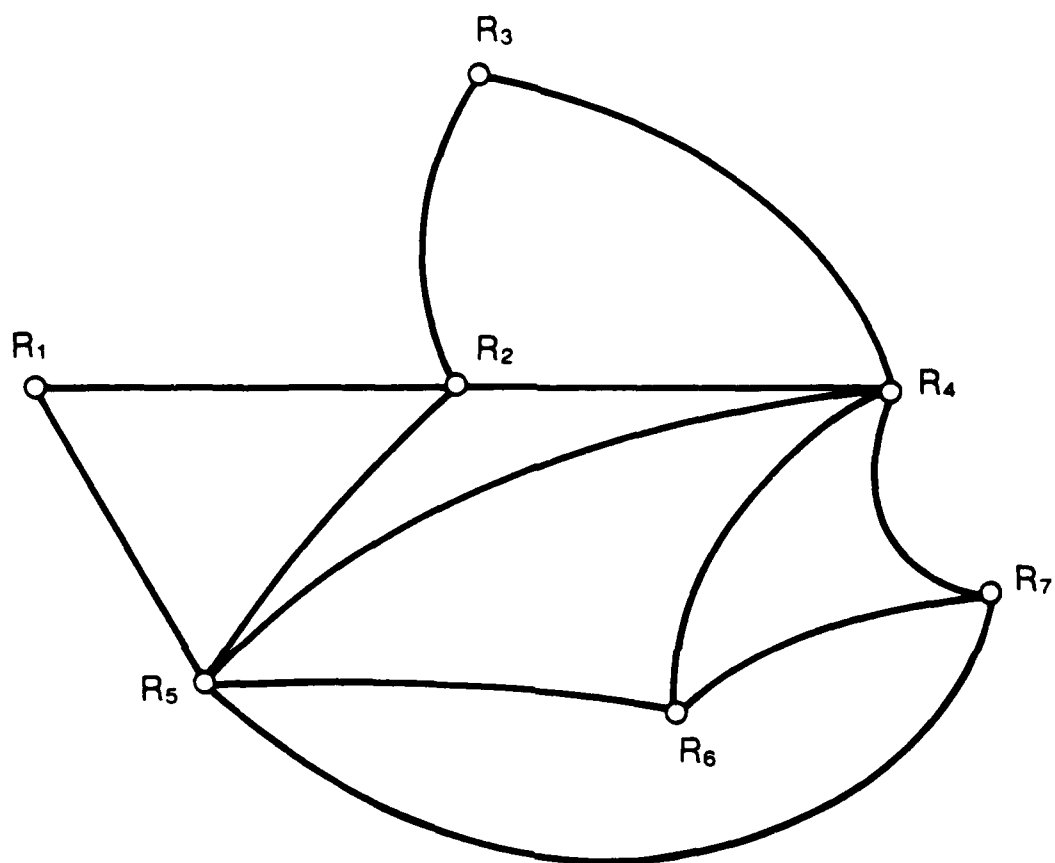
Figure 8

Illustration of Additional Results of Iterative Region Segmentation Procedure.



An Initial Segmentation of an Image.

Figure 9

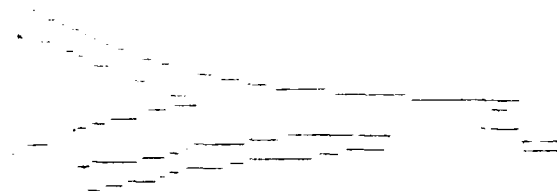


Adjacency Graph for Segmented Image in Figure 9.

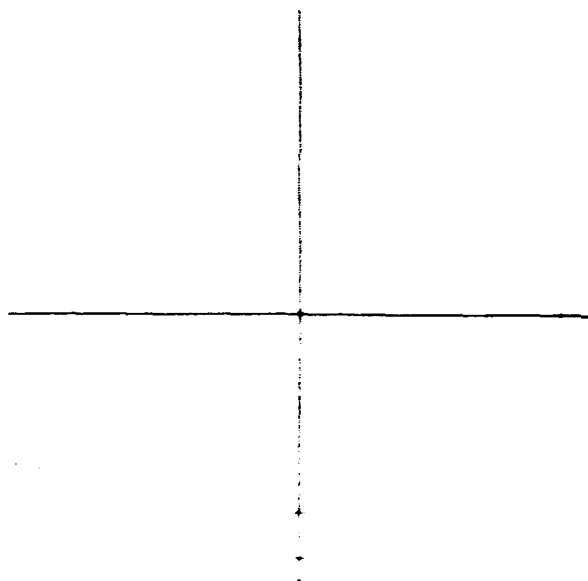
Figure 10



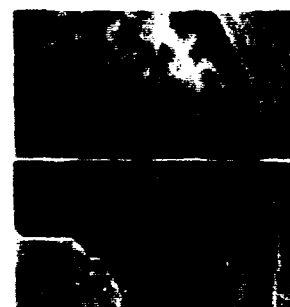
a.) Original Test Image 2



b.) Manually Extracted Road Network



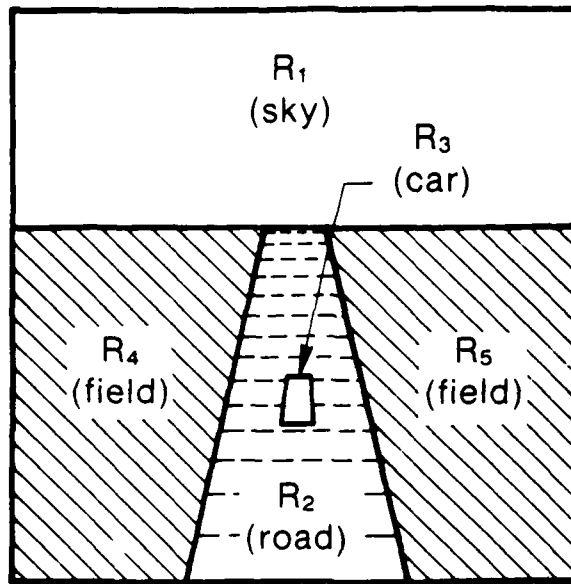
c.) Subdivision of Original Region into Quadrants



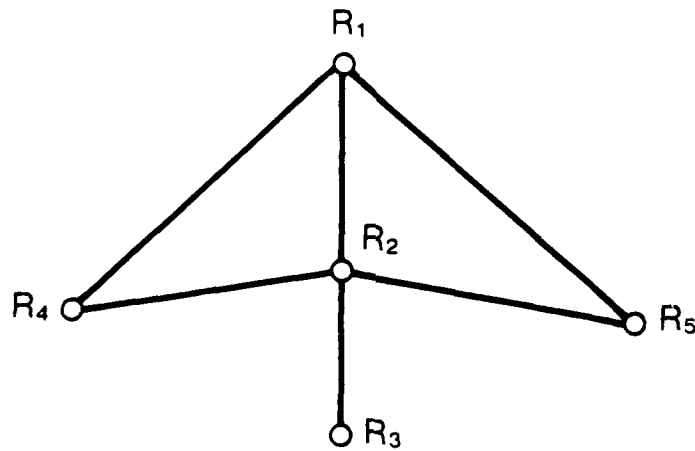
d.) Final Subregion with Neighboring Regions

Figure 11

Illustration of Spatial Subdivision Process.



a.) Schematic Image.



b.) Corresponding Adjacency Graph.

Figure 12

A Schematic Image and Its
Corresponding Adjacency Graph.

B Application of AI Techniques to Image Segmentation and
Region Identification

Report Submitted by: Dr. G. Nagy
Electrical, Computer and Systems Engineering
Rensselaer Polytechnic Institute
Troy, New York 12180
(518)276-6078

Table of Contents	Page
I Introduction.	37
II Digital Image Segmentation Using a Knowledge Base.	38
III Visibility-Oriented Criteria for Terrain Characterization.	44
Discussion.	46
Conclusions	51
References.	52

PART B

APPLICATION OF AI TECHNIQUES TO IMAGE SEGMENTATION AND REGION IDENTIFICATION*

Dr. G. Nagy
Electrical, Computer and Systems Engineering Department
Rensselaer Polytechnic Institute
Troy, New York 12180

I. Introduction:

This report covers the portion of the PRI-NAIC project under supervision during the period January 1-December 30, 1986. I was charged to the project .25 FTE during the academic year until my resignation from the project effective September 30, 1986 (i.e., 5.5 months at quarter time) and five weeks during summer 1986. Other participants in the project were the following:

Prof. M. Krishnamoorthy, RPI CS Dept., 3 wks. summer 1986; 0.1 FTE. Jan.-May '86;
Prof. T. Spencer, RPI-CS Dept., 3 wks. summer 1986; 0.1 FTE, Jan.-May '86;
Prof. S. Seth, UNL CS Dept., consultant, 10 days, summer 1986;
Mr. J. Kanai, grad. student, RPI-ECSE, 2 mos. summer 1986; .5 FTE, Sept. '86;
Mr. J. Yu, grad. student, RPI-ECSE, 4.5 MOS.; .25 fte, fall '86;
Mr. D. Allen, grad. student, RPI-ECSE, .25 FTE, 9 months.

The amount charged to the project was just under one man-year of effort during the period under consideration. Unfunded contributors to the project include Mr. N. Ferraiuolo, a recent graduate of RPI, and Professor D. Embly, BYU CS Department.

The principal theme of the research was the application of AI techniques--knowledge representation, heuristic search, and expert systems--to coupling image segmentation with the identification of isolated regions of an image. Two subsidiary themes also pursued were (1) the classification of topographic terrain features using global rather than local cues, and (2) the integration of digital images and ancillary information into existing, commercially available, relational database management systems. None of these projects was completed, since our initial plans were based on the expectation of three additional years of funding at the 1985 level.

* This work was supported in part by RADC under Contract No. F30602-85-C-0008.

II. Digital Image Segmentation Using a Knowledge Base:

From the reports of other investigators of automated photointerpretation, it appeared clear from the outset that a frontal attack on the problem, using available tools, would result at best in a demonstration of object location in a few selected photographs whose structure and content were incorporated into the software. To our knowledge, no one has succeeded in building a system capable of processing photographs even in a small problem domain (for example, airports), under the condition that the photographs are completely new to the system and to the research team.

Accordingly, the approach chosen was to work initially with a simpler class of images--digitized documents--and to concentrate our efforts on developing a system capable of extracting the structure of many and diverse images of this type. The interpretation of a 2000 x 2000 pixel array representing a digital document requires, in fact, a considerable degree of expertise, and is by no means a trivial task for an automated system. It is expected, however, to be easier than the interpretation of arbitrary aerial photographs, at least partly because of the high contrast and the dominance of orthogonal straight-line features rather than curves and shaded regions. Furthermore, the knowledge base is one shared by all readers of technical material and layout editors, and does not require a highly specialized and rare (particularly in a university environment) skill as does photointerpretation. Since our primary objective is the development of improved interaction between segmentation and classification, rather than improved techniques for either segmentation or classification in vitro, we consider the above task an ideal vehicle for our research.

Restated in terms of digitized documents, the interpretation problem is the following: given a set of digitized pages from a particular technical journal, demarcate each member of a class of application-dependent items such as title, author, first and second level subtitles, figure captions, abstract, acknowledgments, tables, photographs, line-drawings, program segments, and equations. It is assumed, of course, that the system has no recourse to optical character recognition: each component must be identified only on the basis of its size, shape, and geometrical relation to other components. The knowledge base consists of two parts: one is generic, and represents general information about technical document layout; the other one is publication-specific, and represents the layout practices and conventions shared by a family of digitized pages. It is expected that as the utilization of generic layout knowledge becomes more sophisticated, less and less data will have to be entered and stored for specific types of publications.

In this framework, the problem can be divided into a series of subproblems:

local segmentation method;

structural representation for a particular segmentation pattern;

operators that alter a specific segmentation pattern in a given direction;

representation of generic knowledge in terms of segmentation patterns;

representation of publication-specific knowledge in terms of segmentation patterns;

labeling schemata for specific regions based on the current segmentation pattern and the stored knowledge;

means of utilizing the layout knowledge to alter the segmentation pattern of a given document until a consistent set of labels can be assigned to each component of interest;

evaluations of the results.

Our progress during the last year will now be discussed under these headings.

Local Segmentation

The objective of the pixel-neighborhood based segmentation scheme is to divide the image into a set of nested rectangles in time linear with the product of the total number of pixels and the number of nesting levels. A further desideratum is to have the segmentation method operate on a compressed (e.g., RLC with Huffman coding) representation. A family of algorithms using different neighborhood sizes was investigated, but the simplest, based on thresholded black/white and white/black transitions in the projected profile of each rectangle, was deemed sufficient to allow further investigation of the more important problems.

The algorithm, coded in C for an IBM PC and a VAX 11/780, was used to segment CCITT Test Document #5 (appended), a two-column technical article with some figures and equations, down to the character-segment level. The image array consisted of 1728 x 2048 pixels. The code generated 8345 rectangular segments in about 7 minutes on a 6 MHz PC/AT. Methods of improving and speeding-up the algorithm were investigated but not implemented.

Structural Representation of the Image

The properties of a hierarchic data structure, the X-Y tree, were defined. The X-Y tree is similar in concept to the widely used quad-tree and its derivatives, with the important difference that variable location of the nested divisions of the X-Y tree allows representation of the structural components of the image. Further processing then involves only the X-Y tree, rather than the original pixel array. Since many operations involve only higher levels of the tree, this represents an important degree of abstraction.

Concrete representations of the X-Y tree were written in Pascal, C, and BASIC. A number of sample documents, including the CCITT test document mentioned above, were coded. The CCITT document resulted in about 8000 nodes, most of which were, of course, leaf nodes.

It is expected that the segmentation scheme will generally result in correct leaf nodes, but the structure of the document will not be appropriately represented, without feedback from the labeling phase, by the X-Y tree configuration. The original tree, resulting from "uninformed" segmentation, is called a physical tree. The corrected tree is called a logical tree.

Tree Operations

Kanai, Krishnamoorthy and Spencer were able to demonstrate a set of operations that are sufficient to transform incrementally any given X-Y tree into another X-Y tree with the same leaf nodes. The computational complexity of the algorithm is still under study, and the transformation algorithms have not yet been developed to a point that would warrant implementation. The operators were presented by Kanai at the Electronic Imaging Conference in Washington in October 1986.

Kanai has also investigated the performance of a number of other algorithms, including various types of neighbor-finding operations, on the X-Y tree. One algorithm was implemented in Pascal. His current view is that these algorithms are not computationally more complex than the corresponding operations on quad-trees, though the multiplicative constants are larger.

Generic Knowledge Representation

The development of a set of generic document constraints in terms of X-Y trees was undertaken. Such constraints govern the general structure of printed technical reports and articles. For instance, the horizontal composition of characters leads to words; horizontal composition of words leads to lines; vertical lines of approximately the same length constitute paragraphs; paragraphs and single-column figures are assembled into columns, and so forth. Detailed definitions are also required for line-drawings, equations, and tables. All of these notions can be coded into predicates where the variables are the contents of the nodes of an X-Y tree.

The generic knowledge represents all of the logical X-Y trees that would constitute representations of valid printed documents from any source.

Some idea of the amount of data necessary to specify generic layout features may be gained by inspecting the source code of general-purpose document formatters such as TeX or TROFF. We are only in the first stages of this task, but Yu coded a few simple rules for words and lines in EXSYS, a PC-based expert system. He demonstrated that generic labels could be assigned to a small fragment of the CCITT document.

Publication-Specific Knowledge Representation

The development of a set of publication-specific constraints was also undertaken. Specific constraints take into account the consistency of the layout of a class of documents, such as pages of the "Research Contributions" sections of C. ACM. They include the placement of the title, topic, author(s)' affiliations, author(s)' research interests, acknowledgment, responsible editor, abstract, date, page, number, copyright notice, subtitles, figure placement, and so forth.

The publication-specific constraints thus represent all the logical X-Y trees that could be considered legal for a given family of documents.

Labeling Schemata

Among the most difficult tasks facing us is the automation of knowledge acquisition. Although we have discussed a number of approaches, including the use of learning techniques from sample pages, "reverse engineering" document preparation macros, and translating the layout-editor's style book, we have not made progress in this direction. Therefore, we extracted the necessary information from human experts as a prerequisite to coding it in a form suitable for the labeling process.

In order to avoid having to develop our own inference engine, we have used available expert systems to label document components according to the constraints (rules) specified in the knowledge base. Krishnamoorthy coded in OPS-5 several dozen publication-specific rules (for the Research Contributions section of C. ACM, 1983) detailing the layout of single and multiple line titles and single multiple authors. The rules were coded by detailed examination of 5 selected articles. A blind test was then conducted using 5 other articles not previously seen by the coding and rule-writing team. Although this was a minuscule sample, we were encouraged by the fact that the titles and authors were correctly identified on the test document without any modification of the code.

Yu examined a number of commercially available expert systems compatible with the RPI computing environment. This work is documented in his thesis. Two systems, M1 and EXSYS, were purchased. Since entering the X-Y tree for a specific document proved cumbersome in M1, a test on a segment of the CCITT document was conducted on EXSYS. In this test some generic rules were coded

to recognize low-level layout constructs. The main result of the test was recognition of the amount of effort involved in writing generic rules in the expert system's specification language. If automatic methods of establishing the knowledge base (from sample pages, style manuals, or document formatter macros) cannot be established, then it will be necessary to develop a high-level specification language to expedite this task.

Tree Transformation

The naive segmentation scheme described above will not generally group document components according to their semantic value. However, Spencer and Kanai demonstrated that given correct segmentation at the leaf level (which is quite realistic), a physical tree can be transformed into a logical tree with the same leaf nodes using only two types of operations. Furthermore, this transformation can also be performed on any subtree. Our goal now is to use feedback from the expert system (i.e., an indication that a given subtree is not a valid entity) to carry out transformations of the physical tree until a valid configuration is obtained. We intend to carry out this approach according to both a top-down strategy, using publication-specific rules, and a bottom-up strategy, using generic rules.

The alternative to tree transformations would be to resegment any portion of the digitized image that cannot be labeled by the inference engine. The advantage of using tree transformations is that there is no need to access the image at the pixel level. In the CCITT document, for example, the manipulation is carried out in terms of the 8000 or so nested blocks rather than the 3,500,000 pixels. Furthermore, the X-Y tree provides the structure to formulate the knowledge base at a relatively high level compared to the video.

Validation

We are not at the point yet where we are ready to validate our results, but we have designed a series of experiments to do so. Our intention is to generate documents using a high-quality document formatter such as TeX or TROFF and a laser-printer. The resulting document will then be scanned on an Eikonix printer in the RPI Image Processing Laboratory, and segmented and labeled according to the methods described above. The description of the document produced in this manner will then be compared to the macro calls in the formatter.

In order to separate the effects of scanning alignment and inaccuracies from that of the analysis, we also intend to submit the pixel array sent to the laser printer to the same processing steps (except, of course, digitization). We have already used this technique to produce mock-ups of documents at a lower resolution (with a 100 lpi dot-matrix printer). This allows us to generate realistic document images under completely controlled conditions, which would be very difficult with aerial photographs.

Another important tool that we are developing is interactive segmentation of real (digitized) documents using a mouse and a high-resolution display. The facility to edit and label X-Y trees will provide us with an end-to-end document processing facility that will, in principle, allow us to make the transition gradually from completely manual to completely automated analysis. This technique is, of course, directly applicable to photographs.

Relevance to Automated Aerial Photointerpretation

Does the research discussed above have relevance to aerial surveillance for intelligence purposes? Although digitized document analysis has some valuable applications on its own, which we are pursuing in concert with Xerox, IBM, Nippon Telephone, and SUNY Buffalo, here we will consider only its implications for photointerpretation.

First of all, it is clear that natural scenes do not obey the rectilinear constraints imposed by the X-Y tree. However, the important feature of the X-Y tree for the downstream processing is its hierarchical nature. It is not far-fetched to conceive of segmentation methods for aerial photographs that result in multi-level nested regions. Furthermore, it is possible to devise tree transformations that would preserve the hierarchical nature of the data structure under regroupment of the lower levels. This is essential for carrying out operations at the highest possible level of abstraction.

The generic knowledge base for images would include such common-sense items as continuity and uniform width for long linear features (roads, rivers), square corners for rectilinear objects (buildings, most street corners, even crop fields), consistency of shadow directions, orthogonality for "crossings" (bridges, overpasses), row structures and road access for agricultural areas.

The specific knowledge base could be as detailed as a map of the area, or more general like attachment of cloverleaves to highways, periodicity of urban areas, proximity of control towers to runways, roughly equal size of cars and parking spots, circular symmetry of fuel containers, terraced fields in a given geographic area, features of desert landscapes, branching linear structure of railroad terminals, etc. Initially, sufficient challenge would be provided by very specific scenes such as photographs from different points of view and time of day and year of a single airport, university campus, or harbor installation. The next step would be to attempt to expand the knowledge base to describe a family of scenes with similar semantic content, such as a group of small urban airports.

What we hope to gain from our work with documents are:

- (1) a better understanding of the nature of data structures suitable to represent the results of low-level segmentation;
- (2) tools for interfacing a complex geometric segmentation structure with an inference engine;

- (3) ideas for a descriptive language for compiling an image-oriented knowledge base in terms of segmentation primitives;
- (4) a theory of background-foreground relationships;
- (5) insights and methodology on using feedback from the labeling or classification phase to rearrange the segmentation boundaries without resegmenting at the pixel level.

All of these represent difficult research problems. We believe, however, that any advances that we can make on the document problem will bear benefits for automated image interpretation.

III. Visibility-Oriented Criteria for Terrain Characterization:

The "visibility region" of a viewpoint on a surface (a single-valued real function of two independent variables, $z = f(x,y)$) is well defined. It contains all of the points that can be joined to the viewpoint by means of a line-segment that does not pass through the surface. The visibility region is, in general, neither convex nor singly-connected. In principle, the visibility region of every point on the surface can be computed. In practice, to ease the computation and storage of visibility regions, the surface can be approximated by a triangulated irregular network (TIN) as a set of piecewise-linear surface patches, and the viewpoints confined to the nodes of the network. In this case, the horizontal projection of the area visible from a viewpoint consists of polygons.

In a TIN, the terrain surface is represented by irregularly-spaced data points, each consisting of triples (x, y, z) . A triangulation of the data divides the data into disjoint triangles by introducing edges between the vertices (data points). Each edge is adjacent to exactly two triangles, unless the edge is on the convex hull (i.e., the boundary) of the data set. Because of its favorable properties for interpolation, Delaunay triangulation (the dual of the Voronoi tessellation of the projected data points) is the accepted standard. In the sequel, it is assumed that the surface has been Delaunay-triangulated, and the vertices and edges are represented in a suitable data structure.

Computed visibility regions have at least five different interesting types of applications:

1. Visibility for its own sake. Examples are the determination of the minimum number of observation points (e.g., firetowers) necessary to view an entire region. One might also be interested in scenic locations, or in paths with maximum or minimum visibility between origin and destination points.
2. Line-of-sight communications. One can determine the locations for the minimum number of television transmitters for an area, or the optimal

location of receivers. With portable transceivers, one might be interested in the locus of travel of a party, each of whose members must remain in uninterrupted communications with the others. The location of radar, laser, and sonar surveillance systems also belongs to this category.

3. Orientation and navigation. The profile of the horizon is a natural and simply extracted measurement that can be readily used by an observer to locate him/herself with respect to a topographic map. Orienting oneself in this manner is a prerequisite for successful navigation.

4. Data compression. In order to reduce the number of data values in a digital elevation model, one may be able to use visibility considerations to determine which points to keep and which to discard.

5. Extraction of significant terrain features. This class of applications is more speculative: we conjecture that the location and relation of visibility regions provides adequate information for determining important topographic terrain features, such as peaks, ridges, and valleys. Hence visibility information could be used both for the extraction of sketch maps from digital terrain models, and for the gross characterization (i.e., mountainous, hilly, alluvial, mesa) of terrain types.

The purpose of this research is to formalize such problems, investigate methods of solution, determine the computational cost of alternatives, and develop algorithms for specific applications.

Prior work

Digital terrain models are discussed in (Mark 1978 and Nagy 1979). Precise classifications of local topographic features are formulated in (Peucker 1975, Johnston 1975, Greder 1976, Nackman 1985, Frank 1986), and similar definitions are applied to grey-scale images in (Paton 1975, Watson 1984). Peucker advocated the notion of surface specific points, and Nackman developed a formal structure based on critical point configuration graphs: both are based on partial derivatives.

The relation between the "empty circle" criterion for triangulation and Voronoi diagrams was first demonstrated in (Delaunay 1934). Peucker is generally credited with developing triangulated irregular networks (Peucker 1978): the important notion of ordering the triangles with respect to a node was introduced in (Gold 1978). A survey of Delaunay algorithms and data structures may be found in (De Floriani 1985b)

There are three ways to relate the visibility problem on TINs to previous literature: by expanding 2-D visibility results, by modifying grid-based 2 1/2-D results, or by simplifying 3-D visibility results. Several researchers have examined visibility in planar polygons (El Gindy 1981 and 1983, Burton 1982). However, these algorithms depend on the 2-D assumption

that an edge closer to the viewpoint necessarily hides one further away in the same direction.

Grid-based methods are used mainly in generating perspective views of single-valued functions of two variables (Kubert 1968, Wright 1973, Anderson 1982). The algorithms are discretized to take advantage of the uniform sample spacing; removing the uniformity destroys the ordering property on which the algorithms are based. In a DTM based on such a grid rather than upon TINs, these methods would solve the visibility problem adequately.

Largely due to interest in computer graphics, there has been a great deal of work on visibility in three dimensions. A clear distinction, however, must be made between image space and object space visibility algorithms. Algorithms of the first type (image space) determine only how an image of the model will appear from a given viewpoint. They report the limits of visible areas as coordinates on an image and not on the model; therefore, they are not appropriate for this application.

The remainder, object-space algorithms, label surface patches according to their visibility from the selected viewpoint (Sutherland 1974, Weiler 1977, Sechrest 1983). These could be used with no modification to solve the 2 1/2-D problem. There are, however, sufficient simplifications possible with very little computational cost to warrant a new approach. The most immediate simplification is to note that there are no BOTTOM surfaces; therefore, any triangle observed from the underside must be invisible.

Recent bounds on worst-case algorithms are presented in (Devai 1984, 1986a, 1986b), and in (McKenna 1986). It is generally believed, however, that hidden-line and hidden-surface algorithms with optimal worst-case performance are inferior to non-optimal algorithms in the "average" case.

It is apparent that once visibility regions have been extracted, the choice of the best observation points is related to the facilities location and set covering problems of operations research (Handler 1979). The importance of visibility criteria for site location is discussed in the context of geographic information systems in (Creamer 1985), and for a military application of expert systems, nap-of-the-earth helicopter flights, in (Garvey 86). Our approach was presented at the Second Symposium on Spatial Data Handling (De Floriani 1986): in brief, during the year David Allen completed a Pascal program to extract the visibility regions from a surface approximated by triangular planar patches.

Discussion

The computation of visibility regions on a surface is an interesting problem in itself. As we have seen, it is related but not identical to widely researched tasks in computational geometry and computer graphics. The 2 1/2 dimensional problem we consider is intermediate between the full 3-D problem considered in the display of solid objects and the 2-D visibility problems posed by Toussaint and El Gindy. A potentially important new aspect is the

spatial coherence between adjacent viewpoints, over and above the spatial coherence of adjacent areas seen from the same viewpoint.

Our attention was originally drawn to computational geometry (in 1979) by a line-of-sight communication problem: what would be a good data structure for storing the topography of Italy for computing good locations for transmitter stations? In addition to television signals, many other electromagnetic signals used for communications, ranging and imaging propagate in straight lines. Particularly challenging is the determination of the locus of the trajectories of multiple agents interested in maximum dispersal of the party subject to preservation of line-of-sight communications.

The problem of locating an observer by means of observations taken with a digital imaging or range-finder instrument is complicated by the uncontrollable variability in ambient illumination and in the directional surface reflectance of the terrain. It is therefore advantageous to consider methods which depend only on the relatively easily determined visual horizon of the observer. We propose to investigate this problem with respect to both orientation and autonomous navigation.

Data compaction in digital terrain models eludes simple solution. The principal investigator and his colleagues have developed an algorithm for this purpose. The algorithm was based on hierarchical subdivisions into smaller and smaller triangular patches (De Floriani 1984, 1985a). This algorithm exhibited very good performance in terms of average vertical deviation from the original data. However, the resulting terrain visually appeared quite different from the original! The problem with the hierarchical subdivision was the introduction of ridge and valley lines that simply did not exist in the original but had a strong visual impact in the approximation. It was this experience, in fact, that lead us to consider visibility criteria as a means of preserving important features of the data.

Peaks, ridges, and valleys are universally recognized as significant terrain features. The significance of such features is usually determined in terms of their size relative to other such nearby features. As mentioned, however, digital elevation models normally represent the elevation of the terrain to a given degree of accuracy, without special emphasis of prominent terrain features. Such models that do attempt to extract significant features tend to approach the problem from a localized perspective, essentially applying discrete approximations to methods derived from the differential calculus for finding extrema. We propose, however, to represent such features at the expense of accurate reconstruction of the terrain at other points. In other words, our model will provide an abstraction based on prominent features.

How much of the surface one can see from any given point on that surface is an important topographic characteristic that one tends to observe subconsciously. Nevertheless, it is difficult to sketch the visibility regions of specific points on even a simple terrain model. Furthermore,

although it is easy to identify terrain features such as ridges and valleys using a 3-D physical model, it is considerably harder to sketch them on contour plots. We will try to use our program for extracting the visibility regions of selected points from location and elevation data for the purpose of determining such terrain features. The extracted features may be used either for the automatic generation of sketch maps or as key points and constraint edges for economical triangulated irregular networks.

Some connections between visibility and topography are the following. Points where the immediate neighborhood is invisible are convexities: many adjacent convexities constitute a dome. Large, multiply-connected vistas are properties of dominant peaks and ridges. In pits and valleys, the prospect is singly-connected and tends to change gradually. Thus visibility considerations suggest where the valley ends and the mountain begins--one can argue that one is out of the valley as soon as new vistas, over adjacent ridges, open up! If two points have the same region of visibility, then they are in the same valley; if they do not, then there is a ridge between them. Horizons that form the common boundary of the visibility regions of many observation points are usually significant ridges. The shape, orientation, and symmetries of regions of visibility also provide valuable clues to geological formation.

Of course, in addition to identifying features, one must ascertain the relations between them. These relations are locally hierarchical, but globally form a network.

Most of the methods found in the literature for topographic feature extraction ("geomorphology" or "topographic morphometry") are based on characteristic slope angles, local relief, spectral coefficients, or direction and strength of azimuthal trends. Our contention is that visibility models offer a better chance to extract significant terrain features than do methods based on local extrema and curvature.

For some types of terrain the model, as described, is undesirably fine-grained. Observation points may be established for every mole hill and gopher hole. It is possible to increase the grain size by the simple expedient of allowing the observer to view the terrain from a certain preset height, as from an observation tower. This height then becomes a key parameter of the model. Note that if too large a value is chosen, then the features become obliterated; from a high-flying airplane, the topography is barely observable.

Current Research Tasks

1. A critical component of the entire project is the algorithm for extracting the visibility region of a viewpoint. Although we already have developed and tested one algorithm for this purpose, we know how we can improve it in several significant aspects.

- a. Implement an efficient method (including the necessary data structures) for sorting triangles according to their visibility precedence with respect to the viewpoint. In particular, determine whether Delaunay triangulation guarantees being able to grow star-shaped regions one triangle at a time (the resulting spatial ordering is of considerable interest in itself).
 - b. Attempt to find heuristics that will take advantage of the fact that the visibility regions of adjacent vertices are usually almost identical. This should have a dramatic impact on the average-case performance of region finding, and is essential for processing large digital terrain models.
 - c. Compare experimentally (using USGS DEMs), and if possible theoretically, the average-case performance of our algorithms with that of the more general hidden-surface methods used in graphics.
 - d. Investigate the performance of algorithms that compute only an approximation to the visibility region by considering a triangular facet either entirely visible or entirely invisible.
2. We shall investigate direct applications of visibility to locating a minimal set of observation points and a maximal set of hiding places.
- a. Develop a data structure suitable for determining the union and intersection of visibility regions to serve as input to available facilities location programs.
 - b. Develop an algorithm to find the minimal set of observation points and the maximal set of hiding places.
 - c. Examine the dependence of the number of observation points/hiding places on the "tower height" parameter for various terrain types.
3. We shall assess applicability of pre-computed visibility maps to line-of-sight communication problems.
- a. Find the minimal number of transmitters for a given distribution of receivers and vice-versa (this is similar, but not identical, to 2b).
 - b. Develop an algorithm to compute a "visibility metric" (i.e., the number of necessary intermediate relay points) between any two surface points. Study the properties of this distance measure.
 - c. Determine the locus of coverage of a communicating party of n members moving from point A to point B.
 - d. Formalize the concept of visibility region to a curve on the surface and use it to compute maximum and minimum visibility paths between two points.

4. We shall study the applicability of visibility methods to orientation and navigation. This is strictly an exploratory venture; we will collaborate with Professor C.N. Shen, who has worked for many years on navigation problems connected with the Mars Rover.

5. We will attempt to extract significant topographic features.

- a. Extract peaks by considering i) the size and connectivity of the visibility regions of vertices relative to what they would be if they were at a lower elevation; ii) the inclusion relationships between the visibility regions of adjacent vertices; iii) the location of vertices relative to ridges.
- b. Extract ridges as the boundaries of multiple regions of visibility.
- c. Extract pits and valleys by considering i) the peaks and ridges on the obverse surface; ii) regions of minimum visibility.
- d. Formulate a data structure that behaves hierarchically in a local neighborhood, (i.e. define the mutual relation of a peak or ridge dominating or being dominated by another peak or ridge) but behaves as a network globally (i.e., it partitions the regions of influence of distant features of the same importance). One possibility is to apply the concept of structured graphs, which has been extensively studied by our colleague De Floriani.
- e. Compare empirically our methods with those obtained by methods based on the generalization of local extrema, i.e., surface-specific points and critical point configuration graphs. Seek the opinion of geographers and cartographers on the usefulness and validity of the features extracted by our methods.

Conclusions

It is argued that the global nature of visibility criteria offers promise of their eventual application to the automatic identification of topographic features. Several other potential applications of visibility models were discussed; the most immediate, determination of a minimal set of points of observation, is equivalent to the generic set-covering problem.

The basic step, determination of visibility regions, is computationally intensive even with a simplified piecewise linear terrain model. However, we can accelerate the computation of the region of visibility of each vertex through heuristic preprocessing and edge-ordering methods. We can also accelerate the selection of the observation points by approximating the visibility region as a set of elemental surface patches that are either completely visible or completely invisible.

The next step is to exploit the smoothness of the terrain relative to the sampling interval to link the computation of the visibility regions of neighboring points. Only with an efficient algorithm can we hope to test our ideas for topographic feature extraction on data representative of actual topographies.

Only slightly more difficult than the optimal location of observation points are a set of problems associated with line-of-sight transmission. We are confident that once we have a good algorithm for the determination of visibility regions, these problems will prove tractable. An interesting aspect is that of multiple-hop transmissions, which leads to the concept of a visibility metric.

Finding the location of an observer by comparing the observed visual horizon with that computed from a stored model is the inverse problem of computing the visibility regions. Under what circumstances can the location of the observer be determined uniquely? Furthermore, there must be more effective methods for determining the location than by computing the horizon from all possible viewpoints and performing a comparison.

Data reduction in digital terrain models has a long history of research and is closely related to the theory of numerical approximation of functions. Least-squares and maximum deviation approaches are common. However, if one wishes to display the resulting approximate model, it seems reasonable to take into consideration aspects related to the visual features of the terrain.

The reduction of a topographic map to a sketch map is, in a sense, the ultimate data compression. However, the extraction of topographic features depends strongly on the definitions adopted for such objects. We will test the conjecture that visibility-based definitions of topographic features exhibit good correspondence with both intuitive notions and with accepted geographic nomenclature.

References

- Anderson, 1982. "Hidden Line Elimination in Projected Grid Surfaces," ACM Trans. Graphics 1, 4, pp. 274-291.
- Burton & Smith, 1982. "Hidden-Line Algorithms for Hyperspace," Siam J. of Comput. 11, pp. 71-80
- Creamer, 1985. "The Upper Klethia Valley: Computer Generated Maps of Site Location," SAA Meeting, Denver, 1985.
- De Floriani, Falcidieno, Pienovi and Nagy, 1984. A Hierarchical Structure for Surface Approximation, Computers and Graphics 8, 2, pp. 182-193, 1984
- De Floriani, Falcidieno, Pienovi and Nagy, 1985 a. "Efficient Selection, Storage, and Retrieval of Irregularly Distributed Elevation Data," Computers and Geosciences 11, 6, pp. 667-673, 1985.
- De Floriani, Falcidieno & Pienovi, 1985 b. "Delaunay-based Representation of Surfaces Defined Over Arbitrarily Shaped Domains," Computer Vision, Graphics and Image Processing 32, pp. 127-140.
- De Floriani, Falcidieno, Pienovi, Allen, Nagy, 1986. "A Visibility-based Model for Terrain Features," Proc. Second Int. Symp. on Spatial Data Handling, Seattle, pp. 235-250, July 1986.
- Delaunay, 1934. "Sur la Sphere Vide," Bull. Acad. Sciences USSR, Cl. Sci. Mat. Nat., pp. 793-800.
- Devai, 1984. "Complexity of Two Dimensional Visibility Computations," MICAD '84, 3, Paris 1984.
- Devai, 1986 a. "Quadratic Bounds for Hidden Line Elimination," Proc. Second Annual Symp. on Computational Geometry, Yorktown Heights, NY, pp.269-275, June 1986.
- Devai, 1986 b. "Expected-time analysis of a Worst-case Optimal Hidden-surface Algorithm," Proc. STRUCAD 86, Paris, October 1986.
- El Gindy & Avis, 1981. "A Linear Algorithm for Computing the Visibility Region of a Polygon from a Point," J. Algorithms 2, pp. 186-197.
- El Gindy, Avis & Toussaint, 1983. "Application of a Two-Dimensional Hidden Line Algorithm to Other Geometrical Problems," Computing 31, pp. 191-202.
- Frank, Palmer & Robinson, 1986. "Formal Methods for the Accurate Definition of Some Fundamental Terms in Physical Geography," Proc. Second Int. Symp. on Spatial Data Handling, Seattle, pp. 583-599.
- Garvey, 1986. "Evidential Reasoning for Land-Use Classification," Proc. Workshop on Analytical Methods in Remote Sensing for Geographic Information Systems, Paris, pp. 171-202.

Gold and Maydell, 1978. "Triangulation and Spatial Ordering in Computer Cartography,": Proc. Can. Cartographic Association Third Annual Meeting, Vancouver, pp. 170-175.

Greider, 1976. "TOPO III: A Fortran Program for Terrain Analysis," Comput. Geosci. 2, pp. 195-209.

Handler & Mirchandani, 1979. Location on Networks: Theory and Algorithms, MIT Press, Cambridge.

Johnston & Rosenfeld, 1975. "Digital Detection of Pits, Peaks, Ridges and Ravines," IEEE Trans. Systems, Man and Cybernetics 5, pp. 672-680.

Kubert, Szabo & Guilieri, 1968. "The Perspective Representation of Functions of Two Variables," J. ACM, April 1968.

Mark, 1978. "Concepts of Data Structures for Digital Terrain Models," Proc. Digital Terrain Model Symposium, Falls Church, VA, Am. Society of Photogrammetry, pp. 24-31.

McKenna, 1986. "Worst-case Optimal Hidden-surface Removal," Report JHU/EECS-86/05, The Johns Hopkins University, Baltimore, MD.

Nachman, 1984. "Two-Dimensional Critical Point Configuration Graphs," IEEE Trans. Pattern Analysis and Mach. Int. 6, 4, pp. 442-450, July 1984.

Nagy & Wagle, 1979. "Geographic Data Processing," ACM Computing Surveys 11, 2, pp. 139-181.

Paton, 1975. "Picture Description Using Legendre Polynomials," Computer Graphics and Image Processing 4, pp. 40-54.

Peucker & Douglas, 1975. "Detection of Surface-Specific Points by Local Parallel Processing of Discrete Terrain Elevation Data," Computer Graphics and Image Processing 4, pp. 375-387.

Peucker, Fowler, Little & Mark, 1978. "The Triangulated Irregular Network," Proc. ASP-ACSM Symp. on DTMs, St. Louis.

Sechrest & Greenberg, 1983. "A Visible Polygon Reconstruction Algorithm," Computer Graphics 17, 3, pp. 65-68.

Sutherland, Sproull & Schumacker, 1974. "A Characterization of Ten Hidden-Surface Algorithms," Computing Surveys 6, 1.

Watson, Laffey & Haralick, 1984. "Topographic Classification of Digital Image Intensity Surfaces Using Generalized Splines and the Discrete Cosine Transformation," Computer Graphics and Image Processing 13, pp. 143-167.

Weiler & Atherton, 1977. "Hidden Surface Elimination Using Polygon Area Sorting," Proc. SIGGraph 1977 pp. 214-222.

Wright, 1973. "A Two-Space Solution to the Hidden Line Problem for Functions of Two Variables," IEEE Trans. Comp. 22 (January)

Cela est d'autant plus valable que $T\Delta f$ est plus grand. A cet égard la figure 2 représente la vraie courbe donnant $|\phi(f)|$ en fonction de f pour les valeurs numériques indiquées page précédente.

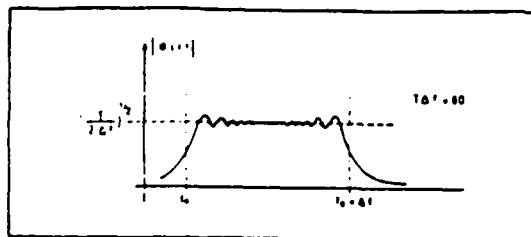


FIG. 2

Dans ce cas, le filtre adapté pourra être constitué, conformément à la figure 3, par la cascade :

— d'un filtre passe-bande de transfert unité pour $f_0 \leq f \leq f_0 + \Delta f$ et de transfert quasi nul pour $f < f_0$ et $f > f_0 + \Delta f$, filtre ne modifiant pas la phase des composants le traversant ;

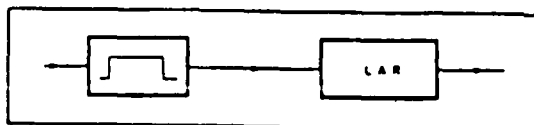


FIG. 3

— filtre suivi d'une ligne à retard (LAR) dispersive ayant un temps de propagation de groupe T_R décroissant linéairement avec la fréquence f suivant l'expression :

$$T_R = T_0 + (f_0 - f) \frac{T}{\Delta f} \quad (\text{avec } T_0 > T)$$

(voir fig. 4).

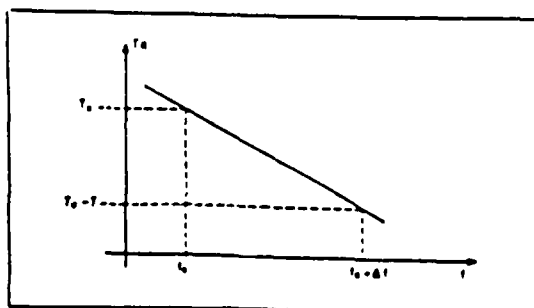


FIG. 4

telle ligne à retard est donnée par

$$\phi = -2\pi \int_0^f T_R df$$

$$\phi = -2\pi \left[T_0 + \frac{f_0 T}{\Delta f} \right] f + \pi \frac{T}{\Delta f} f^2$$

Et cette phase est bien l'opposé de $\phi(f)$.

à un déphasage constant près (sans importance) et à un retard T_0 près (inévitables).

Un signal utile $S(t)$ traversant un tel filtre adapté donne à la sortie (à un retard T_0 près et à un déphasage près de la porteuse) un signal dont la transformée de Fourier est réelle, constante entre f_0 et $f_0 + \Delta f$, et nulle de part et d'autre de f_0 et de $f_0 + \Delta f$, c'est-à-dire un signal de fréquence porteuse $f_0 + \Delta f/2$ et dont l'enveloppe a la forme indiquée à la figure 5, où l'on a représenté simultanément le signal $S(t)$ et le signal $S_1(t)$ correspondant obtenu à la sortie du filtre adapté. On comprend le nom de récepteur à compression d'impulsion donné à ce genre de filtre adapté : la « largeur » (à 3 dB) du signal comprimé étant égale à $1/\Delta f$, le rapport de compression

$$\text{est de } \frac{T}{1/\Delta f} = T\Delta f$$

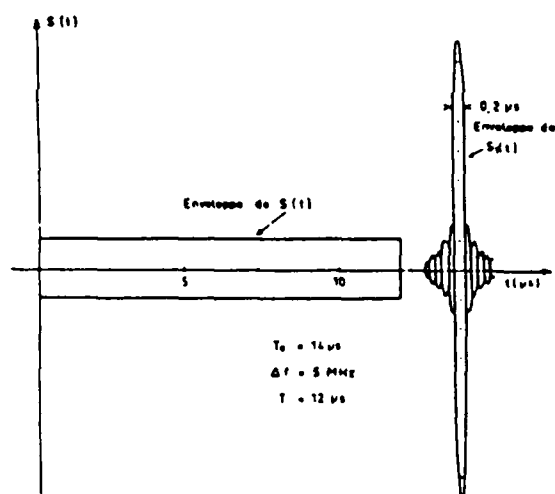
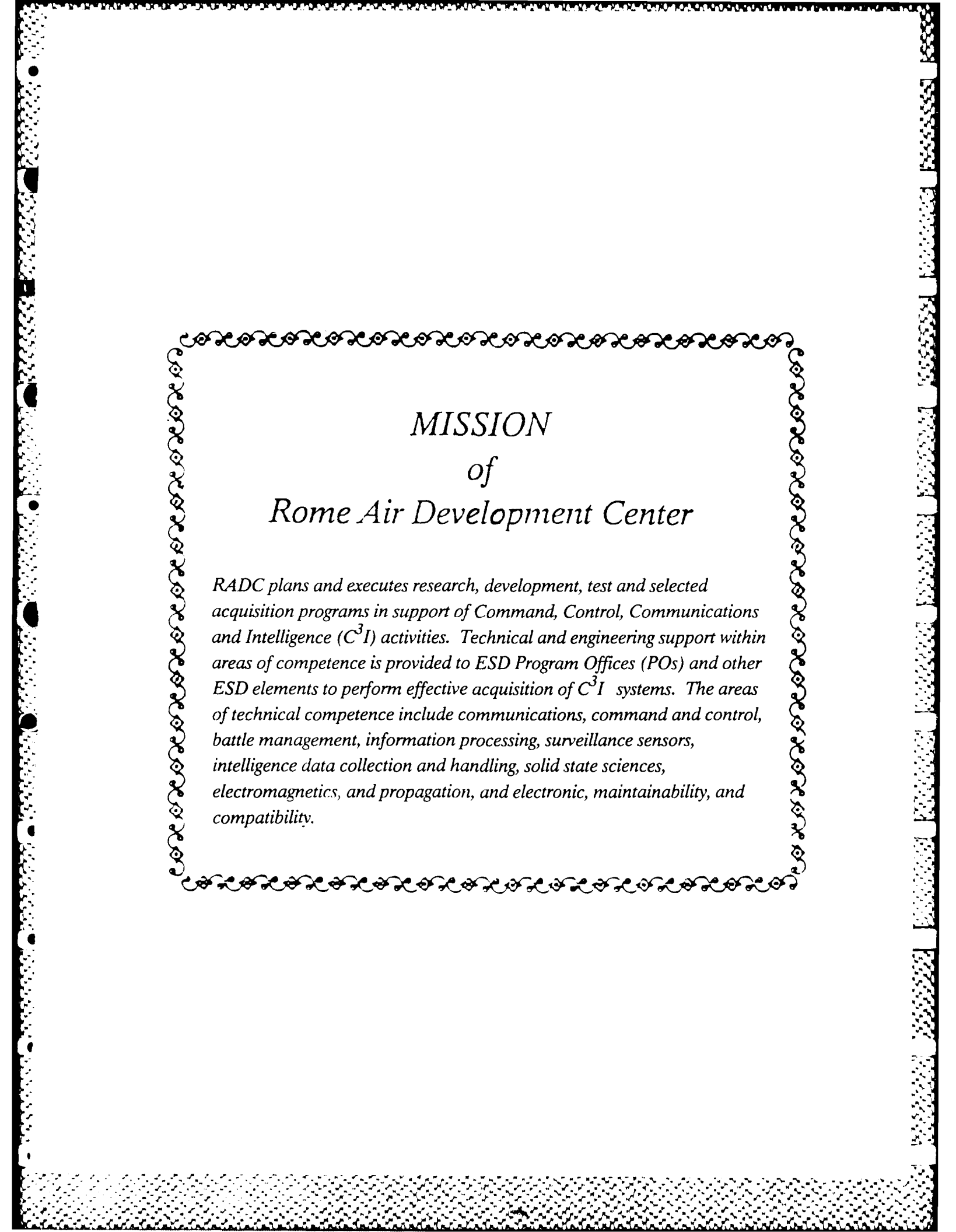


FIG. 5

On saisit physiquement le phénomène de compression en réalisant que lorsque le signal $S(t)$ entre dans la ligne à retard (LAR) la fréquence qui entre la première à l'instant 0 est la fréquence basse f_0 , qui met un temps T_0 pour traverser. La fréquence f entre à l'instant $t = (f - f_0) \frac{T}{\Delta f}$ et elle met un temps

$T_0 - (f - f_0) \frac{T}{\Delta f}$ pour traverser, ce qui la fait ressortir à l'instant T , également. Ainsi donc, le signal $S_1(t)$



MISSION of Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control, Communications and Intelligence (C³I) activities. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C³I systems. The areas of technical competence include communications, command and control, battle management, information processing, surveillance sensors, intelligence data collection and handling, solid state sciences, electromagnetics, and propagation, and electronic, maintainability, and compatibility.

END

DATE

10-88

DTIC