TAX: Prototype Expert System for Terrain Analysis

By Demetre P. Argialas,[1] Associate Member, ASCE, and Ravi Narasimhan[2]

Abstract: Terrain analysis is a time-consuming, costly, and labor-intensive process requiring special skills and training. Furthermore, an enormous amount of remotely sensed data is routinely generated by satellite and airborne sensors which can be used for terrain analysis. Thus, there is an urgent need for an automated approach to analyzing these data and model human reasoning. A rule-based expert system methodology has been developed and the Terrain Analysis Expert (TAX) has been implemented for modeling interpretation logic involved in identifying landforms from aerial images. Knowledge about the geographic location of the image was used to arrive at hypotheses about the landform of the site manifested on the aerial image. These hypotheses were then established or rejected based on the degree of match between the hypothesized landform's pattern elements and those of the site. The site was declared to be the landform with which it had the best match. The pattern elements of the site were obtained interactively from the analyst. A probabilistic method was designed for handling uncertainties in the observed pattern element values and their role in the identification of landforms. The results indicated that a rule-based expert system is appropriate for representing image interpretation logic involved in terrain analysis.

Introduction

Terrain analysis is the systematic study of image patterns relating to the origin, morphologic history, and composition of distinct terrain units, called landforms (Way 1978; Mintzer and Messmore 1984). Among the various approaches to terrain analysis, the landform-pattern element approach has been more prominent in the United States (Way 1978; Mintzer and Messmore 1984). The landform-pattern element approach is based on the premise that soil and rock patterns are repetitive in nature and similar materials create similar terrain patterns, called landforms. Any two landforms derived from the same soil and bedrock, or deposited by a similar process, and existing under the same climatic conditions, exhibit similar physical and visual features on aerial images, called "pattern elements" (Mintzer and Messmore 1984). The pattern elements examined in the landform-pattern element approach include topographic form, drainage pattern type, gully characteristics, soil tone, landcover type, vegetation type, and other special features that may be present.

Terrain analysis takes into account and provides information about physical site factors such as geologic type and structure, soil type and its

[1]Asst. Prof., Dept. of Civ. Engrg. and Remote Sensing and Image Processing Lab., Louisiana State Univ., Baton Rouge, LA 70803.

[2]Grad. Res. Asst., Dept. of Civ. Engrg., Louisiana State Univ., Baton Rouge, LA 70803.

Note. Discussion open until December 1, 1988. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on December 29, 1987. This paper is part of the *Journal of Aerospace Engineering*, Vol. 1, No. 3, July, 1988. ©ASCE, ISSN 0893-1321/88/0003-0151/$1.00 + $.15 per page. Paper No. 22581.

properties, vegetation type, drainage pattern type, and others. This information is used by civil engineers and planners for site development and identifying areas which require ground investigations such as borings and other types of field surveys.

Terrain analysis is both an art and a science. While some researchers have laid down procedures for identifying landforms and their composition, the complexity of the problem is such that there are few instances where clear-cut rules and procedures can be formulated (Way 1978; Lillesand and Kiefer 1979; Mintzer and Messmore 1984; Hoffman 1987). Manual extraction and identification of pattern elements, which is a precursor to landform identification, is tedious and requires expertise. Even when the pattern elements are readily identifiable, an expert interpreter is required to identify the landform comprising the site.

Advancements in artificial intelligence research and the subsequent emergence of expert systems have provided a new powerful tool for the development of computer programs that can capture expertise in many fields and tasks (Winston 1984; Harmon and King 1985).

Knowledge-based expert systems (KBES) are a field of artificial intelligence (Winston 1984; Charniak and McDermott 1985) that emphasize specific, but difficult problem solving requiring expertise (Hayes-Roth et al. 1983). The success of these expert systems is largely determined by the effective representation of domain knowledge (Harmon and King 1985).

The most widely used knowledge representation scheme is the one of rule-based systems (Harmon and King 1985; Jackson 1986). In such a system, the problem solving strategy is represented as sets of rules that will be checked against a collection of facts or knowledge about the current situation. Rule-based knowledge representation centers on the use of *if* ("condition statements") *then* ("action statements"). When the current problem situation satisfies or matches the *if* part of a rule, the action specified by the *then* part of the rule is performed. It is common for the execution of a set of rules to result in a new set of facts which is added to the current list of facts, which trigger other rules. In a rule-based system rules can be employed in a forward or backward chaining method. In forward chaining, rules are matched against facts to establish new facts. In backward chaining, the system starts with what it wants to prove and tries to establish the facts it needs to prove it.

Expert systems are usually employed in domains where facts, rules, and, consequently, conclusions are rarely certain or exact. Inexact reasoning procedures have therefore been developed to complement the knowledge representation and inferencing mechanisms of rule-based systems. Some of the established procedures for handling inexact reasoning that have been demonstrated in the well-known expert systems, such as MYCIN (Shortliffe 1976; Buchanan and Shortliffe 1984), PROSPECTOR (Duda 1980; Reboh 1981), and HYDRO (Reboh et al. 1982), employ heuristic techniques for handling certainties. These heuristic techniques, provide a way for representing uncertainties in facts, in combination of facts, in rules of inference, and in facts supported independently by several rules (Shortliffe 1976; Reboh 1981; Buchanan and Shortliffe 1984; Winston 1984; Charniak and McDermott 1985).

Applications of expert systems in civil engineering have been reviewed by Kostem and Maher (1986), Adeli (1987) and Maher (1987). Expert systems have been successfully employed for representation of knowledge related to interpretation tasks, including interpretation of urban scenes (McKeown 1984), site evaluations for mineral resources (Duda 1980) and military intelligence (Hall and Benz 1985).

Although progress has been made toward the computational interpretation of certain terrain features (Argialas 1986; Argialas et al. 1988), limited computational approaches have been developed to model terrain analysis logic, that is, the problem solving strategy of expert terrain analysts. Mark (1976) demonstrated that the pattern element approach is adaptable to a procedural representation. Leighty (1973) employed a logical approach for terrain pattern recognition and Leighty (1979) has suggested the use of rule-based systems for terrain analysis problem solving.

To further these efforts, an expert system approach was pursued for computational modeling of the terrain analysis problem solving process. The objective of this research effort was the development of the Terrain Analysis Expert (TAX) for landform identification from aerial images. TAX was implemented employing a rule-based production system architecture.

METHODOLOGY

Given an aerial image of a terrain site, the goal of a typical consulting session with the Terrain Analysis Expert (TAX) was to infer the landform type of the site. To limit the scope of the problem, it was assumed that only one landform type existed on that image. The approach followed in identifying the landform of the site was the landform-pattern element approach (Mintzer and Messmore 1984). In this approach, the analyst usually applies hypothesis testing in the following manner. At first he formulates hypotheses about the landforms likely to occur in the physiographic section in which the site is located, by drawing upon his experience and auxiliary information specific to the physiographic section. Then he searches the site characteristics on the aerial images to find a match between the expected pattern elements of one of the hypothesized landforms, as documented in texts and guides, and the observed site characteristics. The analyst continues this procedure, until all the pattern elements are examined. If there is a significant degree of match between the expected and the observed pattern elements, the identity of the landform of the site is established. Otherwise, the next landform in the hypothesis list is investigated for a match.

The basic premise underlying this research effort was that expert systems offer the computational paradigm for representing the terrain analysis problem-solving process described by the pattern-element approach. Specifically, an expert system approach was undertaken because the landform-pattern element approach requires knowledge that is largely empirical, heuristic, and incomplete and computer representation of such knowledge cannot easily be held to rigid and exact descriptions available through procedural languages. Instead, it is greatly facilitated by symbolic representation, symbolic logic, and heuristic search.

Furthermore, the assessment of the landform of a site is a process not easily amenable to rigorous and complete modeling. Instead uncertainties are introduced during problem solving in both the identification of the individual pattern elements and the synthesis of the pattern elements in inferring the landform. For example, the distinction between a dendritic and a rectangular drainage pattern or the difference between a medium and coarse drainage texture is not a clear-cut decision. The difficulty primarily stems from the fact that the terms used for describing terrain pattern elements are qualitative and are not easily amenable to precise quantification. Drainage patterns employ qualitative descriptions such as "gently curving mainstream," "streams radiating like the spokes of a wheel," and others (Argialas 1986; Argialas et al. 1988).

The other type of uncertainty is related to the synthesis of the pattern elements in order to infer a landform. While there are some typical cases, where observation of even a single pattern element causes the terrain analyst to identify unequivocally a landform (for example, the presence of sinkholes indicates a humid-limestone), the majority of landforms display pattern elements which are not unique to that landform. Even if all the pattern elements are unambiguously identified, it might still be difficult to clearly differentiate between two landforms based on a complete and rigid match of pattern elements. An example of such a similarity is between limestone and shale, which have very similar pattern elements in an arid climate.

To handle these two types of uncertainties, certainty values were associated with each pattern element value observed on an aerial image. Moreover, probability values were associated with each pattern element in the models of landforms to express its strength to the identification of a particular landform type. The approach was similar to the one described in PROSPECTOR (Duda 1980; Reboh 1981).

Six landform types have been chosen for focusing the knowledge acquisition process. These types were the humid and arid forms of sandstone, shale, and limestone. The domain knowledge was composed of facts and procedures collected from terrain analysis books (Way 1978; Lillesand and Kiefer 1979), reports (Mintzer and Messmore 1984), the experience of the writers, and an interview with an expert photointerpreter.

Conceptual Models

TAX's knowledge was described with models of terrain-related objects and decision rules pertaining to problem solving in terrain analysis. Models were developed for describing the landform of the site, each pattern element of the site, the relationships among physiographic sections and landforms, and between landforms and their pattern elements. Facts and decision rules with uncertain knowledge sources were identified and methods were developed for their representation.

Models were designed to represent the association between physiographic sections, their expected landform types, and their associated probabilities, based on information derived from physiographic and geomorphologic books and maps (Lobeck 1932; Fenneman 1938). These models were represented as shown in Table 1.

Models of landforms were constructed to describe the relationship between landforms and their expected pattern elements. Such a description for humid sandstone, shale, and limestone is given in Table 2. This description was composed of the expected value of the pattern elements,

TABLE 1. Model of Physiographic Section Employed in TAX

Physiographic section (1)	Landform type (2)	Probability (3)
Cumberland Plateau	Humid sandstone	0.45
Cumberland Plateau	Humid shale	0.45
Cumberland Plateau	Humid limestone	0.10

TABLE 2. Models of Humid Sandstone, Shale, and Limestone Employed in TAX

Pattern element (1)	Pattern element value (2)	P (E/H) Sandstone (3)	Shale (4)	Limestone (5)
Topography	Steep slopes	0.6	0.15	0.5
	Medium slopes	0.2	0.7	0.25
	Flat/undulating	0.2	0.15	0.25
Drainage type	Dendritic	0.6	0.8	0.1
	Rectangular	0.2	0.1	0
	Angular	0.2	0.1	0.1
	Internal	0	0	0.8
Drainage texture	Coarse	0.6	0.1	0.1
	Medium	0.3	0.3	0
	Fine	0.1	0.6	0
	Undefined	0	0	0.9
Soil tone	Light	0.7	0.2	0.3
	Medium	0.2	0.6	0.5
	Dark	0.1	0.2	0.2
Land use valleys	Cultivated	0.3	0.7	0.8
	Forested	0.5	0.1	0.1
	Urban	0.2	0.2	0.1
Land use slopes	Cultivated	0.1	0.1	0.7
	Forested	0.9	0.8	0.2
	Urban	0	0.1	0.1
Gully type	V-shaped	0.8	0.1	0.5
	Sag and swale	0.1	0.8	0
	U-shaped	0.1	0.1	0.5
Gully amount	None	0.3	0	0.8
	Few	0.7	0.2	0.2
	Many	0.0	0.8	0

and an estimation of the degree by which these pattern element values provided evidence in support of that landform. The latter was represented with two probability values: (1) The probability of the occurrence of the pattern element value in that landform, or the probability of the evidence given the hypothesis $P(E/H)$; and (2) the probability of the occurrence of the same pattern element value in all other landforms, or the probability of the evidence given the absence of a hypothesis $P(E/^{-}H)$. When multiple values were possible for the same pattern element all the significant values were represented in the model. Table 2, for example, shows that the possible values of topography for humid sandstone were steep slopes, medium slopes, and flat or undulating, with corresponding $P(E/H)$ values of 0.6, 0.2, and 0.2.

The values of $P(E/H)$ were initially extracted from books and reports. For example, it is known that humid sandstone (SS-h) may exhibit, depending on geomorphologic conditions, either dendritic, rectangular, or angular drainage pattern (Way 1978; Lillesand and Kiefer 1979; Mintzer and Messmore 1984). This was reflected by assigning appropriate $P(E/H)$ values for each of those drainage types of humid sandstone. These values were later refined by the expert. The values of $P(E/^{-}H)$ were computed by taking into account the available physiographic information concerning the list of hypothesized landforms of the site, based on the relations between physiographic sections and landforms. The problem solving strategy was then modeled as rules pertaining to these terrain related concepts or objects.

Formal Reasoning

At the outset, the problem of formulating rules for landform identification seems deceptively simple. A formalism such as the one shown seems adequate.

Rule A:

If	topography is steep slopes;
and	drainage-pattern is dendritic;
and	soil-tone is light;
and	land use is forested;
. . .	. . . .
. . .	. . . .
then	the landform of the site is sandstone.

On closer examination, one finds that landforms do not have such rigid descriptions. For instance, the drainage pattern of sandstones may be dendritic, angular or rectangular. An improvement over this formulation would be to account for alternative values by specifying that the value of a pattern element could be one among a set of values (enclosed in [. . .]).

Rule B:

If	topography is [steep slopes/medium slopes];
and	drainage pattern is [dendritic/rectangular/angular];
and	soil tone is [light/medium];
and	land use is [forested];
. . .	. . . .
. . .	. . . .
then	the landform of the site is sandstone.

This formulation will however not be able to take care of rare cases. If the specification of the pattern element values is made to encompass all possible cases, then more than one landform may match the description of the site. Therefore some kind of a measure is necessary to indicate the level of confidence one has on the assertion of the landform of the site.

In the following formulation, different confidence factors (CF) are assigned to the description of a landform based on different combinations of pattern element values.

Rule C1:

If	topography is [steep slopes];
and	drainage pattern is [angular];

and soil tone is [light];
and land use is [forested];
. . . .
. . . .
then the landform of the site is sandstone, $CF = 90$.

Rule C2:
If topography is [steep slopes];
and drainage pattern is [dendritic];
and soil tone is [light];
and land use is [forested];
. . . .
. . . .
then the landform of the site is sandstone, $CF = 70$.

This formulation still does not take care of the confidence value that the analyst has on the observed pattern element value. For example, for a given site, he might be "definitely certain" that the drainage pattern is dendritic, where for another site he might ascertain that he is "moderately certain" that the drainage pattern is dendritic. To represent this confidence of the analyst on the value of the pattern elements, certainty values are associated to each observed pattern element, such as:

Rule D1:
If topography is [steep slopes], $CF = 80$;
and drainage pattern is [angular], $CF = 85$;
and soil tone is [light], $CF = 20$;
and land use is [forested], $CF = 100$;
. . . .
. . . .
then the landform of the site is sandstone, $CF = 90$.

Rule D2:
If topography is [steep slopes], $CF = 100$;
and drainage pattern is [dendritic], $CF = 100$;
and soil tone is [light], $CF = 100$;
and land use is [forested], $CF = 100$;
. . . .
. . . .
then the landform of the site is sandstone, $CF = 100$.

If one were to enumerate explicitly each of these possible cases, the number of rules one would have to write, to encompass all the cases for all the landforms, would become so large as to make the system infeasible. An alternative method is therefore called for, which would account for the uncertainties in the pattern element values and in the assertion of the landform of the site. In this method, the landform of the site was considered to be the hypothesis (H) and the pattern elements were considered as evidences (E) which strengthen or weaken this hypothesis. Each evidence had associated with it two numbers (LS, LN) which were a measure of how strongly the evidence affected the confidence in the hypothesis:

If E then H (to degree) LS, LN

This means that evidence E suggests the hypothesis H to a degree specified by the certainty factor LS and LN. The number LS indicated how encouraging it was for our belief in the hypothesis to find the evidence present, while LN indicated how discouraging it was to find the evidence absent. The two numbers, LS and LN, specified the sufficiency and the necessity measures, respectively, and were computed from the conditional probabilities [$P(E/H)$ and $P(E/{}^{-}H)$] provided by the expert.

In a more general form, if another number C indicating the confidence in the assertion of the pattern element is employed, the preceding rule takes the form:

Rule E:
- If — the topography of the site is steep slopes, with certainty C (steep slopes);
- and — the current hypothesis for the landform of the site is sandstone, with certainty C (sandstone);
- then — modify the certainty C (sandstone) by calling a certainty computing procedure that will take into account C (steep slopes), C (sandstone), LS, and LN for steep slopes in sandstone.

TAX employed rules like rule E in a backward reasoning procedure to identify a landform. A flow diagram illustrating this problem solving strategy is shown in Fig. 1. A sample of TAX's rules pertaining to the effect of topography of the site on updating the a priori certainty associated with the hypothesis of a landform is presented as follows.

1. Hypothesize a landform type based on physiography.
   - If — there exists a landform type in the knowledge base which occurs in the same physiographic section as the one given by the analyst,
   - then — create an object landform-of-the-site and initialize its probability to the a priori probability of the occurrence of that landform type in that physiographic section.

2. Query site topography from analyst.
   - If — an acceptable certainty value of topography has not as yet been obtained for one of the values of topography appearing in the model of the landform,
   - then — obtain the value of topography of the site and its associated certainty value of topography by querying the analyst.

3. Infer site topography if already there.
   - If — the certainty value of a pattern element has been obtained from the analyst while it was attempting to establish another landform type for the landform-of-the-site,
   - then — the same certainty value of that pattern element is used for the current hypothesis of the landform-of-the-site.

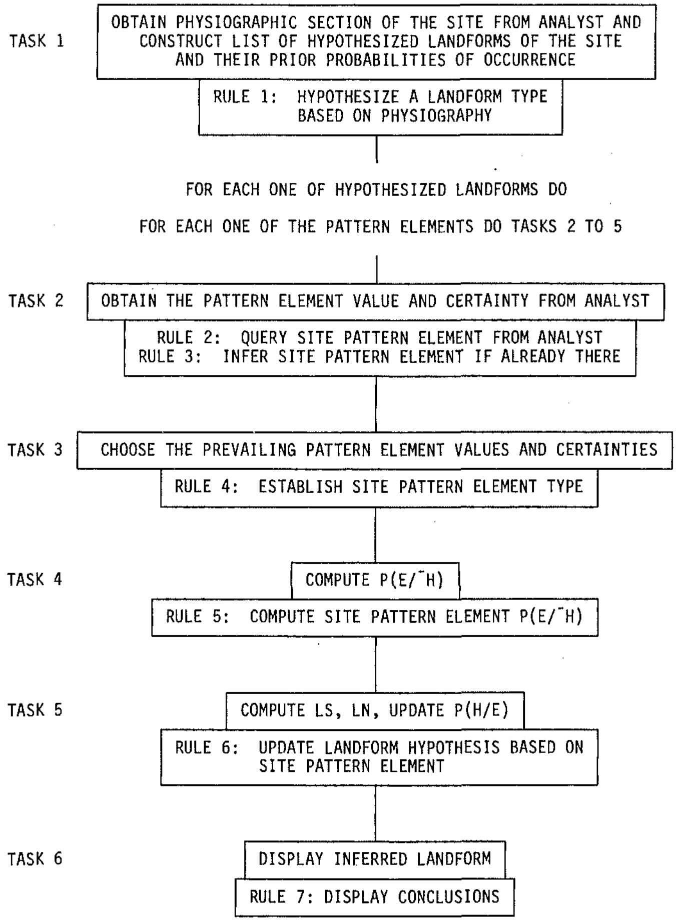

FIG. 1. Flow Diagram Illustrating TAX's Logical Organization

4. Establish site topography type.

If there are multiple instantiations of topography of the site pertaining to the same landform type but with different values for the attributes of topography and certainty value of topography,

then choose the one instantiation with the highest certainty value of topography as being the best and select the corresponding value of topography as the value of the topography of the site.

5. Compute site topography $P(E/{}^{-}H)$.

If there is another landform type belonging to the same physiographic section as the currently established landform type and the value of topography established as the best for the site (rule 4) also occurs in this landform, then compute $P(E/^{-}H)$, where E is the value of topography and H is the currently established landform type.

6. Update hypothesis based on site topography.
   If there is a value of topography for some landform type which best matches the topography of the site, but has not as yet been used for updating the value of probability for that landform type,
   then update the probability of that landform type by calling on an external LISP function "probability compute."

7. Display conclusions.
   If there is a landform type whose probability value is greater than the probability value of the other landform types,
   then display the landform type as the landform-of-the-site and display the probability value associated with the conclusions.

At first, the a priori certainty associated with the hypothesis of a landform was estimated from information related to the physiography of the site. For instance, the landforms that are likely to occur in the Cumberland Plateau physiographic section are sandstone, shale, and limestone with approximate probabilities of occurrence 0.45, 0.45, and 0.10, respectively (Table 1) (Lobeck 1932; Fenneman 1938). The a priori certainty of each hypothesized landform was therefore initialized to the probability of the occurrence of the landform in that physiographic section (rule 1). TAX then chose the landforms in this hypotheses list, one by one, and attempted to establish each one of them, by matching the pattern elements of the site with the models of the landform.

This matching of the pattern elements of the site to the pattern elements of each of the hypothesized landforms took place by first querying the analyst for a certainty value (between -3 and 3) for each pattern element value (rule 2). The sign of the certainty value indicated the presence or absence of the pattern element value, and its magnitude implied the level of confidence of the analyst in his assertion. A certainty of -3 indicated that the pattern element was certainly absent, 3 indicated definite presence, and 0 indicated that nothing could be said about the pattern element value.

If the model of a landform contained multiple values for a pattern element, then TAX queried the analyst for all the values or until a certainty of 2 or more was given by the analyst for a particular pattern element value. In either case, the pattern element value with the highest certainty value was selected (rule 4) and its $P(E/H)$ and $P(E/^{-}H)$ values were computed (rule 5) to modify the a priori certainty associated with the hypothesis of the landform (rule 6). This procedure was repeated for all pattern elements and for all hypothesized landforms. The landform that had the highest a posteriori certainty associated with it was declared to be the landform of the site (rule 7).

Fig. 2 illustrates the procedure for the hypothesis of humid-sandstone (H). The evidences (E) that contribute to this hypothesis are topography, drainage type, drainage texture, and others. The landform model may contain multiple values for a pattern element; for instance the drainage

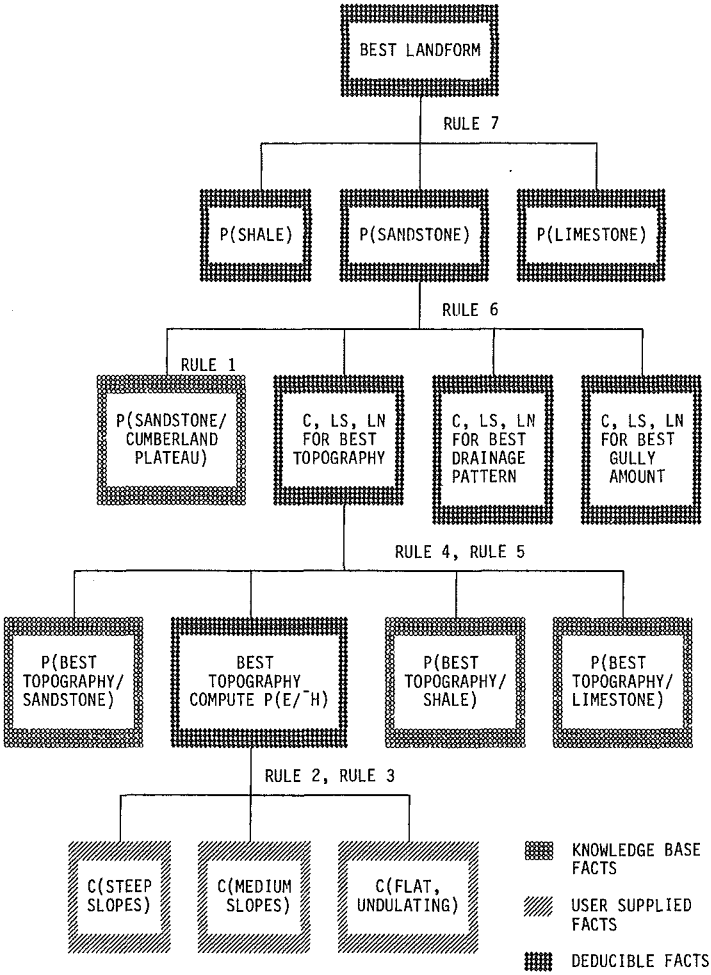

FIG. 2. Portion of TAX's Inference Network Showing Decision Making Under the Hypothesis of Humid Sandstone

could be coarse or medium textured. However, only the value for which the analyst gives the highest certainty will be used to update the certainty associated with the hypothesis.

The a priori certainty associated with the hypothesis of sandstone, $C(H)$, is first transformed to certain ratio, $CR(H)$

$$CR(H) = \frac{C(H)}{1 - C(H)} \quad (1)$$

The $P(E/^{-}H)$ value for each landform-pattern-element pair was computed by summing the probabilities of observing the pattern element in each of the other landforms in the hypothesis list, and dividing the sum by the cumulative probability of occurrence of the other landforms in that physiographic section

$$P(E/{}^{-}H) = \frac{\sum_{i=1}^{i=n} P(E/H_i) * P(H_i)}{\sum_{i=1}^{i=n} P(H_i)} \quad \text{(2)}$$

If the evidence was definitely present, the a priori certainty ratio was multiplied by the sufficiency measure (LS) of the evidence to obtain the a posteriori certainty ratio, $CR(H/E)$

$$CR(H/E) = CR(H) * LS \quad \text{(3)}$$

where LS = the measure of how encouraging it was to find an evidence in establishing the hypothesis and it was computed as

$$LS = \frac{P(E/H)}{P(E/{}^{-}H)} \quad \text{(4)}$$

If the evidence was definitely absent, the a priori certainty ratio was multiplied by the necessity measure (LN) to obtain the a posteriori certainty ratio. LN was the measure of how discouraging it was to find that an evidence was absent and it was computed as

$$LN = \frac{1 - P(E/H)}{1 - P(E/{}^{-}H)} \quad \text{(5)}$$

However, in case of uncertainty in the evidence, expressed as E', the resulting certainty was interpolated between the two cases of perfect certainty using a piecewise linear function (Duda 1980; Reboh 1981)

$$C(H/E') = C(H) + \frac{C(H) - C(H/{}^{-}E)}{3} * C(E) \qquad \text{if } C(E) < 0 \quad \text{(6)}$$

$$C(H/E') = C(H) + \frac{C(H/E) - C(H)}{3} * C(E) \qquad \text{if } C(E) \geq 0 \quad \text{(7)}$$

where $C(H/E')$ = the a posteriori certainty of the hypothesis H given an evidence E' with certainty $C(E')$; $C(H/E)$ = the a posteriori certainty of the hypothesis H given the definite presence of evidence E; and $C(H/{}^{-}E)$ was the a posteriori certainty of the hypothesis H given the definite absence of evidence E.

Fig. 3 illustrates the effect of the evidence, topography, on the certainty associated with the hypothesis, that the landform is humid-sandstone. The a priori certainty associated with this hypothesis is 0.45 (on a scale of 0–1). The three possible values for topography of humid-sandstone, (i.e., steep slopes, medium slopes and flat undulating) and the analyst given certainties for these pattern element values are shown in Fig. 4 and Table 3. Since the certainty associated with steep slopes dominated, its $P(E/H)$ and $P(E/{}^{-}H)$ were used for modifying the a priori certainty of the humid-sandstone hypothesis as follows:

J. Aerosp. Eng., 1988, 1(3): 151-170

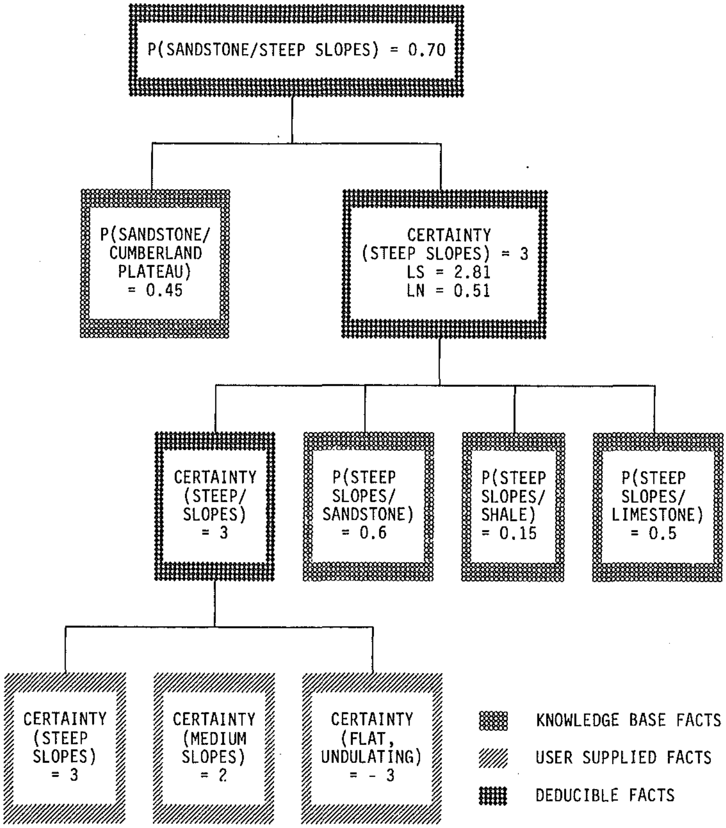

FIG. 3. Portion of TAX's Inference Network Illustrating Effect of Topography on Certainty of Hypothesis of Humid Sandstone

$$P(\text{steep slopes/not sandstone}) = [P(\text{steep slopes/shale}) * P(\text{shale}) + P(\text{steep slopes/limestone}) * P(\text{limestone})]/[P(\text{shale}) + P(\text{limestone})] = \frac{0.15 * 0.45 + 0.5 * 0.1}{(0.45 + 0.1)} = 0.21 \quad (8)$$

TABLE 3. Results of TAX's Evaluation of Hypothesis of Humid-Sandstone with Certainty Values (C) Supplied by Analyst 1

Pattern element (1)	Value (2)	C (3)	$P(E/H)$ (4)	$P(E/^{-}H)$ (5)	LS (6)	LN (7)	A priori certainty (8)	A posteriori certainty (9)
Topography	Steep slopes	3	0.6	0.21	2.8	0.51	0.45	0.70
Drainage type	Angular	2	0.2	0.1	2.0	0.89	0.70	0.78
Drainage texture	Coarse	3	0.6	0.1	6.0	0.44	0.78	0.95
Soil tone	Medium	1	0.2	0.58	0.34	1.91	0.95	0.93
Land use slopes	Forested	3	0.9	0.69	1.3	0.32	0.93	0.94
Land use valleys	Forested	3	0.5	0.1	5.0	0.55	0.94	0.99
Gully type	V-shaped	3	0.8	0.17	4.63	0.24	0.99	0.99
Gully amount	Few	1	0.7	0.2	3.5	0.38	0.99	0.99

J. Aerosp. Eng., 1988, 1(3): 151-170

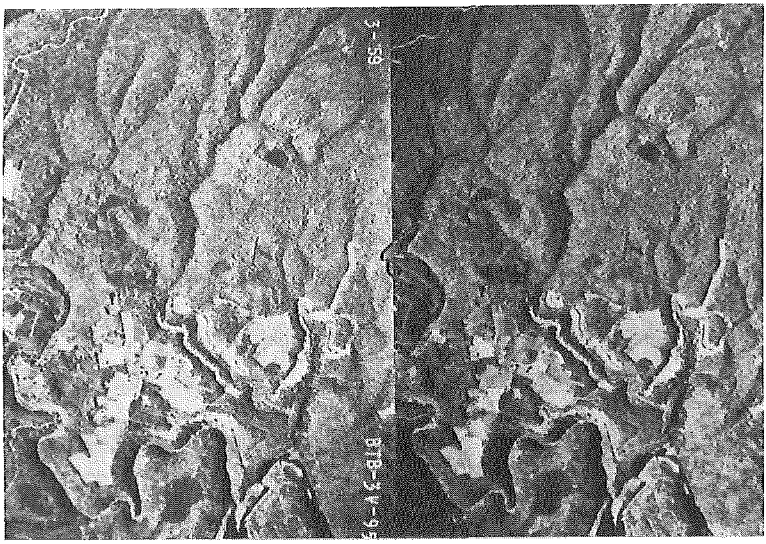

FIG. 4. Stereopair BTB-3V-95/96 Showing Terrain of Humid Sandstone

$$LS = \frac{P(\text{steep slopes/sandstone})}{P(\text{steep slopes/not sandstone})} = \frac{0.6}{0.21} = 2.8 \quad (9)$$

$$CR(\text{sandstone}) = \frac{C(\text{sandstone})}{1 - C(\text{sandstone})} = \frac{0.45}{0.55} = 0.82 \quad (10)$$

$$CR(\text{sandstone/steep slopes}) = 0.82 * 2.8 = 2.3 \quad [\text{for } C(\text{steep slopes}) = 3] \quad (11)$$

$$C(\text{sandstone/steep slopes}) = 0.70 \quad [\text{for } C(\text{steep slopes}) = 3] \quad (12)$$

Testing and Evaluation

The production system language OPS5 was selected and the models and decision rules were formally represented. Landform-related facts were represented as OPS5 objects and the problem solving strategy was represented through production rules (*If-Then*) pertaining to these objects (Argialas and Narasimhan 1988). This version of TAX has been implemented on a VAX 11/780 minicomputer. The OPS5 language is also available for microcomputers. Other expert system tools such as Insight 2+, Knowledge Engineering System (KES), Knowledge Engineering Environment (KEE), and Intelligence Compiler (IC) are available for micro-, mini-, and mainframe computer systems (Maher 1987).

The way chosen for testing TAX for the consistency and accuracy of the embedded knowledge with that of "experts," and the accuracy and correctness of its conclusions was to ask different analysts to interact with TAX, and then compare the different evaluations of the site. This testing has also provided for an evaluation of how TAX reconciles differences in opinion and differences in certainty values provided by the analysts.

The test site is shown on the stereopair of Fig. 4. These aerial images were obtained from the Agricultural Stabilization and Conservation Ser-

TABLE 4. Results of TAX's Evaluation of Hypotheses of Humid Sandstone, Shale, and Limestone

Analyst (1)	A Posteriori Certainty Sandstone (2)	Shale (3)	Limestone (4)
1	0.99	0.001	0.002
2	0.99	0.01	0.007
3	0.93	0.01	0.133
4	0.99	0.002	0.005
5	0.99	0.02	0.001
Best case	0.99	0.0004	0.002
Worst case	0.91	0.15	0.002
1 (equal a priori probabilities)	0.99	0.002	0.01

vice, Aerial Photographic Division. The code of the stereopair was BTB-3V-95/96 (1959). It represents flat-lying upland sandstone in humid climate (Mintzer and Messmore 1984). The site is part of the Cumberland Plateau Section of the Appalachian Plateau. It is characterized by topography of sharp-sloped ridges, and jointed, incised valleys. The drainage pattern texture is coarse, and the type is modified dendritic to partly angular. The land use is forested rangeland.

Five analysts were asked to provide TAX with certainties for the pattern element values of the test site. Analyst 1 has had years of experience in terrain analysis; he was the most experienced photointerpreter participating in this experiment. Analyst 2 was the senior writer of this paper who has had extensive experience in terrain analysis. Analyst 3 was the junior writer of the paper with graduate-level experience in terrain analysis. Analysts 4 and 5 have had graduate courses and moderate experience in terrain analysis. All analysts agreed that the landform of the site was humid sandstone.

The input required during execution of the program was composed of the physiographic section and of the values and certainties of the pattern elements concerning the test site, as those were interpreted by the analysts from the stereopair of aerial images.

Table 4 shows the results of TAX's evaluation of the site for the various cases. By comparing the results of the five analysts, the following have been observed: (1) TAX inferred the actual landform of the site (i.e., sandstone) for all analysts; (2) the a posteriori certainties of the identified landform were the same (0.99) for all analysts, except analyst 3 (0.93); and (3) TAX computed insignificant certainties for the hypotheses of humid shale and limestone, as expected, since these were not representing the landform of the site.

For further evaluation of the system, a sensitivity analysis was performed by having TAX repeat the evaluation two more times with hypothetical analysts. The first evaluation was done by considering the best values of certainty supporting the hypothesis of humid sandstone for all pattern element values, among all the values chosen by the five analysts. The second evaluation was done by considering the worst values of certainty supporting the hypothesis of humid sandstone for all pattern element values, among all the values chosen by the five analysts. Thus,

J. Aerosp. Eng., 1988, 1(3): 151-170

TABLE 5. Results of TAX's Evaluation of Hypothesis of Humid-Shale with Certainty Values (C) Supplied by Analyst 1

Pattern element (1)	Value (2)	C (3)	$P(E/H)$ (4)	$P(E/^{-}H)$ (5)	LS (6)	LN (7)	A priori certainty (8)	A posteriori certainty (9)
Topography	Steep slopes	3	0.15	0.58	0.26	2.03	0.45	0.17
Drainage type	Angular	2	0.1	0.18	0.55	1.1	0.17	0.13
Drainage texture	Coarse	3	0.1	0.51	0.2	1.83	0.13	0.03
Soil tone	Medium	1	0.6	0.25	2.36	0.54	0.03	0.04
Land use slopes	Forested	3	0.8	0.77	1.03	0.88	0.04	0.04
Land use valleys	Forested	3	0.1	0.43	0.23	1.57	0.04	0.01
Gully type	V-shaped	3	0.1	0.75	0.13	3.54	0.01	0.001
Gully amount	Few	1	0.2	0.61	0.33	2.05	0.001	0.001

upper (0.99) and lower (0.91) bounds were computed for the certainty value of the hypothesis of the humid sandstone. This comparison has revealed that, even in the worst case, TAX correctly evaluated the landform of the site. However, even after all eight pattern elements were accounted for, the a posteriori certainty of the hypothesis did not rise above 0.91. The certainty of humid shale had slightly increased, but it was still insignificant (0.15).

A third evaluation was made, by assuming that the prior certainties concerning the occurrence of the three hypothesized landforms (sandstone, shale, and limestone) were equal (that is, each was equal to 0.33). This evaluation was made to test the effect of the a priori probabilities of the hypothesized landforms of the site, as these were extracted from the physiographic information, on the a posteriori certainty of the landform of the site. This comparison has revealed that TAX still concluded in support of humid sandstone with no increase in the certainties of the other two landforms (Table 4). This implies that information regarding the relative distribution of the hypothesized landforms, based on physiographic information, may not be critical to the final conclusion.

The following are interpretations concerning the relation among the values of $P(E/H)$, $P(E/^{-}H)$, LS, LN, a priori, and a posteriori certainties derived from the results appearing in Tables 3 and 5.

LS values greater than 1 implied that the observed pattern element value favors the hypothesis of that landform. Hence, the a posteriori certainty of the landform is increased over the a priori certainty value. For example, the LS value for coarse drainage texture in humid sandstone was large (6.0) because the $P(E/H)$ value (0.6) was much higher than the $P(E/^{-}H)$ value (0.1) (Table 3). Higher $P(E/H)$ value implied that in most sandstones it has been observed that the drainage texture is coarse. Lower $P(E/^{-}H)$ value implied that a coarse drainage texture is not commonly found in the other landforms occurring in the same physiographic section (e.g., shale and limestone). Since the drainage texture of the test site has been interpreted as coarse, with certainty +3, and since its LS value was large (6.0), the certainty of the hypothesis of humid sandstone was increased from its a priori value of 0.78 to its a posteriori value of 0.95 (Table 3). Similar observations can be made for the rest of the pattern elements of Table 3 having relatively large LS, and small LN values (e.g., steep slopes, angular drainage pattern, etc.).

LS values less than 1 implied that the observed pattern element value was unfavorable for the hypothesis of that landform. Hence, the a posteriori certainty of the landform was decreased in relation to the a priori certainty value. For example, the LS value for medium soil tone in humid sandstone was low (0.34), because the $P(E/H)$ value (0.2) was much lower than the $P(E/^{-}H)$ value (0.58) (Table 3). The low $P(E/H)$ value implied that in most sandstones it has been observed that medium soil tone is not commonly found where the light soil tones are commonly found. This has been expressed in the landform models by the larger $P(E/H)$ value (0.7) for light soil tones (Table 2). The higher $P(E/^{-}H)$ value (0.58) for medium soil tone implied that a medium soil tone is commonly found in other landforms occurring in the same physiographic section (e.g., shale, limestone). Since the soil tone of the test site has been interpreted as medium with certainty +1, and since its LS value (0.34) was lower than its LN value (1.91), the certainty of the hypothesis of humid sandstone was decreased from its a priori value of 0.95 to its a posteriori value of 0.93 (Table 3).

When LS was approaching 1, the observation of the evidence of a particular pattern element value had no effect on the certainty value of the hypothesis of the current landform. For example, since $LS = 1.03$ for the observation of the evidence of forested land use on slopes for establishing the hypothesis of shale, the a priori and a posteriori certainties were the same (Table 5).

An LS value approaching 0, like the one for V-shaped gullies concerning the hypothesis of shale (0.13) (Table 5), implied that the hypothesis was false when the evidence was true. Indeed, it is highly improbable to find V-shaped gullies in humid shales, as this was indicated by its very low $P(E/H)$ value (0.1) employed in the landform models (Table 2).

Practical Applications

Rule-based expert systems have been demonstrated to be a valuable tool for representing the expert's knowledge relating to the identification of landforms from aerial images. This information about the landform of the site can be used to derive expected engineering properties such as soil type, depth to water table, depth to bedrock, etc. It is anticipated that these engineering attributes could also be represented and modeled using an expert system approach. The elements of such an expert system can further be synthesized to produce information about groundwater resources, engineering site suitability, environmental impact assessment, and others.

Conclusions

A comparison of the present approach to other approaches cannot be carried out, since there are no published results of previous efforts. Comparisons can only be made with the qualitative models.

The qualitative landform-pattern element models suffer mainly from lack of flexibility, and inability to adjust to the specific situations of any particular site. Some pattern elements that may be obvious at one site may not be at another. If not considered properly, they may undermine the entire interpretation. The nature of the data and knowledge in terrain analysis is such that it precludes an exact form of reasoning. Terrain analysis implicitly incorporates elements of probabilities and plausible reasoning, since it often requires the subjective judgment of expert analysts. Because uncertainties in the pattern element values are not accounted for explicitly in the qualitative approach, it is difficult to estimate the confidence of an inference. TAX, however, attempted to remove these limitations of the qualitative models by utilizing plausible reasoning, uncertainty handling procedures and by relying upon its knowledge base to resolve problems of uncertainties, and missing data.

Qualitative descriptions of the pattern elements were represented by associating certainty values to the observation of the pattern element values. These were very similar to the ones employed in qualitative terrain analysis (Mintzer and Messmore 1984).

Inexact models were introduced for describing landforms in relation to their pattern elements by associating probability values to the pattern element values, expressing their strength in support of a hypothesis of a landform.

If no prior or auxiliary information had been employed in TAX, it would have been necessary to test the hypothesis of all possible landforms (perhaps more than 40) for being candidate landforms of the site. By employing a priori information related to the physiography of the site, only three landform hypotheses were formulated for this particular test site. This represents a significant reduction of the search space of the plausible landform hypotheses.

The evaluation performed has shown that a variety of cases can be handled satisfactorily by the prototype system. This result suggests that with increased knowledge and refinement of the reasoning structures, a high performance expert system is attainable.

One of the primary goals in developing TAX was to develop an adequate structure for representing and reasoning about objects in the terrain analysis domain. As such, the structure described here represents a platform on which future work will take place. However, considerable experience with other knowledge representation schemes and more detailed knowledge of the terrain analysis process will be required.

Terrain analysis has been a field where experience has no substitute, making it particularly suitable for the application of knowledge-based experts systems. This study has shown that expert systems can bring to terrain analysis the very real possibility of capturing and accumulating the knowledge of their true experts and making this knowledge available to others in a usable form.

Acknowledgments

The work reported here was supported partially by two Louisiana Department of Transportation Research Center projects (Contracts 87-2GT and 87-1H[B]) and by the Department of Civil Engineering and the Remote Sensing and Image Processing Laboratory at Louisiana State University. The assistance of Dr. Roger K. Seals, Prof. Ara Arman, and Dr. Charles A. Harlow is gratefully acknowledged. The writers thank Olin W. Mintzer, research civil engineer at the Engineering Topographic Laboratories for his help, for he has been our expert. He has kindly offered his estimates for both the probabilities of the knowledge base concerning the landform-pattern element pairs and his estimation of certainty values of the pattern element values of the test site. The assistance of R. Greene and R. Czek in interpreting the test site and providing their certainty values in testing the system is greatly appreciated. We thank Susan Sartwell for the excellent typing of the final text, tables, and figures. We also thank the two reviewers for their constructive comments.

Appendix. References

Adeli, H. (1987). *Expert systems in construction and structural engineering.* Chapman and Hall, London, England.

Argialas, D. (1986). "Computer-assisted recognition of drainage patterns." *Technical Paper*, Annual Convention of American Society for Photogrammetry and Remote Sensing, vol. 4, 435–444.

Argialas, D., Lyon, J., and Mintzer, O. (1988). "Quantitative description and classification of drainage patterns." *Photogrammetric Engrg. and Remote Sensing*, 54(4), (in press).

Argialas, D., and Narasimhan, R. (1988). "A production system model for terrain analysis knowledge representation." *Microcomputers in Civ. Engrg.*, 3(2).

Belcher, D. (1948). "The engineering significance of landforms." *Highway Res. Board Bull.*, 13, 125.

Buchanan, B., and Shortliffe, E., eds. (1984). *Rule-based expert systems: The MYCIN experiments of the Stanford heuristic programming project.* Addison-Wesley, Reading, Mass.

Charniak, E., and McDermott, D. (1985). *Introduction to artificial intelligence.* Addison-Wesley, Reading, Mass.

Duda, R. (1980). "The PROSPECTOR system for mineral exploration." *Final Report*, SRI International, Menlo Park, Calif.

Fenneman, N. (1938). *Physiography of the eastern United States.* McGraw-Hill Book Co., New York, N.Y.

Hall, C., and Benz, S. (1985). "A knowledge-based photointerpretation system." *Nat. Defense*, LXIX(404), 54–62.

Harmon, P., and King, D. (1985). *Expert systems: artificial intelligence in business.* John Wiley & Sons, New York, N.Y.

Hayes-Roth, F., Waterman, D., and Lenat, D. (1983). *Building expert systems.* Addison-Wesley, Reading, Mass.

Hoffman, R. (1987). "The problem of extracting the knowledge of experts. *AI Magazine*, 8(2), 53–67.

Jackson, P. (1986). *Introduction to expert systems.* Addison-Wesley, Reading, Mass.

Kostem, C., and Maher, M., eds. (1986). *Expert systems in civil engineering.* ASCE, New York, N.Y.

Leighty, R. (1979). "Research for information extraction from aerial imagery." *Remote sensing symposium*, U.S. Army Corps of Engineers, Engineer Topographic Laboratory, Reston, Va.

Leighty, R. (1973). "A logical approach toward terrain pattern recognition for engineering purposes," dissertation presented to Ohio State University, Columbus, Ohio, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Lillesand, T., and Kiefer, R. (1979). *Remote sensing and image interpretation.* John Wiley & Sons, New York, N.Y.

Lobeck, A. (1932). *Atlas of American geology.* The Geographical Press, Columbia University, New York, N.Y.

Maher, M., ed. (1987). *Expert systems for civil engineers.* ASCE, New York, N.Y.

Mark, J. (1976). "Computer analysis of photo pattern elements." *Photogrammetric Engrg. and Remote Sensing*, 42(4), 545–556.

McKeown, D. (1984). "Knowledge-based aerial photo interpretation." *Photogrammetria*, 39, 91–93.

Mintzer, O., and Messmore, J. (1984). *Terrain analysis procedural guide for surface configuration.* U.S. Army Corps of Engineers, Engineering Topographic Laboratories, Fort Belvoir, Va.

Reboh, R. (1981). "Knowledge engineering techniques and tools in the prospector

environment." *Technical Note 243*, SRI International, Menlo Park, Calif.
Reboh, R., Reiter, J., and Gaschnig, J. (1982). *Development of a knowledge based interface to a hydrological simulation program*. SRI International, Menlo Park, Calif.
Shortliffe, E. (1976). *Computer-based medical consultations: MYCIN*. Elsevier Publishing Co., New York, N.Y.
Way, D. (1978). *Terrain analysis*. Hutchinson and Ross, Stroudsburg, Pa.
Winston, P. (1984). *Artificial intelligence*. Addison-Wesley, Reading, Mass.

J. Aerosp. Eng., 1988, 1(3): 151-170