PEER-REVIEWED ARTICLE

The Fourth Dimension in Digital Photogrammetry (DP)

A. M. Al-garni

Abstract

Automatic image interpretation is viewed in this paper as a fourth dimension—the knowledge dimension—that should be integrated with the three-dimensional (3D) analytical geometry of land surfaces in digital photogrammetry (DP). The fourth dimension contains quantitative and qualitative knowledge. As a small portion of ongoing research, digital terrain models (DTMs) are the first quantitative portion of knowledge that is integrated with some heuristic knowledge to form a continuously modified Knowledge-Based Image Interpretation System (KBI²S). The system is developed and three image features have been tested on the system, giving an encouraging new dimension to DP.

Introduction

Digital photogrammetry (DP) can be viewed as analytical photogrammetry that has been modified to provide an increase in automation. This is attained by replacing analog input by digital input so that computers can process and manipulate data with less human intervention. Following the steps of analytical photogrammetry, we are dealing in DP with the three-dimensional (3D) analytical contents of images, putting aside the image interpretation field (IIF) as a standalone field with which computers have had very little to do. In this paper IIF is considered as a new dimension that must be integrated with the conventional analytical operations of DP so that computers process some relevant knowledge about image features. A conceptual model that interprets image features based on DTMs and other visual items of knowledge that can be overlaid on DTMs of areas of interest is presented.

As of today not much effort has been devoted to introducing artificial intelligence (AI) into digital photogrammetry (DP) (Al-garni, 1992; Argialas *et al.*, 1989). Even recent publications concerned with digital stations, digital photogrammetric processing (Gruen, 1989), and softcopy photogrammetric workstations (Dowman *et al.*, 1992; Skalet *et al.*, 1992; Miller *et al.*, 1992; Schenk *et al.*, 1992) have placed little emphases on the knowledge dimension and image interpretation. In a global view, the pyramid of DP can be classified into three major phases according to the operations conducted on digital images. In order, these phases are early vision, intermediate vision (stereopsis), and late vision. Extensive research concerning the early and intermediate vision phases has been published (Schenk, 1992). However, the late vision phase has received less attention in the field of DP. This ignorance is dangerous because the early and intermediate vision phases could be re-evaluated when we, as photogrammetrists, come to integrate and base the late vision phase on the early and intermediate phases. One should keep the ultimate objective of the photogrammetric operations in mind while conducting academic research in any of the phases of the DP pyramid.

The late vision process in DP consists of different layers of knowledge that lead to explicit information about image features. This process comprises the main core of image interpretation. Image interpretation, from the viewpoint of digital photogrammetry, is an integrated process that combines and integrates the processes of the early, intermediate, and late vision stages. That is, an image interpretation model should be developed to smoothly combine the analytical aspects of digital photogrammetry and the heuristic aspects of image interpretation.

In this paper a conceptual theme of an optimal interpretation model which combines the analytical and heuristic characteristics of aerial digital images is presented. This is accomplished through tree-like image analysis nodes that are activated or deactivated by a specialized inference mechanism. The model consists of a global image analyzer that has local knowledge classifiers.

Development of the Image Interpretation Model

Image interpretation has developed as a standalone field that has been well established as a manual process. This field has a distinguishing character that must be understood; that is, a human interpreter gains his skills after years and years of experience. Accordingly, human expertise is considered as the real core of image interpretation.

In order for a machine to duplicate human expertise in image interpretation, there are two main conditions that have to be maintained. First, images have to be prepared in a form that a computer can deal with and manipulate. That is, analog photographs must be converted into digital form. Second, human expertise has to be acquired and formulated according to specific rules so that knowledge coding can be conducted. Naturally, the two conditions are preserved in image interpretation. Accordingly, a conceptual interpretation model has been developed by this research.

In this study, different layers or templates are developed to perform quantitative and qualitative image analysis on a region level. Figure 1 shows *n* templates with *m* regions. The templates represent *n* knowledge bases that are merged by the expert system to create one template with unified explicit information (called the ID template). That is, the model has been developed to deal with image regions rather than image pixels because the ultimate objective of image interpretation

Photogrammetric Engineering & Remote Sensing,
Vol. 61, No. 1, January 1995, pp. 57–62.

0099-1112/95/6101-57$3.00/0

Civil Engineering Department, College of Engineering, King Suad University, P.O. Box 800, Riyadh 11421, Saudi Arabia.

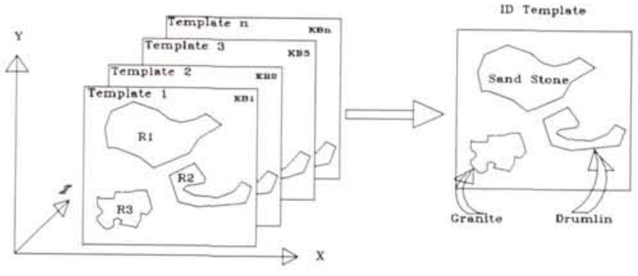

Figure 1. Template overlays and evidence accumulation.

is explicit information (obtained from regions) as opposed to implicit information (embedded in individual image pixels).

The processed information (**PI**), as accumulative evidence for a region (*Ri*), is a function of the number of independent templates of knowledge bases (*f(KBT)*) (as shown in Figure 1) times the specific terrain analysis value (T_m). The interpretation model of the KBI²S is expressed in the following mathematical form:

$$\xi = (\Theta_{Ri})_{max}\omega \tag{1}$$

$$\Theta_{Ri} = \Sigma\theta_{ii} \tag{2}$$

where ξ is the identity of an image feature and $\omega \in \psi$. All features that the expert system can identify are contained in ψ as a permanent knowledge base in the system. Also, Θ_{Ri} is the summation of the diagonal elements (θ_{ii}) of the matrix **PI**. Equations 1 and 2 are further analyzed as follows:

$$[\mathbf{PI}]_{Ri} = T_m f(\,KBT) \tag{3}$$

$$f\,(KBT)_{Ri} = [F_i][\eta_m] \tag{4}$$

From Equations 3 and 4 we obtain

$$[\mathbf{PI}]_{m \times m} = [\mathbf{T}_m]_{m \times 1}[\mathbf{F}_i]_{1 \times m}[\eta_m]_{m \times m} \tag{5}$$

with the condition that $T_{mi} \times F_i = 1$ and $T_{mi}F_j = 0$; [**PI**] is a diagonal knowledge matrix containing heuristic and analytical information. That is, in matrix form

$$[\mathbf{PI}]_{Ri} = \begin{pmatrix} T_{m1} \\ T_{m2} \\ \cdot \\ \cdot \\ \cdot \\ T_{mm} \end{pmatrix} (F_1\ F_2\ \ldots\ F_m) \begin{pmatrix} n_{11} & 0 & 0 & \ldots & 0 \\ 0 & n_{22} & 0 & \ldots & 0 \\ 0 & 0 & n_{33} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & n_{mm} \end{pmatrix}$$

The symbol η_{ii} stands for probability values associated with $T_{mi} \times F_i$ which are obtained through adequate observations of an acknowledged expert in the field while conducting practical image interpretation (Al-garni, 1992). Likewise, T_{mi} is a global terrain analysis attribute that has been recognized by experts to be a distinguishing criteria in image interpretation that could add to evidence accumulation. Finally, F_i is a tree-like node containing values that should be attributed to every T_{mi} element for a specific image feature. The developed expert system has been implemented to contain three main types of modules and to manipulate their contents. Each of these modules deals with a specific task:

- A registration module for geometrical and coding purposes;
- A height information model (or digital terrain model (DTM)) for topographic analyses purposes; and
- A heuristic and analytical module that is initiated by the inference mechanism as required by any of the modules above. Also, it deals with all relevant expertise and rules of thumb of experts.

The rest of this paper highlights these modules and their AI mechanisms in KBI²S.

Registration Module and Its Heuristic Agent

The image interpretation model is developed in such a way that it allows different spatial data and heuristic information to be combined to form an integrated knowledge base for relevant delineated regions. This objective (creating independent, but communicating, knowledge bases) can be achieved if all relevant regions are treated in object space. In this way, heuristic information, as well as any other kind of image data for particular pixels, is addressable in terms of object space as a common space for a variety of information sources.

Two important steps must be conducted so that a common space is established. The first step involves a process of georeferencing image regions to an object coordinate system (e.g., registering images to maps). In this step relevant information that concerns each region is registered to the corresponding object template. Accordingly, an integrated knowledge base that consists of image data, heuristic information, geophysical data, and other sensor data is developed through a georeferencing operation. The second step involves what I called the *geocoding operation*. This operation can be viewed differently according to the object space template. For instance, if pixels of digital images that form homogeneous regions are addressable in terms of a map coordinate system, then the geocoding system here is referenced to maps using map terminologies for knowledge coding.

Accordingly, we have two different coordinate systems that should be related to one another. The two systems are the image coordinate system $f(x,y)$ and the template (e.g., object) coordinate system $g(u,v)$. Here f and g form a pair of mapping functions that can relate the two coordinate systems (Richards, 1986) as follows:

$$u = f(x,y) \tag{6}$$

$$v = g(x,y) \tag{7}$$

The definition of the two functions (f,g) depends on the type of sensors which scan the images. However, both share the usage of certain types of control points. For instance, CCD cameras that scan aerial photographs can utilize the well established methods of transformation found in analytical photogrammetry, provided that a calibration process for the camera is accomplished.

KBI²S contains three methods for registration purposes, each of which is initiated by the inference engine based on several criteria. These registration methods vary based on the type of input of digital images. For instance, the following collinearity condition equations (found in basic standard texts of photogrammetry such as Moffitt *et al.* (1980)) may be selected by the expert system if the input of digital images is scanned aerial photographs:

$$[\mathbf{X\ Y\ Z}]^{\mathrm{T}} = k\mathbf{M}[\mathbf{x\ y\ z}]^{\mathrm{T}}; \tag{8}$$

$$[X\ Y\ Z]_i^{\mathrm{T}} = [(X_i - X_c)\ (Y_i - Y_c)\ (Z_i - Z_c)]_i^{\mathrm{T}}; \tag{9}$$

$$[x\ y\ z]_i^{\mathrm{T}} = [(x_i - x_o)\ (y_i - y_o)\ (-f)]^{\mathrm{T}} \tag{10}$$

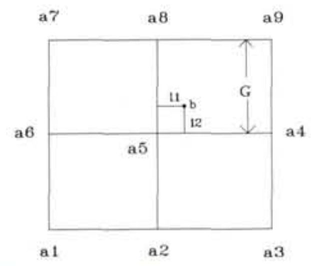

Figure 2. DTM interpolation methods.

$$x_i = -f\frac{m_{11}X+m_{12}Y+m_{13}Z}{m_{31}X+m_{32}Y+m_{33}Z} \quad (11)$$

$$y_i = -f\frac{m_{21}X+m_{22}Y+m_{23}Z}{m_{31}X+m_{32}Y+m_{33}Z} \quad (12)$$

where **M** is 3 by 3 rotation matrix; X_c, Y_c, and Z_c are the coordinates of the exposure station; X_i, Y_i, and Z_i are the coordinates of an object point i; x_i and y_i are the photo-coordinates of the point; and x_o, y_o, and f are the interior elements of the camera. The collinearity condition may be replaced by the projectivity condition (Equations 13 and 14) if certain advantages of the latter are realized by the expert system: i.e.,

$$x_2 = \frac{a_1x_1 + b_1y_1 + c_1}{a_ox_1 + b_oy_1 + 1} \quad (13)$$

$$y_2 = \frac{a_2x_1 + b_2y_1 + c_2}{a_ox_1 + b_oy_1 + 1} \quad (14)$$

where x_2, y_2 and x_1, y_1 are two coordinate systems based on which the transformation coefficients (a_o, b_o, a_1, b_1, c_1, a_2, b_2, and c_2) are computed.

In some other cases, defined by the heuristic agent of the system, mapping polynomials and ground control points are used for registration purposes. For example, a Landsat image (MSS or TM) can be registered to a map using the following general polynomial transformation forms:

$$u = k_o + k_1x + k_2y + k_3xy + k_4x^2 + k_5y^2 + \cdots \quad (15)$$

$$v = c_o + c_1x + c_2y + c_3xy + c_4x^2 + c_5y^2 + \cdots \quad (16)$$

provided that the minimum requirements of scale and control points are preserved (map scale and image scale correspond to a scale that is considered useful for images where proper information can be extracted). Again, the inference engine mechanism of the KBI²S decides on the degree of the polynomial based on several criteria such as density and distribution of the ground control points in the plane of the image. Moreover, complexity (rigidness) of the topographic forms of a land surface constitutes another criteria that affects the decision of the expert system regarding the order (degree) of the polynomial used.

DTM Module Directed by a Heuristic Agent

From the viewpoint of intelligence in DP, a DTM can play a great role in image interpretation. The developed DTM module comprises an important layer of knowledge with flexible accommodation of proper DTM algorithms. The flexibility of this module is supervised by a proper heuristic layer of knowledge. That is, regions of homogeneous characteristics are pre-delineated. The classifiers, whether they be heuristic or analytical, lead to different groups of classes recognized on the image. Each group has its regional seeds that are assessed geometrically from the viewpoint of planimetric and vertical information.

Geometrical analyses of image features are essential to quantitative terrain analysis where many features and ground forms possess very distinguishing topographic geometry. For instance, drumlins, volcanic land forms, and eskers have unique geometric and topographic characteristics. Moreover, the size of image features is an important clue to feature identities. Also, a DTM reveals the forms of drainage systems based on which soil properties can be analyzed (Morris, 1991).

While in the registration mode of regions, maximum geometrical dimensions of different image features are computed. Moreover, point levels are estimated by proper DTM algorithms for the same regions. The geometry of regions may affect which algorithm should be used for DTM computations. Interpolation methods for DTM computations can be as simple as a uniform grid system method. On the other hand, elevation information can be computed by triangular facet, random point, contour line, or spline methods.

The decision regarding proper selection of DTM methods is initiated by specialist rules created in the heuristic module of the KBI²S. The decision is made based on two main factors: topography and economy. For instance, topographic forms of regions, such as grid-like flat regions, irregular regions, and semi-regular regions, imply which algorithm should be used. Moreover, an optimal decision will consider the two-fold economic issue of the interpretation process. That is, a close analysis of the trade-off between required accuracies and the computational burden should be considered. For flat regular regions, uniform grid system methods can be used (see Figure 2). In this case, the mathematical model is formed to compute the elevation of a point located at (x,y) in the image. If points a_1, a_2, a_3, ..., a_9 have known elevations (z_i) and we need to compute the elevation z_b (Murchison, 1977) of point b at known planimetric position (x_b,y_b), then

$$z_b = z_5(1 - \frac{l_1}{G}) + z_4\frac{l_2}{G}(1 - \frac{l_1}{G}) + z_8\frac{l_1}{G}(1 - \frac{l_2}{G}) + z_9\,\frac{l_2}{G}\,(\frac{l_1}{G}) \quad (17)$$

In cases where at least three points surrounding the unknown point can be located, the triangular facet method can be used to acquire the elevation information. In this method, measured coordinates of selected points form vertices of varying triangular planes. It is the place of the inference mechanism of the system to define criteria for proper selection so that false levels are avoided. Mainly, break-lines that give slope changes in topography imply preferable point locations. In Figure 3, (x_b,y_b) are the coordinates of point b that

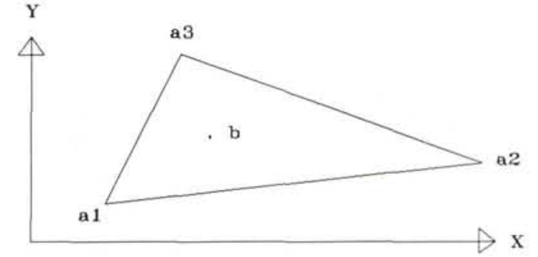

Figure 3. Triangular facet method for DTM interpolation.

has an unknown elevation. The following plane equation (reported by Murchison (1977) as a triangular facet method of acquiring DEM) is used to acquire (z_b):

$$z_b = c_1x_b + c_2y_b + c_3 \tag{18}$$

where c_1, c_2, and c_3 are constants. A minimum of three points, a_1, a_2, and a_3, can be used to form three equations to have a unique solution for the unknown constants: i.e.,

$$z_1 = c_1x_1 + c_1y_1 + c_1 \tag{19}$$

$$z_2 = c_2x_2 + c_2y_2 + c_2 \tag{20}$$

$$z_3 = c_3x_3 + c_3y_3 + c_3 \tag{21}$$

After some manipulations, we get

$$c_1 = \frac{(z_1 - z_2) - (z_1 - z_3)(y_1 - y_2)}{(x_1 - x_2)\,(y_1 - y_3) - (x_1 - x_3)(y_1 - y_2)} \tag{22}$$

$$c_2 = \frac{(z_1 - z_2) - c_1(x_1 - x_2)}{(y_1 - y_2)} \tag{23}$$

$$c_3 = z_1 - c_1x_1 - c_2y_1 \tag{24}$$

Another feature of the expert system is the ability of the inference engine to initiate a random point interpolation method when appropriate. This method is used to generate a DTM using a second-degree equation of a surface based on the principles of least squares. This method is flexible and can be beneficial where random points are picked by human experts. The mathematical model is expressed as follows:

$$z = c_1x^2 + c_2xy + c_3y^2 + c_4x + c_5y + c_6 \tag{25}$$

where c_1, ..., c_6 are the unknowns. Having $n > 6$ points with known (x,y,z) coordinates surrounding the unknown point, the elevation of the point can be computed. Selecting proper weighing methods such as the inverse-of-square distance from relevant points, Equation 25 may be generalized in a matrix form where observation equations and normal equations are formed as follows:

$$z = \mathbf{N}^{-1}\mathbf{U} \tag{26}$$

where **N** and **U** are as follows:

$$\mathbf{N}^{-1} = (\mathbf{A}^{T}W\mathbf{A})^{-1} \tag{27}$$

$$\mathbf{U} = (\mathbf{A}^{T}Wl) \tag{28}$$

where **A** is the design matrix formed from elements of Equation 25. If the image distance is accepted as a criteria for weighting purposes, then

$$W = \frac{1}{d^2} \tag{29}$$

Simulation and Experimentation

The expert system was developed to apply the previously explained theory for image interpretation. The general configuration of the system (called KBI²S) is presented in Figure 4. It is a PC-based interactive expert system with a considerable degree of automation. The system is frame-based (Hayes-Roth *et al.*, 1983; Rich *et al.*, 1991) and written in a LISP-based language where a backward search strategy (Al-garni *et al.*, 1992) of solution is implemented. Different image features have been simulated, processed, and tested on the system.

A total of three image features were simulated based on data acquired from standard image interpretation sources (Al-garni, 1992; Way, 1973; Zuidam, 1985; Mintzer, 1989). These ideal features were tested on the system by different panels. In this way, we knew the ID of each feature before we asked the computer to reveal the identities, but the panels who input relevant information to test the system had no prior information about the identity of the features.

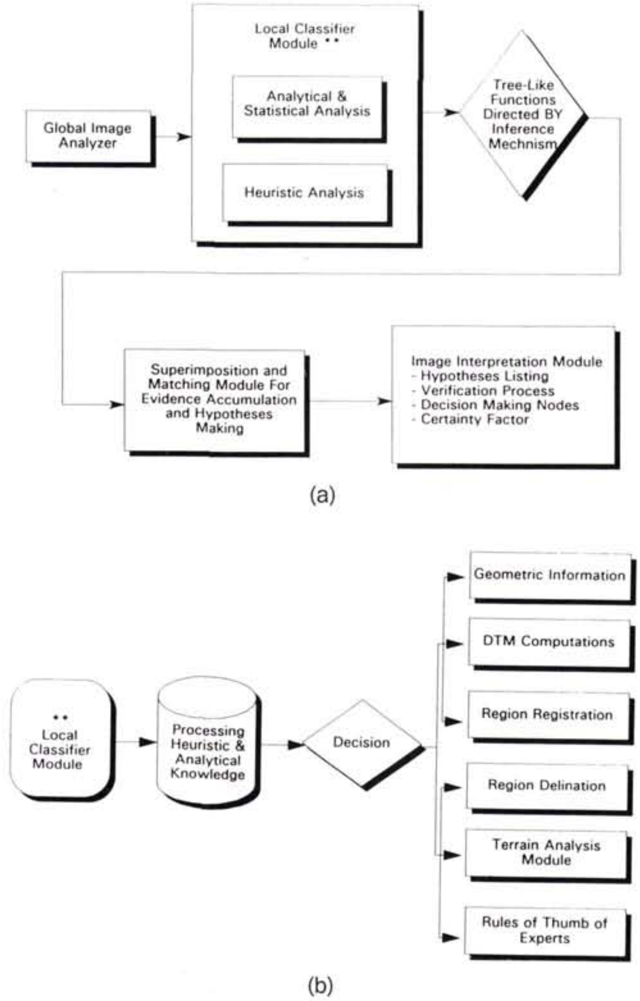

Figure 4. (a) General configuration of KBI²S. (b) Detailed description of the local classifier module in Figure 4a.

Digital elevation models for three sites are shown in Figure 5. These DTMs were overlaid on three different templates of knowledge. The first was a terrain analysis template which contained generic descriptions of surfaces. The second template contained analytical aspects of topographic surfaces in a 3D environment. The third template contained heuristic knowledge and expertise that influenced decision making regarding the ID of image features.

Table 1 includes an example of pieces of evidence that each template provided. Accumulative evidence is provided by the inference mechanism based on hypothesis and verification of hypothesis. Results obtained were compared with the *a priori* ID of these features and found to be in full correspondence (see Table 2).

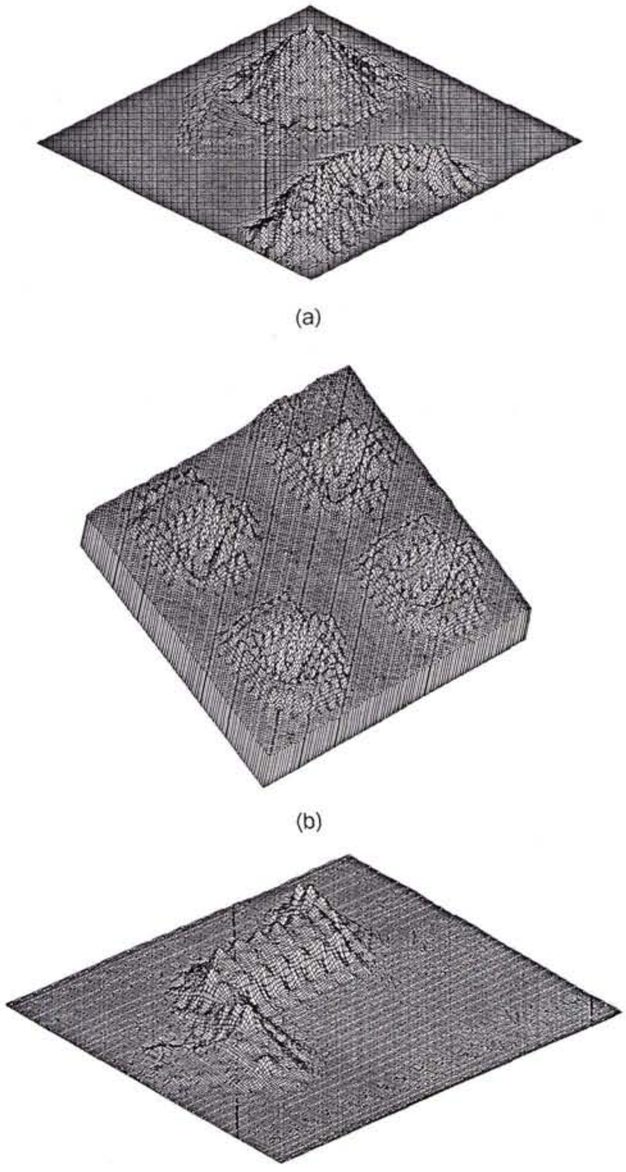

Figure 5. (a) DTM simulation for granite site. (b) DTM simulation for limestone site. (c) DTM simulation for esker formation.

Table 1. ID Template Knowledge Extracted by the KBI^{2}S

Region # (Image #)	Terrain Analysis Template	Analytical Aspects Template	Heuristic Knowledge Template
I	Elongated Snake Like Shapes with Internal Drainage	10-50*m* Width 1-5*m* Height >100*m* Length	Light Tones and Natural Cover
II	Surface Depression or Small Sinkholes	<1.2*m* Depression 0.4-3*m*Radius 0.0-0.3*m* Edge Height	Internal Drainage, Mottled Tone, and Cultivated
III	A-Shaped Hill with U-Shaped Gullies and Natural Cover	0.80*m* Height in a Large Form of Land Masses	Dark Dray Tone and Dendritic Curved End Drainage Pattern

Table 2. Comparison between Computer ID Template and Human Inference

Image # or Region #	Input Template	*A Priori* ID	Computer ID	Certainty Factor (%cf)
I	Table I Columns 2, 3, 4	Esker	Esker	92
II	Table I Columns 2, 3, 4	Limestone	Limestone	98
III	Table I Columns 2, 3, 4	Granite Large Masses	Igneous Rocks (Granite)	90

Conclusions

KBI^{2}S utilizes DTMs as a quantitative knowledge factor that I believe will open a new era to the automation aspects of image interpretation because topographic attributes of land surfaces contain the main clues and evidence for feature identities. Digital photogrammetry provides the means to improve the way image interpretation is conducted. AI adds a fourth dimension (expertise merged with analytical aspects of digital photogrammetry) to digital photogrammetry. KBI^{2}S provides a new insight to the integration of different stages of computer vision in digital photogrammetry. The system is expandable and will be modified to accept and manipulate a wider range of analytical and heuristic knowledge from digital images. The results obtained are very encouraging because all features were correctly and automatically identified by the system.

References

Al-garni, A., 1992. *Image Interpretation for Landforms Using Expert Systems and Terrain Analysis,* Ph.D. Dissertation, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio.

Al-garni, A., T. Schenk, and D. Way, 1992. Control Strategies for an Expert System to Interpret Landforms, *International Archives of Photogrammetry and Remote Sensing,* ISPRS, Washington, D.C., Volume XXIX, Commission VII, pp. 605–613.

Argialas, D., and R. Narasimhan, 1988. TAX: Prototype Expert System for Terrain Analysis, *Journal of Aerospace Engineering,* American Society of Civil Engineers, 1(3):151–170.

Dowman, I., H. Ebner, and C. Heipke, 1992. Overview of European Developments in Digital Photogrammetric Workstations, *Photogrammetric Engineering & Remote Sensing,* 58(1):51–56.

Gruen, A., 1989. Digital Photogrammetric Processing Systems: Current Status and Prospects, *Photogrammetric Engineering & Remote Sensing,* 55(5):581–586.

Hayes-Roth, F., D. Waterman, and D. Lenat (editors), 1983. *Building Expert Systems,* Addison-Wesley Publishing Company, Inc., London.

Miller, S., U. Helava, and K. Devenacia, 1992. Softcopy Photogrammetric Workstations, *Photogrammetric Engineering & Remote Sensing,* 58(1):77–83.

Mintzer, O., 1989. Research In Terrain Knowledge Representation for Image Interpretation and Terrain Analysis, *U.S. Army Sympo-*

sium on Artificial Intelligence Research for Exploitation of Battlefield Environment, El Paso, Texas, pp. 277–293.

Moffitt, F., and E. Mikhail, 1980. *Photogrammetry,* Harper & Row, Publishers, New York.

Morris, K., 1991. Using Knowledge-Base Rules to Map the Three- Dimensional Nature of Geological Features, *Photogrammetric Engineering & Remote Sensing,* 57(9):1209–1216.

Murchison, D., 1977. *Surveying and Photogrammetry, Computation for Civil Engineers,* Newnes-Butterworths, Boston.

Rich, A., and K. Knight, 1991. *Artificial Intelligence,* McGraw-Hill, Inc., New York.

Richards, J., 1986. *Remote Sensing Digital Image Analysis, An Introduction,* Springer-Verlag, New York.

Schenk, T., 1992. *Machine Vision and Close Range Photogrammetry. Research Activities in Digital Photogrammetry at The Ohio State University,* A Collection of Papers Presented at the XVII Congress of ISPRS, Department of Geodetic Science & Surveying, Report No. 418.

Schenk, T., and C. Toth, 1992. Conceptual Issues of Softcopy Photogrammetric Workstations, *Photogrammetric Engineering & Remote Sensing,* 58(1):101–110.

Skalet, C., G. Lee, and L. Lander, 1992. Implementation of Softcopy Photogrammetric Workstations at the U.S. Geological Survey, *Photogrammetric Engineering & Remote Sensing,* 58(1):57–63.

Way, D., 1973. *Terrain Analysis,* Second Edition, [Publisher, Location].

Zuidam, R., 1985. *Aerial Photo-Interpretation in Terrain Analysis and Geomorphologic Mapping,* Smits Publishers, The Hague.

(Received 11 March 1993; accepted 20 October 1993)

A. Al-garni

A. Al-garni received the Bachelor of Science Degree in Civil Engineering, Majoring in Surveying, at King Saud University (KSU) in Saudi Arabia, Riyadh, in 1983. He was selected as a Teacher Assistant for one year in the department. In 1984 he received a scholarship from KSU to study for the Masters and Ph.D. degrees in the United States. He received the MS degree from The Ohio State University, majoring in Remote Sensing. He received another MS degree, from the Department of Geodetic Science and Surveying, The Ohio State University. By 1992 he had received the Ph.D. degree from The Ohio State University, majoring in digital and analytical photogrammetry, and conducting Artificial Intelligence (AI) research for image interpretation. Presently he is an assistant professor at KSU, a partial consultant to the military survey department, and a committee member in establishing the national mapping center in Saudi Arabia.
