USING DIFFERENT SOURCES OF INFORMATION IN AUTOMATED LINEAR FEATURE EXTRACTION FROM REMOTE SENSING DATA

P. O. Simoni, V. Rodrigues, P.P. B. Oliveira

Instituto de Pesquisas Espaciais - INPE/MCT
Laboratório Associado de Computação e Matemática Aplicada
Caixa Postal 515, 12201, São José dos Campos, SP, Brasil

ABSTRACT

The extraction of linear features from LANDSAT satellite images for structural geology applications is analysed. It is also discussed the influence of a number of factors such as sensor type, position of the observer, shadowing, and knowledge about dominant structures, over the inference of relief convexity, the detection of linear features and the type of feature. The integration of these factors are then studied in the context of an automated linear feature extraction process under specification and it is argued about the desirability of a knowledge representation model for combining all the available sources of information that are important for the extraction process.

1. INTRODUCTION

The modeling of structural feature extraction process from remote sensing data by geologists is a very complex task, and it is not known up to this time the exact extent to which many objective or subjective factors influence it. Despite the fact that uniformity in the results coming from many persons is not guaranteed to be attained, other problems are intrinsic to the task of extracting lineations and lineaments, such as its hard and tedious character and the requirement of high visual acuity from geologists.

Any attempt to develop a computer system that would help geologists in this task will necessarily face a set of complex problems to be solved.

Some early attempts to develop image processing systems to help in the linear feature extraction process are commented next along with their limitations. The research developed by Moore and Waltz (1983) led to an empiric method for lineament detection and enhancement, aimed at determining the geologic and hydrologic significance of the results, and they observed some limitations of their approach, related to the enhancement of meaningless lines, to the enhancement of linear features in undesirable directions and to the difficulty of obtaining results in shadowed or uniformly covered flat areas and in dissected terrain; they also observed that there is no necessary correlation between image contrast of a linear element and the geologic significance that may be assigned to it. Ehrich's (1979) system employs global information for generating piecewise linear approximations to lineaments and growing them; this system avoids the use of (convolution) filters because they set in advance the sizes of the elements to be detected. In another system (Abrams et alii, 1985), (Nguyen et alii, 1984) 95% of the lineaments are present in the contour image, as well as the main fault directions and the main directions in the rose diagrams, but the contour image is noiser than the hand-drawn one, there are some linear features that are not as long in this image as in the hand-drawn one and there are also others that were not even detected; a number of adjustements were needed in order to arrive at the mentioned performance level but it is not known if the same fit sequence can be applied successfully to other images. Mathematical morphology techniques were used by Flouzat et Moueddene (1986) in order to extract elementary elements and there are situations in which elements were not extracted when searching for the main direction in an anticline.

An analysis of the commented attempts suggests an alternative approach to the problem, a knowledge-based computer vision system in which the application of image processing techniques is guided by high level knowledge and in which their application follow visual perception models.

This paper will deal basically with one important aspect in the design of such a vision system: how to design a software system that would allow the integration of the many information sources that may interact with the extraction process and would also have information symbolically and explicitly processed.

2. THE PROBLEM OF AUTOMATIC LINEAR FEATURE EXTRACTION

The extraction of linear features from aerial photographs and satellite images is a very important step in the broad research usually performed when analysing a region for mineral research. Many kinds of maps, such as geologic, tectonic and metalogenetic for instance are generated in this analysis, in an attempt to correctly identify the structural patterns that corresponds to the relief information.

If this is an important step in the analysis of a region, both for correct geologic understanding about region formation and also for economic assessment, it is desirable to investigate the conditions under which it could be total or partially automated.

Since the early observations by Hobbs (1904) about the fact that maps of large regions depict some features that are related to rock structure, a number of reports approached the problem of delimiting the extent of the involved concepts - lineation, lineament, linear and linear feature. It must be noticed that even nowadays there is no general agreement within the professional community about the scope and exact meaning of these terms (O'Leary et alii, 1976).

Other intrinsic difficulty associated to the automation under study is that the extraction process may be influenced by some subjective and not entirely controllable conditions. It is a known fact that stress may be a cause either for unreliable detection or for time varying behavior of a given individual. Furthermore, evidences were collected indicating that several geologists analysing the same image may build different linear feature maps (Podwysocki et alii, 1975). There are other factors (objective or subjective - such as the determination of local relief convexity information from a single source) that may be taken into account when studying the extraction process performance, as it will be commented in the next sections.

In this paper, the terms lineation, lineament and linear feature will have the following meaning. The term lineament will be used in the sense of simple or composite linear features of a surface, whose parts are aligned in rectilinear or slightly curvilinear relationship (O'Leary et alii, 1976). Lineation is a term related to petrography, and the parallelism or alignment of small elements and the observation or large scale are emphasized; they can be represented only symbolically in a map. Lineament, on the other hand, is associated to geologic elements (foliations, joints, etc); in images, a superficial expression of it is available. Linear features will refer to elements of an unspecified origin.

The influence of the above mentioned subjective factors would be lessened by choosing an objective extraction process. The underlying problem is to assess how objective is a given procedure. One useful assumption is that persons will be offered a supervised training according to an uniform paradigm. These factors will have influence over the system under design even if it is highly automated, because the operator may fall in a situation where he must either check for the validity of the extracted elements or modify part of an automatically generated map.

The just mentioned problems suggests some preliminary features of an automatic system:

a) it must be interactive, in the sense that information about a linear feature being rejected or added by an user should be accepted.

b) it is desirable that an extracted element have attached the evidences used in its inference and also the parameters of algorithms used for the related computations, in order to explain or to accept modifications in the extraction process.

These requirements do not solve the difficult problem of uniformity (the chosen definitions of terms are not yet completely unambiguous up to now for operational purposes, and the system should also be influenced by high level knowledge, for instance), but a frame of reference will be available for comparison purposes among members of a team analysing a region.

3. SOME FACTORS RELATED TO LINEAR FEATURE EXTRACTION

Some subjective factors were already commented and additional ones will be mentioned next.

a) Characterization of features according to sensor type

There is no general agreement about the way for choosing the more suitable kind of sensor to analyse a given region. The outcomes from imagery of different kinds of sensors may differ depending on a number of circumstancial factors. If the region has semi-arid climate, it is a good policy to pick images from both rainy and dry seasons, in order to distinguish seasonal effects from structural ones; the kind of relief also imposes constraints on the preferrable bands (the analysis may be morphology or gray level directed) and elevations. Band 7 is considered useful in general for confirming already detected information. Moreover, there may be situations (lithology separation, for instance) in which the area covered by a given image may contain regions which call for more than one kind of sensor.

Parsons and Yearley (1986) conducted a study about the information that can be extracted by a single operator from LANDSAT images of four bands using a standard extraction technique, and collected evidence that one source does not subsumes another - there are lineaments that belong to only one image. It was apparent from their study, however, that the same preferred orientation can be detected in almost all maps.

When available, radar images allow a great enhancement of geological features and have the advantage of allowing the choice of scan direction. They present however low resolution and difficult extraction in directions close to the scan direction.

b) **Spectral response**

Besides the direct influence of sensor type on the presence of linear features, there are available some evidences about how the spectral response of a given band allows the inference or the rejection of some pixel classification. This may be useful for confirmation purposes (when one is checking a hypothesized drainage, for instance; in this sense, some bands are more suitable for deciding if a given area corresponds to water).

Salisbury and Hunt (1974) studied the spectral behavior of several kinds of rocks under the assumption of weak chemical interactions; their results can not be applied carelessly to Brazilian sites, due to the high occurrence of chemical modifications (due to rain and decomposition).

c) **Sensor resolution**

Another factor related to the sensor is the resolution and the area covered by an image. A high resolution sensor would be desirable, but the available technology limits the resolution that can be attained in near future. Resolution has an effect over many small elements that can not be detected in satellite images.

d) **Season of the year**

Parsons and Yearley (1986) also showed that there is an influence of the season in the population of extracted lineaments - there is no subsumption among images from different seasons; nonetheless the preferred orientation can be obtained from any of the images of different seasons.

e) **Position of sun**

It is often assumed that low elevation (35 - 45 degrees) images are preferred for lineament detection (Moore and Waltz, 1983), because shadows enhance a large class of lineaments - depressions, in eroded terrains.

f) **Trend detection**

Some geologists suggest that if the general trend of the region covered by the image is identified a priori (a known river or mountain formation), the extraction process becomes easier. It must be remarked however that the most evident elements in an image are not necessarily the most important for mineralogical studies.

g) **Relief information**

Stereo information also has influence on the extraction process.When using aerial photographs, stereo image pairs are quite common, When the primary source is satellite images, there is some overlap between adjacent images in the case of LANDSAT imagery (the overlap depends on the latitude) and stereo pairs can be ordered when using SPOT imagery. It was observed however in trained geologists the skill of inferring local relief information from shadows in a single image (almost at once, when presented an image).

Known automatic procedures for depth computation in this kind of images are not reliable yet. The influence of shadowing on the aproximate determination of the local slope, of relative elevations and of drainages from a single multispectral view was studied by Haralick et alii (1985) for a dissected region, taking into account ancillary knowledge such as sun position. The separation of topography and reflectance effects was performed using a surface Lambertian reflectance model. One important aspect related to stereo information is that is not known exactly how much information is needed for this kind of application. Observations suggest that local relief information (about an edge being convex or concave) may be enough, but there are situations where this information is not clear and detailed depth computations may be needed or digital terrain models may be used instead.

h) **Texture characterization**

The particular analysis to be performed on images is such that the "objects" to be identified are not completely characterized by their features before process beginning. The sought elements have a shape almost linear, may appear either as edges or lines, may have varied sizes, may exhibit linear or almost parallel formations (in which the common basis, not present in the image, must be determined) and their acceptance may depend on the global context in which they appear.

The attributes of the linear features to be extracted may vary in a range of values:

- linearity: small nonlinearities are allowed for the smaller elements; the accepted deviation is subjective.
- length: there is no safe indication about the minimum or maximum length of an element. A given element that may be put in correspondence with a fracture will be considered a joint or a fault whether there are evidences about crustal displacement. An useful approximation when no further information is available is that surface elements greater than 1,5 to 2,5 km may be faults.
- image contrast: the contrast between crest sides depends on the slope of the illuminated side in relation to the light source.

It is also important to consider the formation of larger aligned unities (lakes, for instance) which suggest some tectonic activity. The smaller elements will not necessarily be linear in this situation; they may have blob shape, for instance. It is supposed that the following definition of textural element holds in this context: it is the smallest homogeneous and continuous surface that may occur repeatedly. The separation of a greater element in its smaller constituent parts is appropriate whenever these elements can be detected.

4. PROPOSED AUTOMATED SYSTEM

The problem of designing an automated tool to substitute completely or to a great extent human operators in this specific task is considered unsolved up to this time. Instead, it will be attempted to develop an interactive tool which can be upgraded step by step.

Some desirable features of an idealized robust system are listed as follows; it should:

- extract linear features in the whole image or in parts of it by its own initiative
- check for the existence of a feature in a given area whenever requested by an operator (reinforcing the thresholds used by low level processes)
- accept high level information given by an operator (such as one related to the existence of a dominant trend), which may have influence on the search process
- propagate information about either the rejection of features or the presence of trends to already extracted elements

Besides these features,

- available knowledge about a region should be allowed to be considered in case of small contrast elements (such as in fold regions)
- every extracted element should have associated to it the evidences that led to the extraction
- the system must be able to identify what is the most reliable information source in a given situation, when a complementary source must be used, how to compare information coming from many sources and how to weight them. This feature will be the subject of a coming section.

The system being developed will basically have the structure of Figure 1. This structure does not explicit neither the temporary structures needed during inference processes, such as an agenda-like structure for storing tasks to be executed and a structure to propagate evidences, nor the system control flow. The structure for evidences allows the storage of both inferred hypotheses and questions to be answered.

Meta-knowledge will consist for instance of control rules for specifying a rule or a set of rules to be selected next, in the rule related part of the representation.

A given set of observed data may generate multiple (even conflicting) hypotheses; thus there must exist judgement rules for selecting the most likely one. Moreover, identification rules may give conditions for deciding whether a spatial sequence of points and segments may be taken as a lineament or a lineation.

The following information must be associated in the image structure to every extracted element: the algorithm used for extraction, involved thresholds, the window on which a search was done, a flag to inform if the search was performed upon request of a symbolic inference, another

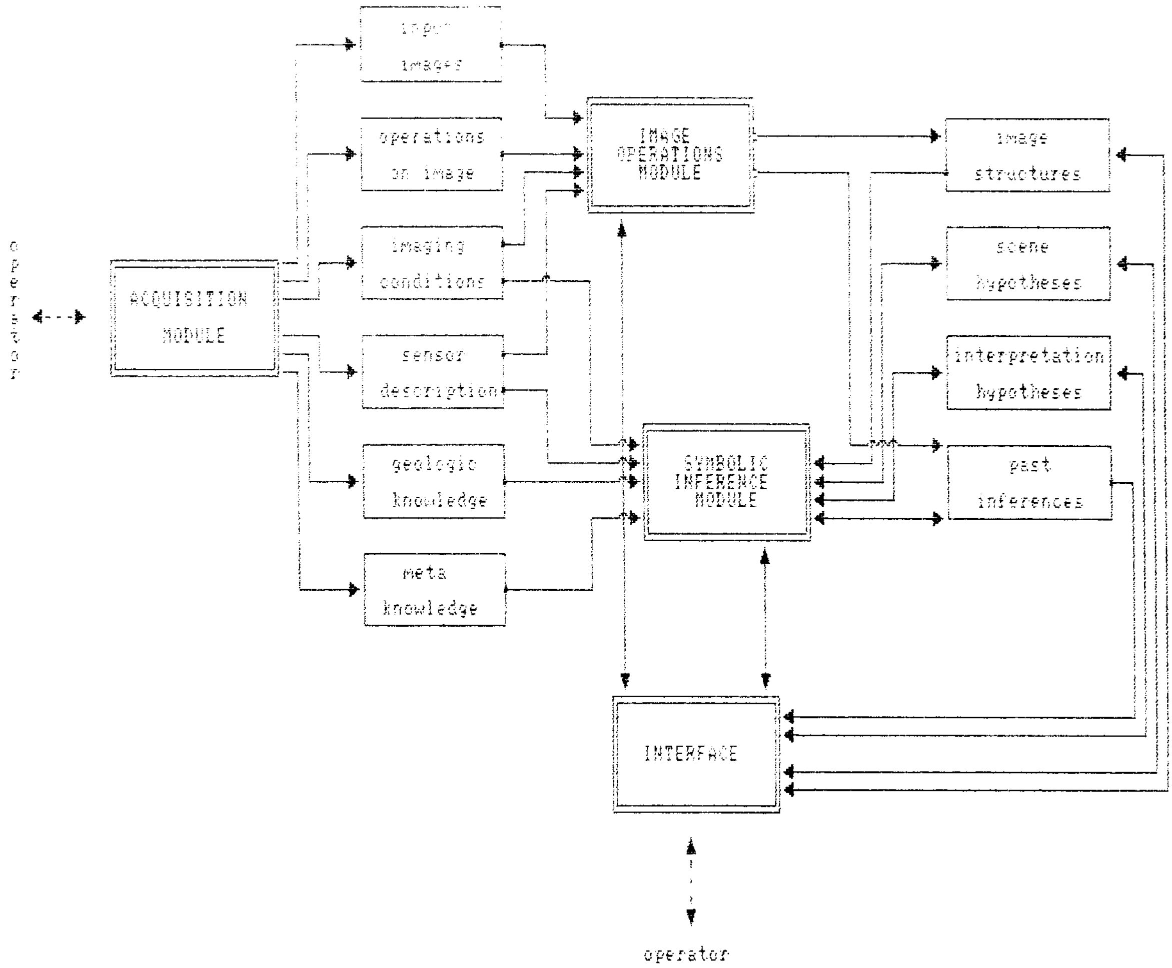

Fig. 1 - System structure.

informing if an interaction with operator occurred and its type, its main attributes (contrast, position, length) and subjective acceptance degree. Furthermore the focus of attention for the region being inspected is stored and available to the inference processes.

All past inference steps must be stored, in order to achieve explanatory purposes. In addition, it will be interesting to keep this information in order to attain nonmonotonic system behavior. This behavior is important because there are global hypotheses that may act on the extraction process, and also global hypotheses that may be dependant on the extracted elements.

Hypotheses may be attached to extracted elements relating to their correspondence in the 3D scene: if they are concave or convex edges, rivers, or even shadow, vegetation or antropomorphic separation borders.

Another kind of hypothesis to be stored, not always available or unique, is about the geologic interpretation of an element or group of elements (such as in discontinuous lineaments or larger formations), which may be obtained by inference or from the operator. For a cluster of small elements, even if not contiguous, its spatial structure is explicity stored. For larger structures, masks are stored to make it easier to refer to the whole structure. Interpretive hypotheses may also be related to larger units, from which the region in a single image is part (such as the situation when one is dealing with the geologic age of a larger region).

Two sets of processing activities are characterized; they are conceptually different, they make access to different storage and different working structures and generate different classes of information. A kind of image related processing is the computation of relief convexity from shadows and position of sun; its outcome is a symbolic information that will be used in other different processing paradigm. It must also be kept stored along with this obtained information the related hypotheses and parameters used in the image processing algorithms.

5. COMBINATION OF INFORMATION FROM MANY SOURCES

It was already observed that the system should be able to combine information from many sources. This requirement was introduced in order to lessen a notorious advantage that human extractors have over the systems developed so far. It does not overcome however other drawbacks, such as those related to the visual process, which had partially been taken into account in the proposed system architecture.

This integration may result in particular control system behavior:

- confirmation of hypothesis: if a hypothesis made about an extracted element is ambiguous (a given line may correspond either to a drainage or to a lineament, for instance), a question may be generated about its validity, and stored knowledge about it (the knowledge that an image from another band can help in deciding) may be invoked in order to clear it up, within a confidence degree range.
- reinforcement of image processing algorithm parameters: the equivalent of a human careful examination may be partially implemented through an appropriate choice of parameters associated to image processing techniques (filter size and edge minimum size or contrast, for instance).
- hypothesis generation: there are some combinations of observed elements that may trigger a high level hypothesis (a particular geometric relation between two linear features may be enough to infer a fault zone, for instance).

Some of the most relevant aspects to be considered are listed next; the exact meaning of the related adjectives (close, neighbor, etc.) is not mentioned but it must be stored in the chosen representation.

- if satellite images are being used and there is a low contrast element then if the local trend direction is known and the element direction is close to the trend direction, then a closer examination should be performed.
- if the general structures or trends are known by a geologist, their main characteristics (kind of structure, aproximate predominant direction, position, length, kind of related geological phenomena) should be acceptable as high level knowledge that may influence the search.
- from a set of neighbor linear features, it may be assigned a geological interpretation - a fault - if there is enough evidences about the displacement; the evidences may be given by the presence of surface foliations. If such an interpretation can be assigned, its specific type (inverse, for instance) should also be described.
- whenever there are two almost neighbor and parallel linear features which can be considered as being positive, and no negative feature is identified between them, it may be the case that the negative feature was hidden by vegetation, image scale or clouds, for instance, and both a closer look at the neighborhood of the positive linear elements in the same image and in images from other bands should be performed.
- in general it is difficult to identify antropomorphic structures; some clues that may be used are that the separation edges are generally rectilinear or smoothly curvilinear and the enclosed region has uniform spectral response (in vegetation characteristic sensor bands) for agricultural sites; for roads, another clue is that the line is also rectilinear or smoothly curvilinear, the spectral response lies in a known interval, there may exist a neighbor region that may correspond to bare soil; etc.
- if hypotheses can be made about the presence of certain kinds of rocks (acidic rocks, for instance), a high contrast between vegetation and rocks is expected.
- the characterization of curved composite shapes (such as the lineaments that may appear in areas with faults and folds) may give a strong evidence about the existence of mineralizations in the separation between two folded zones.
- if satellite images are being used and there is a low contrast element then if sun elevation is not low (35 - 45 degrees) and there is another image that covers the same coordinates with lower elevation, an additional feature to be considered for future versions of the system being proposed is that it should suggest an inspection in this image. The image registration problem will have to be considered in this situation.
- if the primary source of information is radar then the extracted linear elements will probably be lineaments or lineations (and not the separation between two regions with different kinds of vegetation) if they lie in directions other than those ones close to the scan line; on the other hand, they will receive low confidence degree and require a closer examination (or check if there are many other elements is the

same situation) if their direction is close to the scan direction.

6. CONCLUDING REMARKS

The rationale for a proposed automated linear feature extraction model from remote sensing data was exposed and the role of a knowledge based vision approach was analysed taking into account the requirements for a robust extraction process.

Besides the support to linear feature extraction, a knowledge model yields another benefit. The analysis of a region for ore deposit characterization requires many steps, such as the generation of geologic, tectonic and metalogenetic maps, and the identification of both a province and an ore deposit. As this job is usually performed by many persons in different places and times, a method for storing information coming from many sources (satellite and radar images, airplane photos, for instance) with different confidence degrees and different supporting evidences would be undoubtedly useful.

7. ACKNOWLEDGEMENTS

The authors are very thankful specially to the geologists Juércio Tavares de Mattos e Wougran Soares Galvão for their attention, suggestions and criticisms along all the discussions we have had so far concerned with the proposed system and for introducing us the first concepts in geology itself.

8. REFERENCES

ABRAMS, M.; BLUSSON, A.; CARRERE, V.; NGUYEN, T.; RABU, Y. Image processing applications for geological mapping. **IBM Journal of Research and Development**, 29(2):177-187, 1985.

EHRICH, R.W. Detection of global lines and edges in heavily textured images. In: INTERNATIONAL CONFERENCE ON BASEMENT TECTONICS, 2. **Proceedings.** 1979, pp.508-513.

FLOUZAT, G.; MOUEDDENE, K. Computer-aided interpretation of complex geological patterns in remote sensing. In: IGARSS SYMPOSIUM. **Proceedings.** Zurich, 1986, pp. 783-786.

HARALICK, R.M.; CAMPBELL, J.B.; WANG, S. Automatic inference of elevation and drainage models from a satellite image. **Proceedings of the IEEE**, 73(6):1040-1053, 1985.

HOBBS, W.H. Lineaments and the Atlantic border region. **Geological Society of America Bulletin**, 15: 483-506, 1904.

MOORE, G.K.; WALTZ, F.A. Objective procedure for lineament enhancement and extraction. **Photogrammetric Engineering and Remote Sensing**, 49 (5):641-647, 1983.

NGUYEN, P.Y.; SIMON, S.; ASFAR, L.; BLUSSON, A. Extraction de parametres de forme pour la recherche des structures linèaires dans les images-satellite. In: COLLOQUE IMAGE - TRAITEMENT, SYNTHESE, TECHNOLOGIE ET APPLICATIONS, 1. **Proceedings.** Biarritz, 1984, pp. 741-746.

O'LEARY, D.W.; FRIEDMAN, J.D.; POHN, H.A. Lineament, linear, lineation: some proposed new standards for old terms. **Geological Society of America Bulletin**, 87:1463-1469,1976.

PARSONS, A.J.; YEARLEY, R.J. An analysis of geologic lineaments seen on LANDSAT MSS imagery. **International Journal of Remote Sensing**, 7(12):1773-1782, 1986.

PODWYSOCKI, M.H.; MOIK, J.G.; SHOUP, W.D. Quantification of geologic lineaments by manual and machine processing techniques. In: NASA EARTH RESOURCES SURVEY SYMPOSIUM. **Proceedings.** Houston, TX, 1975, pp. 885-903.

SALISBURY, J.W.; HUNT, G.R. Remote sensing of rock type in visible and near-infrared. In: INTERNATIONAL SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT, 9. **Proceedings.** Ann Arbor, MI, ERIM, 1974, v. 3, pp. 1953-1958.