STRUCTURAL MATCHING OF DIGITAL IMAGES AND TERRAIN MODELS

Kostas Papanikolaou
Eugene Derenyi
Department of Surveying Engineering
University of New Brunswick
Fredericton, N.B.
Canada E3B 5A3

ISPRS Commission III

ABSTRACT

A method for the registration of digital images and digital terrain models (DTM) is presented. The method is based on the structural matching of line features. A network of promising ridges is first extracted from both the image and the DTM data. Directional illumination of the topography is one of the techniques used in pursuit of that aim. An automatic matching between these two sets of ridges is then performed. Geometric and neighbouring relations are used as "matching keys". Results of an example are included in the presentation.

1. INTRODUCTION

Registration of remotely sensed images to a reference coordinate system, be it in object space or in another data set, constitutes an important first step in automatic scene analysis or in the analysis of multilayer data sets. It allows the use of one-to-one correspondence in the analysis and facilitates the merging of the results with information originating from other sources. Typically, registration uses discrete points that can be positively identified in the image. The procedure presented here is mainly designed for mountainous or rugged terrain. Features commonly used as control points do not exist here or are in short supply. Topography is a governing factor for surface illumination and reflectance. Therefore, the most prominent features for registration are mountain ridges that cause sudden changes in surface illumination.

A digital terrain model (DTM) in the form of grid elevation data is used as the geocoded basis for image registration. The procedure involves two major steps. A network of prominent ridges is first extracted from the DTM. The same ridges are also detected in the image. Then the problem of registration is transformed into the problem of matching two sets of linear features. In the second step structural matching of ridges is performed based on geometric properties of neighbouring line segments. Since the matching method relates line segments between two corresponding data sets, it can be applied for automatic registration in a more general sense. It can be employed for matching planimetric features, that are detected in an edge enhanced image, against a corresponding set of features in a digital map data base or in another edge enhanced image.

2. FEATURE EXTRACTION

Two different approaches were examined and implemented for the extraction and selection of ridges from the grid elevation models. In the first approach, a set of ridges is extracted directly from the terrain model by local processing of elevation data. In the second approach, an image of the same elevation model, enhanced by relief shading, is used and the detection of ridges is based upon discontinuities of radiance (gray level) values. The latter method equally applies to detection of ridges from digital images.

2.1 Detection of Ridges from Elevation Data

The task is to identify all points in the grid elevation model that lie on ridges. To achieve this in an automated way, an accurate and valid definition of these features is required. Several

suggestions can be found in the scientific literature. In water resources management ridges form the border lines of drainage basins. Consequently, if basins can be identified, then ridges can be traced as their boundaries. Douglas [1986], however, reported that such procedures yield unacceptable results or are prohibitively expensive because they introduce artifacts which must be removed by further processing. For the purpose of image matching it is not necessary to locate the entire ridge structure within a data set. It is sufficient to define traces of ridges locally. A ridge is considered to be a line on the terrain surface connecting two adjacent peaks (local maxima), following a path of the highest altitude. Ridges can also be defined as elongated regions of convexity or lines of divergence of slope. These definitions, however, apply only to continuous surface geometry and their immediate application to a digital elevation model, bounded by the discrete geometry of the sampled grid, is unsatisfactory [Peucker and Douglas, 1975]. The reason is that the recognition of ridges is ambiguous and sensitive to noise. Our implementation uses two methods of equal efficiency. Both are local in the sense that the decision on whether or not a grid point lies on the ridge is made on the basis of the elevations of the point itself and some of its neighbours. Both methods require just a single pass through the data array.

The procedure described by Jensen [1985] for the detection of channels by locating elevation points in the grided data which form "V" shapes, can also be applied to ridges by locating points forming "Λ" shapes. The procedure is as follows: The two diagonal and the two orthogonal cross-sections passing through each grid point, are examined in the 3x3 window around the point. If both endpoints of any cross-section are lower than the tested point in the centre, then the triplet of points forms a " Λ" shape and the point is considered to lie on a ridge. Jensen [1985] also recommends the use of asymmetric cross-sections so as to overcome the tendency of the method to deal only with well defined narrow channels. This approach however, results in noisy clouds of ridge-points. In our case, well defined sharp ridges are to be detected, so testing of asymmetric cross-sections is not included in the program.

The second local operator to define ridges is based on the elimination of points that cannot lie on a ridge. The rationale here is to locate the points to which water would flow from the central point and declare them as non-candidates to be on ridges [Fowler and Little, 1979]. At each time four points are checked, the central one and the three neighbours adjacent to it on the right and below. The lowest point or points among these four are eliminated as not being ridge candidates. Finally, only points lying on ridges are left.

Figure 1 shows a contour map of the grid DTM used in this experiment. It covers a small area in British Columbia. The size of the grid array is 48 lines by 160 pixels per line. Figure 2 shows ridges as detected following Jensen's suggestions, while Figure 3 presents the result of applying Fowler's algorithm. Note the artifacts produced at the border of the area. Post processing editing, as will be discussed later, has been applied to the detected pixels for thinning and noise removal purposes.

2.2 Extraction of Ridges from Images

The second approach uses a relief shaded model of the terrain to detect and extract ridges. The directional illumination used to enhance the topography highlights mountain ridges. Moreover the ridges that are parallel to the direction of the imaginary light source tend to disappear while the ones perpendicular are very distinct. In order to detect all ridges from the elevation model, synthetic images produced with various directions of illumination should be examined. However, illumination of the terrain is also an integral part of the production of real images. Consequently, if the position of the illuminating source at the moment of the image acquisition is used as the location of the imaginary light source for analytically shading the terrain model, the same mountain ridges will be apparent in both the real and the synthetic images. These ridges will become boundaries between bright and dark areas, and as such, they can be extracted using an edge detection scheme.

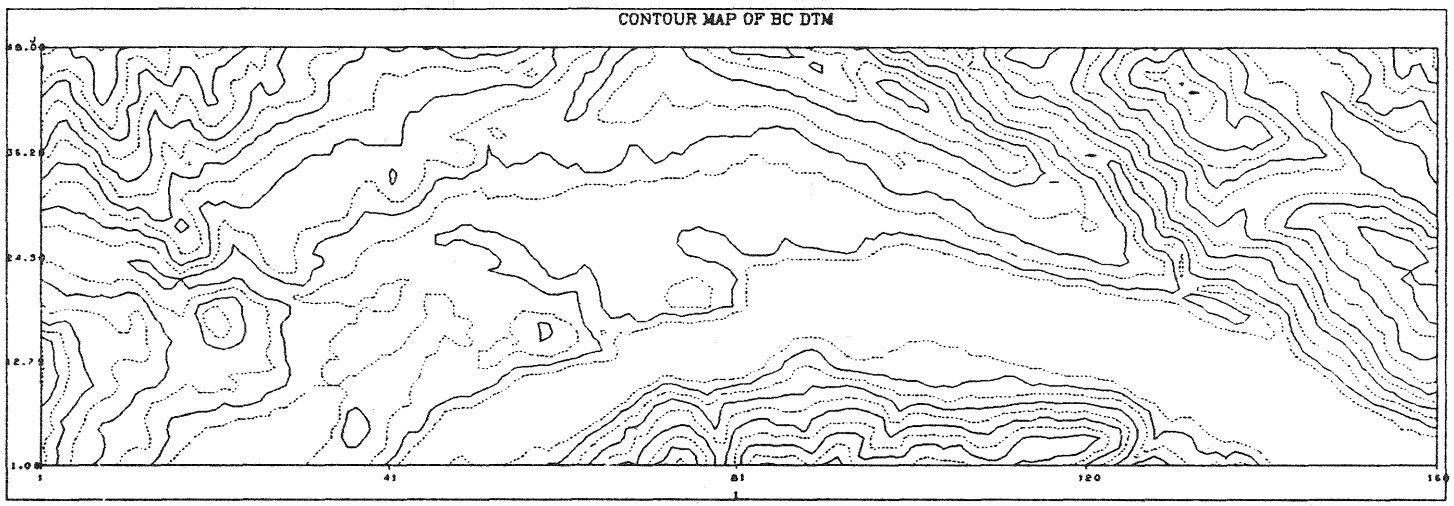

Figure 1: Contour map derived from the grid elevation model.

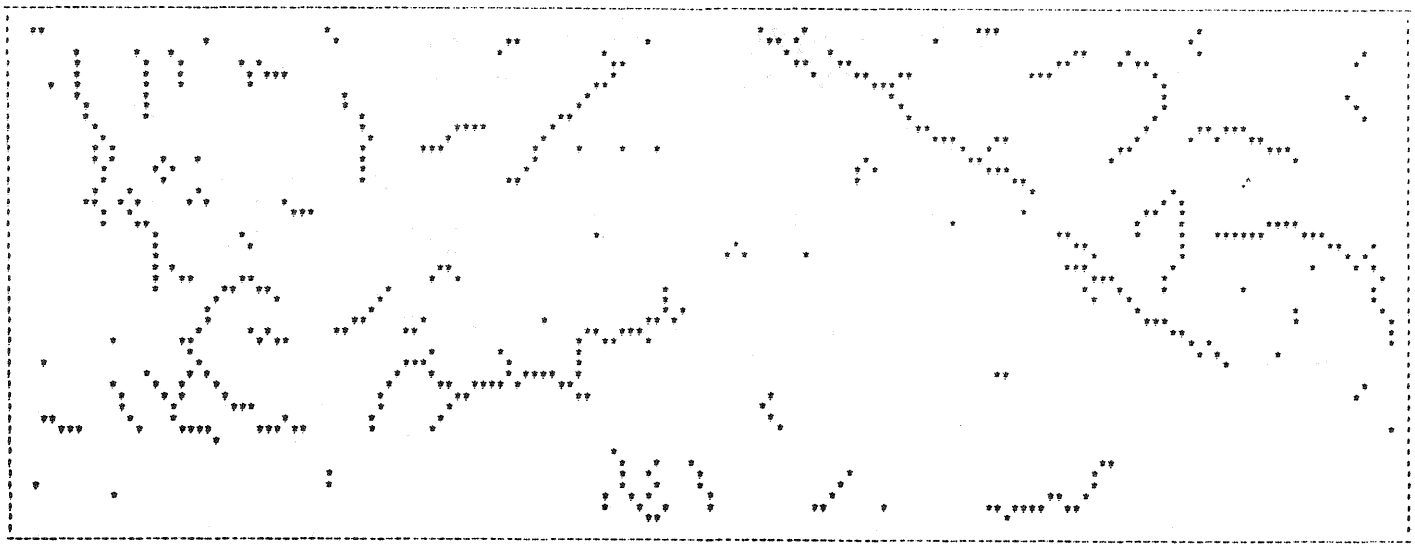

Figure 2 : Ridges detected using Jensen's approach.

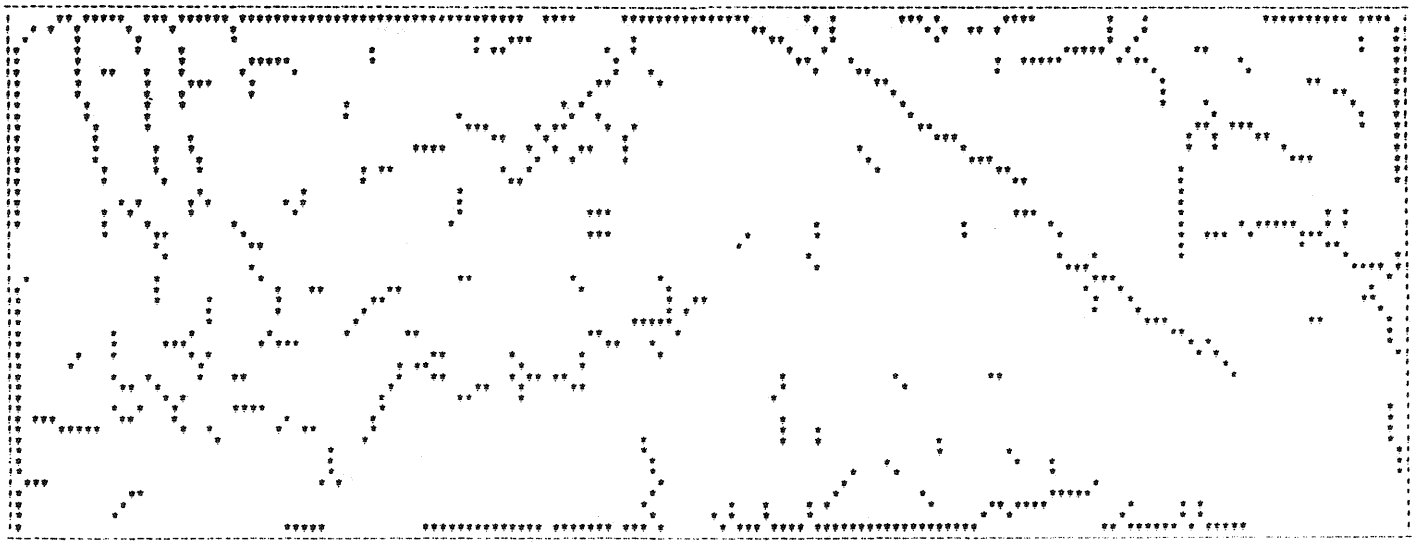

Figure 3 : Ridges detected using Fowler's approach.

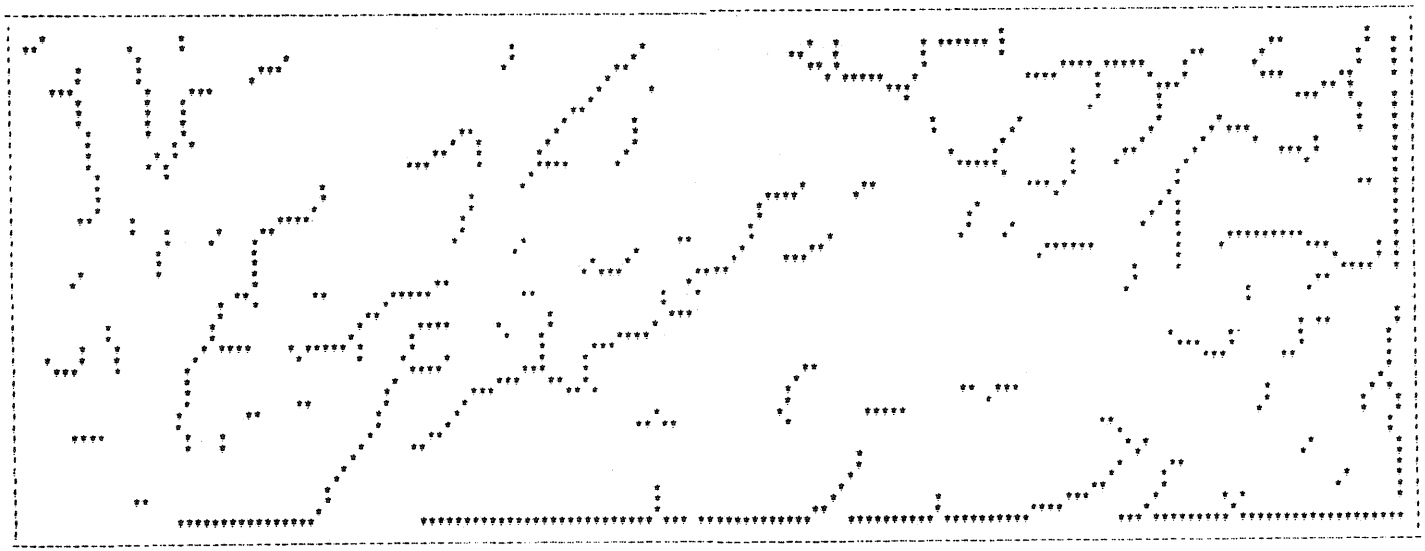

Figure 4 : Ridges extracted form the relief shaded image.

2.2.1 Production of Synthetic Image

Analytic hill shading techniques are employed to produce the synthetic image. The Lambertian model is used, assuming that the surface, being a perfect diffuser, appears equally bright from all viewing directions. According to this model, the reflected light from a surface element depends only on the direction of illumination and the orientation of the element. It is in fact proportional to the cosine of the angle between the local normal and the direction of the light source. The normal to the surface is expressed by the vector (-p, -q, 1), where p and q are the slopes along the two orthogonal directions, x and y. They can be computed locally from the grid DTM [Papanikolaou and Derenyi, 1987]. Furthermore, vector ($-p_O$, $-q_O$, 1) coinciding with the direction of illumination is a function of the orientation angle φ (measured anticlockwise from the East) and the zenith angle θ of the illumination source. Following Horn [1981]:

$$p_O = -\cos\varphi \tan\theta$$
$$q_O = -\sin\varphi \tan\theta$$

Since the cosine of the angle between two vectors is given by their inner product divided by their norms, the function

$$R = \frac{1.0 + p_o p + q_o q}{\sqrt{1 + p_o^2 + q_o^2}\sqrt{1 + p^2 + q^2}}$$

is an estimate for the reflectance value of the surface element. This function gives negative values for surfaces facing away from the sun and thus accounts for self-shadowed areas. In order to have a more realistic synthetic image, shadows cast by other surfaces are also considered. All areas in shadow are located with the use of a hidden-surface algorithm [Papanikolaou and Derenyi, 1987] and are assigned a constant minimum value denoting diffuse illumination.

2.2.2 Detection of Ridges in Synthetic or Real Images

In both images, ridges produce discontinuities in reflectance values. The parts of the surface facing the illumination source give relatively high response while the surface elements facing away from the source reflect only diffuse illumination which is set to a very small (or zero) constant value. Ridges lie on the transitions between bright and dark areas. As such, they can be extracted by an edge detection algorithm in the following manner:

First, a gradient operation is performed to produce an edge enhanced image. The original image is convoluted with the 3x3 Sobel operators [Gonzalez and Wintz, 1987]. At each pixel location (x, y), the two components, G_x, G_y, of the gradient are combined to compute the magnitude of the gradient as $G = (G^2_x+G^2_y)^{1/2}$. This value indicates the rate of maximum change for the image function at (x, y). Then the gradient image is thresholded to eliminate weak responses. As a result, pixels giving high changes in gray-level values - or in other words, pixels lying along sharp transitions in intensity values - will remain. A further constraint is added to the operation: At each pixel the direction of maximum change, a = arctan (G_y/G_x), is also calculated. This direction should point towards the illumination source. It in fact indicates the sequence of the change in the reflectance from dark-to-light or light-to-dark when the detected edge is crossed. This sequence should always be dark-to-light when moving towards the source. In the opposite case, detected pixels constitute transitions which correspond to channels, breaks of slope, or simply form boundaries of cast shadows and are therefore eliminated. Figure 4 illustrates the ridges detected from the synthetic image. Angles φ and θ were set to 135° and 45° respectively for the illumination of the model.

2.3 Post Processing of Detected Pixels

After pixels lying on ridges have been detected on both the terrain model and the digital image, some further processing is needed in order to connect these points into meaningful lines. More precisely, while the outputs of the ridge detection process are bilevel images the

input to the matching algorithm must constitute two sets of line segments. Post processing comprises the transformation of a bitmap of ridges to a set of line segments representing the same ridges.

Since ridge detection produces clouds of ridge pixels rather than lines, a line thinning is performed first to produce lines that are one pixel thick. Such a line is formed by pixels having only two neighbours unless they are endpoints (one neighbour) or junctions (three or more neighbours). The classical thinning algorithm proposed by Pavlidis [1982] is used here. It successively erases all boundary pixels from a thick line until a skeleton of pixels remains. This skeleton is defined by connectivity criteria and represents the thin line. Since the algorithm appears to allow trimming of endpoints a version that preserves endpoints in a preliminary pass is also used. In the second step, strings of coordinates constituting individual lines are identified by line following techniques [Peuquet, 1981]. Each individual line is followed from pixel to pixel until the end of the line is reached. In order to avoid ambiguities as to which branch of the line to follow when a junction is encountered, all junctions are removed before the operation and then reintroduced during line following as endpoints of line segments. Finally the followed lines are filtered into straight line segments [Douglas and Peucker, 1973] using an offset tolerance. Furthermore, segments adjacent to the same junction are merged if they have similar orientation. This is done to avoid breaking or loss of lines as result of junction removal. A length tolerance is also introduced to eliminate very short lines from the matching process. Finally, an orientation test is carried out for ridges detected directly from elevation data, to exclude segments that are parallel to the direction of illumination.

3. MATCHING OF DIGITAL IMAGES AND TERRAIN MODELS

In the previous section it was demonstrated how two sets of line segments portraying mountain ridges can be extracted from digital images and terrain models. These line segments are represented by the coordinate values of their endpoints. In this section a matching procedure that locates corresponding ridge lines in an image and in a terrain model is presented.

Many structural matching methods use local pictorial features such as endpoints or corners as matching keys. In this project, geometric properties of neighbouring line segments are used to establish matching correspondences. This scheme is identified by Mitsuyama et al. [1984] as being less sensitive to noise fluctuation. Furthermore, the ridge structures in the two data sets may not be identical. Some may be misoriented, displaced in position or even entirely missing from one of the sets. The method allows for such cases, since it locates corresponding ridges locally rather than performing a global fit based on the entire sets [Horn and Bachman, 1978] .

Pairs and triplets of adjacent ridge line segments are selected for performing the matching. Definitions of what constitutes adjacent line segments are illustrated in Figure 5. Two segments A and B are neighbours if they can be connected with a straight line without intersecting any other segment within the quadrilateral defined by A and B. Three segments A, B, and C satisfy the adjacency criterion if A--C and C--B are neighbouring pairs and there exists a straight line which intersects only A, B, and C within the same quadrilateral. An a priori knowledge of registration parameters, which is usually available for remotely-sensed data, limits the search area for corresponding segments to windows in the data sets.

In the example presented here, ridge lines extracted from the elevation data (Figure 2) constitute Data Set 1. Ridges detected in the synthetic image are considered simulated real image data and form Data Set 2. The matching of Data Set 2 to Data Set 1 is performed in the following sequence:

(a) Pre-processing of Data Set 1. Triplets of neighbouring segments are identified interactively in Data Set 1, and their adjacency is established by intersecting straight lines, as

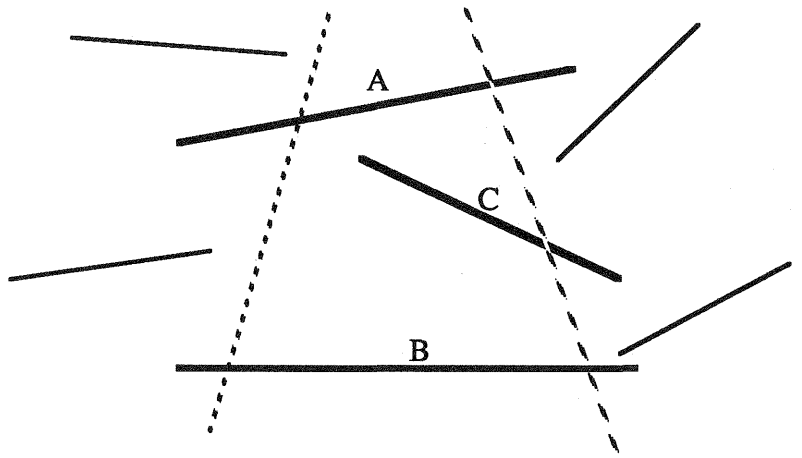

Figure 5 : Neighbouring line segments A and B and A, B, and C.

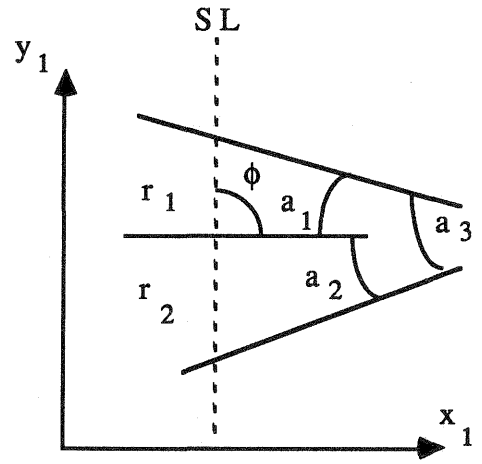

Figure 6 : Geometric properties for a set of three line segments.

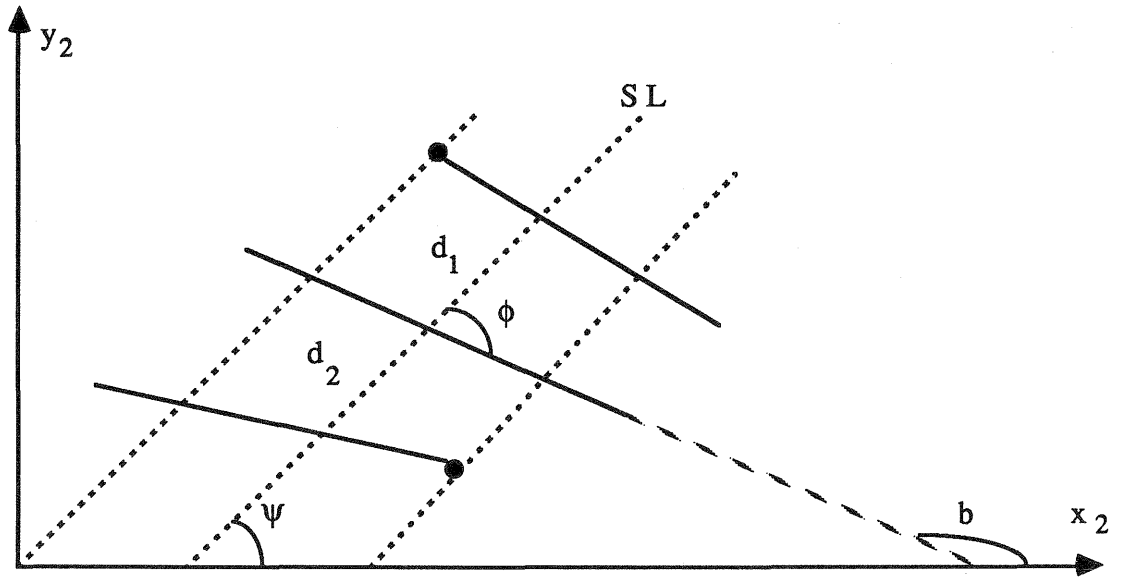

Figure 7 : Step 3 of the matching procedure.

explained above. The most prominent triplets are then selected as "control lines" for matching. This operation is analogous to control point selection in a standard geometric registration. Next, each triplet is rotated so as to bring the straight line intersecting it parallel to the y-axis. Finally, the geometric elements illustrated in Figure 6 are computed for each triplet. Angle ϕ is formed between the intersecting line and the middle segment, while a_1, a_2, a_3 are angles formed between the three line segments in all combinations. The ratio $r = r_2/r_1$ is also calculated. These five values are stored for each triplet..

(b) **Pre-processing of Data Set 2.** Pairs of two neighbouring segments are identified within Data Set 2. This is done automatically in the following way. The entire data set is subdivided into strips parallel to y-axis. These strips are defined by the x-coordinates of endpoints of all line segments within data set .The x-coordinates are ordered by their value and each pair of successive points defines a strip. For each strip, all line segments are found and sorted according to their y-coordinate in the middle of the strip. This sorting results in pairs of neighbouring segments. In order to find all possible pairs, Set 2 has to be rotated several times before applying the above procedure. In our example, however, one rotation of 90° was sufficient. Finnaly, the angle b_1, formed by the two segments, is calculated for each pair .

(c) **Matching.** The objective is to locate the corresponding line segments in Data Set 2 to each set of three line segments defined in Data Set 1. It is implemented in three steps:

Step 1. All pairs of segments formed in the pre-procesisng of Set 2 are candidates for matching. A pair is considered for further processing if

$$|a_1 - b_1| < T$$

where T is a threshold value set to 15° for the experiment.

Step 2. In this step the previously chosen segment pair is extended into a triplet of line segments. All line segments from Set 2, except the ones forming the pair in question, are considered. For each triplet, new angles between segments are calculated and compared against angles a_2 and a_3 of the set of three. If corresponding angles match within the tolerance T the triplet becomes a candidate and is considered for the final step of the procedure.

Step 3. As illustrated in Figure 7, angle ψ formed by the straight line SL of Set 1 and axis-x_2 in Set 2 is calculated. In other words, a line intersecting the middle segment of the triplet and forming an angle ϕ with it is found. Its orientation ψ (counted anticlockwise from axis-x_2) is estimated as

$$\psi = \phi + b - 180°$$

where b is the angle between the middle segment of the candidate triplet and the x_2-axis. This is done in order to define an orientation along which ratio r will be tested. The position of a line having ψ orientation and forming ratio d_2/d_1 equal to r is calculated. Segments d_1 and d_2 are defined on the line by the candidate triplet. This line should lie within the band illustrated in Figure 7 in order to intersect all three segments of the candidate triplet. If this statement holds the tested triplet corresponds exactly to the initial set of three segments from Set 1.

Step 3 proved to be a very sensitive in operation. Line segments tend to lie parallel to each other because of directional illumination. Consequently, even a small displacement among the three line segments disturbs the ratio d_2/d_1 and the requested straight line is located far outside the desired band. The advantage is however, that no false correspondences are generated as one would expect as a result of this structured orientation of ridges. The results of the experiment are shown in Figure 8. Four triplets were matched for the image. The corresponding line segments are superimposed on the contour map for an overall assessment of the described method. It is noticed that although a ridge appears broken in two parts on

the DTM while being uniform on the image, the corresponding parts are still matched correctly.

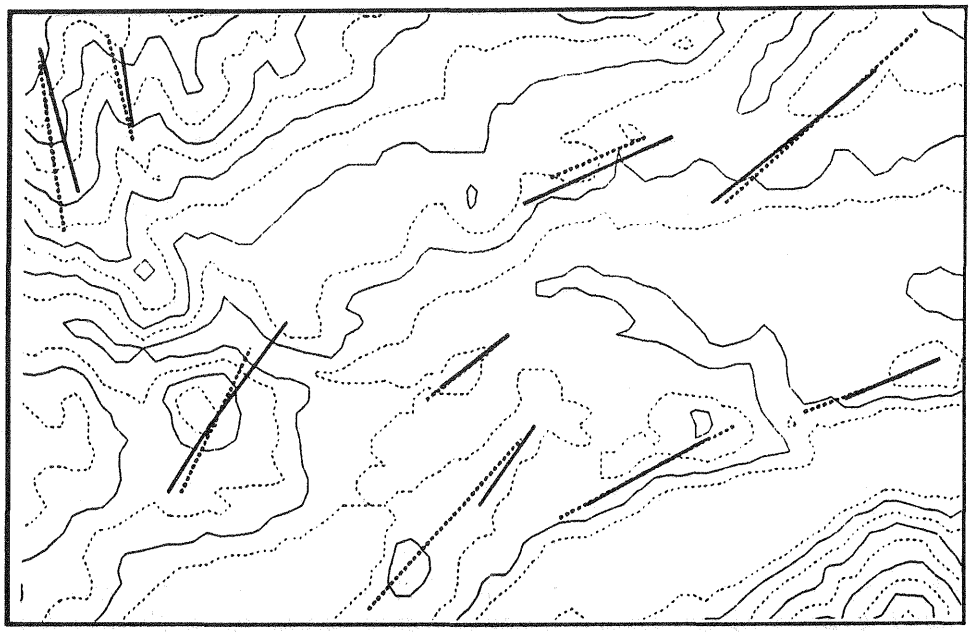

Figure 8 : Matched ridges ; solid lines represent the ones from the DTM, while dashed lines represent ridges from the image.

The final step, which is not yet implemented, involves the computation of the transformation parameters for the registration of the two data sets. Corresponding points should be defined in the matched line segments. Endpoints of segments are unsuitable for this purpose because the geometric relations used as matching keys do not depend on the length of ridges but only on their orientation and location. Intersections between the line segments and the points that produce ratios equal to r should be used instead. In case of nearly parallel segments, their intersections are left out.

4. CONCLUSIONS

This paper describes the methodology for the automatic registration of digital images on digital terrain models. The discussed procedure is feasible when the surface topography is the determining factor of image intensities. In such case mountain ridges are well detectable from the image and therefore constitute prominent features for automatic matching. Extraction of ridges from both digital image and terrain is first performed. Then the two sets of detected structures are matched with the help of geometric properties of neighbouring ridges. The results of the automatic matching between a grid elevation model and a simulated image demonstrated that:

a. The implemented methods detected mountain ridges satisfactory.

b. The matching procedure locates the corresponding features correctly. Furthermore it succeeds when the approximation of ridges has resulted into line segments of uneven number or different length.

Further experiments with real images are required to define the specific types of imagery and terrain that are most appropriate for the presented methodology.

ACKNOWLEDGEMENT

This project was supported by a Natural Sciences and Engineering Research Council, Canada and an Energy, Mines and Resources, Canada grant in aid of research.

REFERENCES

Douglas, D.H. and T.K. Peucker (1973). "Algorithms for the Reduction of the Number of Points Required to Represent a Digital Line or its Caricature." *The Canadian Cartographer*, Vol. 10, No. 2, pp. 112-122.

Douglas, D.H. (1986). "Experiments to Locate Ridges and Channels to Create a New Type of Digital Elevation Model." *Cartographica*, Vol. 23, No. 4, pp. 29-61.

Fowler, R.J. and J.J. Little (1979). "Automatic Extraction of Irregular Network Digital Terrain Models." *Proceedings of SIGGRAPH-79*, Chicago, Aug. 1979, pp. 199-207.

Gonzalez, R.C. and P. Wintz (1987). *Digital Image Processing*. 2nd Edition, Addison-Wesley Publishing Company, Don Mills, Ontario.

Horn, B.K.P. and B.L. Bachman (1978). "Using Synthetic Images to Register Real Images with Surface Models." *Communications of the ACM*, Vol. 21, No. 11, pp. 914-924.

Horn, B.K.P. (1981). "Hill Shading and the Reflectance Map." *IEEE Proceedings* , Vol. 69, pp. 14-47.

Jensen, S.K. (1985). "Automated Derivation of Hydrologic Basin Characteristics from Digital Elevation Model Data." *Digital Representations of Spatial Knowledge/ Proceedings of the Auto-Carto 7*, Washington, D.C., March 11-14, pp. 301-310.

Mitsuyama, T., H. Arita and M. Nagao (1984). "Structural Matching of Line Drawings Using the Geometric Relationship between Line Segments." *Computer Vision, Graphics and Image Processing*, Vol. 27, pp. 177-194.

Papanikolaou, K. and E.E. Derenyi (1987). "GIS in Support of Remote Sensing Technology: Present Applications, Future Possibilities." *GIS '87 San Francisco/ Second Annual International Conference, Exhibits and Workshops on Geographic Information Systems*, October 26-30, pp. 333-339.

Pavlidis, T. (1982). *Algorithms for Graphics and Image Processing*. Computer Science Press, Rockville, Maryland.

Peucker, T.K. and D.H. Douglas (1975). "Detection of Surface-Specific Points by Local Parallel Processing of Discrete Terrain." *Computer Graphics and Image Processing*, Vol. 4, pp. 375-387.

Peuquet, D.J. (1981). "An Examination of Techniques for Reformatting Digital Cartographic Data/Part 1: The Raster-to-Vector Process." *Cartographica*, Vol. 18, No. 1, pp. 34-48.