SATELLITE IMAGE REGISTRATION THROUGH STRUCTURAL MATCHING [1]

Leonardo Sant'Anna Bins
Flávio R. Dias Velasco
Instituto de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201 - São José dos Campos, SP
Brasil

COMMISSION III

1. INTRODUCTION

There are many situations when it is necessary to register two images. For example, when we need to:

a) combine images of the same scene taken with different sensors;

b) analyze temporal variations of the same scene: images are obtained with the same sensor but at different times;

c) recognize patterns: one image correspond to the scene, the other to the pattern (or model) and a match has to be found between the model and part of the image.

The techniques traditionally used for registration of satellite images are based on correlation. These techniques, however, are quite sensitive to differences in scale (resolution), radiometry, and also to geometric transformations like rotation. These restrictions and the high computational cost tend to restrict the scope of application of the correlation-based methods.

In this work we present a method for image registration that, instead of making pixel by pixel comparisons (as in the correlation-based methods), it tries to register two images through the matching of their intermediate-level structural descriptions. By imtermediate-level description we mean a description of the image in terms of homogeneous regions and linear features (objects) and their spatial relationships. The method can be made invariant to a given class of transformations simply by choosing object attribute descriptors and spatial relations that are invariant to the transformations in the class.

The organization of this paper is the following. Section 2 shows how a structural description can be extracted from an image and how two structures can be matched. Some results of image registration using the method are shown in Section 3. The final section evaluates the method and the results and describes directions for future work.

1 This work was supported by SID Informática.

2. MATCHING STRUCTURAL DESCRIPTIONS

An image can be described by the objects it contains and how these objects spatially relate to each other. An object can be either a region, i. e., a two-dimensional set of contiguous pixels or a linear feature, which is one-dimensional. The object can be characterized by its shape and spectral attributes. Some shape attributes are the size, perimeter, orientation, elongation and compactness. The various spectral values distribution moments of the points that constitute the region like mean, covariance, etc can serve as spectral attributes. To make an object description invariant to a transformation it is necessary to choose the attibutes appropriatedly. For example, size and perimeter, can not be used to make the description invariant to scaling. On the other hand, depending on how they are defined, compactness and elongation are good candidates.

Another component of the description is the spatial arrangement of the objects. The position of the objects can not be used directly for obvious reasons. However the relative positions of the objects is a characteristic that is invariant to most transformations.

A structural description of an image can be represented as an oriented graph where the the nodes correspond to the objects and the arcs to the spatial relationships. The problem is, given two descriptions, how to find a mapping between the descriptions. The mapping must associate nodes in the descriptions such that the node and its counterpart have compatible attributes and are related to similar nodes by similar relations. The mapping does not need to map every node of either graph. If more than a mapping exists, then one chooses the best one, i. e., one that maps more nodes.

The problem as stated is a particular case of a classical problem in Computer Vision: the matching of the "world" with the stored "model". Let us call "image graph" one of the descriptions and "model graph" the other one. This problem has been shown to be equivalent to the subgraph isomorphism problem (Ballard and Brown, 1982), which is an NP-complete problem.

One way of solving the matching problem involves the construction of an "association graph". In this graph, each node corresponds to a pair of compatible nodes, one from the image graph and the other from the model graph. There is an arc between two nodes (a,1) and (b,2) if the relationship between "a" and "b" is compatible with the relationship between "1" and "2". Once the association graph is built, the best mapping can be found simply by extracting its largest connected subgraph (Ambler et al., 1975).

Directly looking for the cliques in the graph can be computationally very expensive. One possibility is to use relaxation techniques (Zucker et al., 1977). To this end, we extended the association graph to include both positive and negative reinforcements. The positive links are those of the association graph. A node (a,1) negatively affects a node (b,2) if either "a" equals to "b" or "1" equals to "2". In other words, (a,1) and (a,2) are two competing interpretations of node "a".

The relaxation step proceeds as follows. To each node it is associated a real number: the node "potential". The potential is updated based on the potential of the neighboring nodes and on the kind of relationship between the nodes (if positive or negative). Initially all the nodes have a potential that is depended on the matching of the two objects that originated the node. The updating of the potential is iterated until some convergence criterion is met (for example, no significant change on the potential values). After the process has converged, nodes with high potential most probably correspond to valid mappings; those with low potential correspond to unlikely mappings. The pairs of corresponding objects provide "control points" that can be used to derive the geometric transformation that relates the two images.

3. EXPERIMENTAL RESULTS

An experiment was conducted using satellite images of Porto Alegre city, Rio Grande do Sul state, LANDSAT TM sensor. In this experiment, band 1 was registered with band 2. The first step, object extraction, was done by thresholding the image in its 10 and 90 grey level percentile, resulting in two types of objects (light and dark objects). This method of segmentation was chosen because it yields regions that are invariant (to a degree) to transformations of the grey levels. To reduce the noise in the resulting images, very small objects were ignored.

The attributes chosen were:

- object type (light or dark);
- compactness (perimeter squared divided by area);
- elongatedness (ratio between principal axes).

These attributes are invariant to simple geometric transformations like scaling, translation and rotation.

To construct the structural description of the image, also relations between the objects must be defined. These, as the object attributes, must be chosen so that they are invariant to the likely transformations between the two images. In our case the two images (different bands of the same scene) were geometrically identical. However, to test the method, we chose only relations that were invariant to the same class of tranformations as the attributes. These were the angle between objects and the size relation (larger than).

Figure 1 shows band 1 of the image; Figure 2 shows the segmented images. The segmentation of band 1 resulted in 20 objects and the segmentation of band 2 produced 21 objects. The association graph that was formed had 45 nodes, which is about 10 percent of the maximum possible size (420). After relaxation, only nine nodes "survided". Of these nine nodes, 3 corresponded to wrong associations. After filtering the wrong associations, the centers of these objects can be used as control points of the transformation that relates the two images.

The mean square error for the registration of the two images was 1.28 pixels.

4. CONCLUSIONS

In this work we proposed the use of structural description for registration of satellite images instead of the traditional method of pixel-by-pixel correlation. The main advantage of the method is that it can be made invariant to a specific class of geometric and radiometric transformations, making it suitable for the registration of images obtained in very different conditions. Although the method can be used as the only way to achieve registration, it can be used also as an initial step of a two pass process. The second pass could be a traditional correlation algorithm, to fine tune the result.

As in the case of the experiment, the relaxation step does not guarantee a unique solution, as some objects might be mapped to the wrong ones. However, as the wrong points are few, they can be rejected by procedures similar to those used to reject bad control points in image registration. It is also possible to double check the surviving nodes with more rigorous attributes.

The object attributes and relations play different roles in the method. The importance of the attributes is in reducing the size of the association graph while the relations are the main agents responsible for ambiguity elimination, cutting down the number of possible mappings.

Our future plans include using linear objects instead of two-dimensional regions. We feel that borders and lines are much more robust to radiometric changes than regions. One possible way of extracting the linear features is through the use of the Hough transform (Ballard and Brown, 1982). It is also our intention to evaluate the accuracy of the method for images obtained in different conditions.

5. REFERENCES

Ambler, A. P.; Barrow, H. G.; Brown, C. M.; Burstall, R. M.; Popplestone, R. J., "A versatile computer-controlled assembly system", Artificial Intelligence, 6(2):129-156, 1975.

Ballard, D. H.; Brown, C. M., Computer Vision, Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

Zucker, S. W.; Hummel, R.; Rosenfeld, A., An application of relaxation labeling to line and curve enhancement, IEEE Transactions on Computers, vol. C-26(4):394-403, 1977.

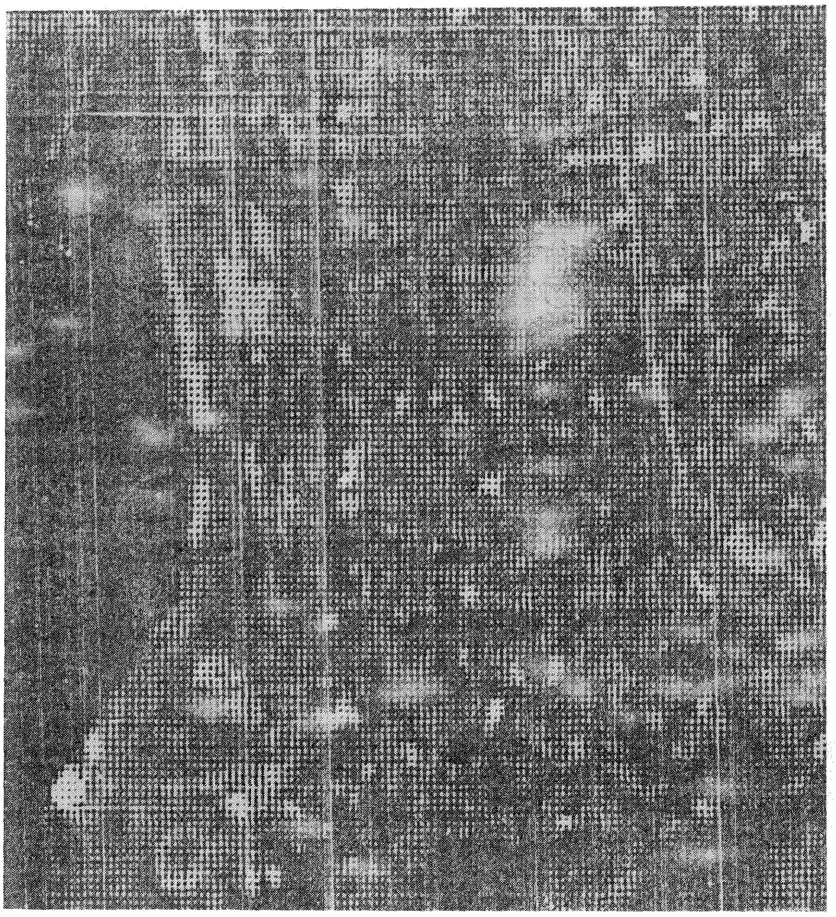

Figure 1. Original image, band 1.

Figure 2 a - Segmented image: band 1.

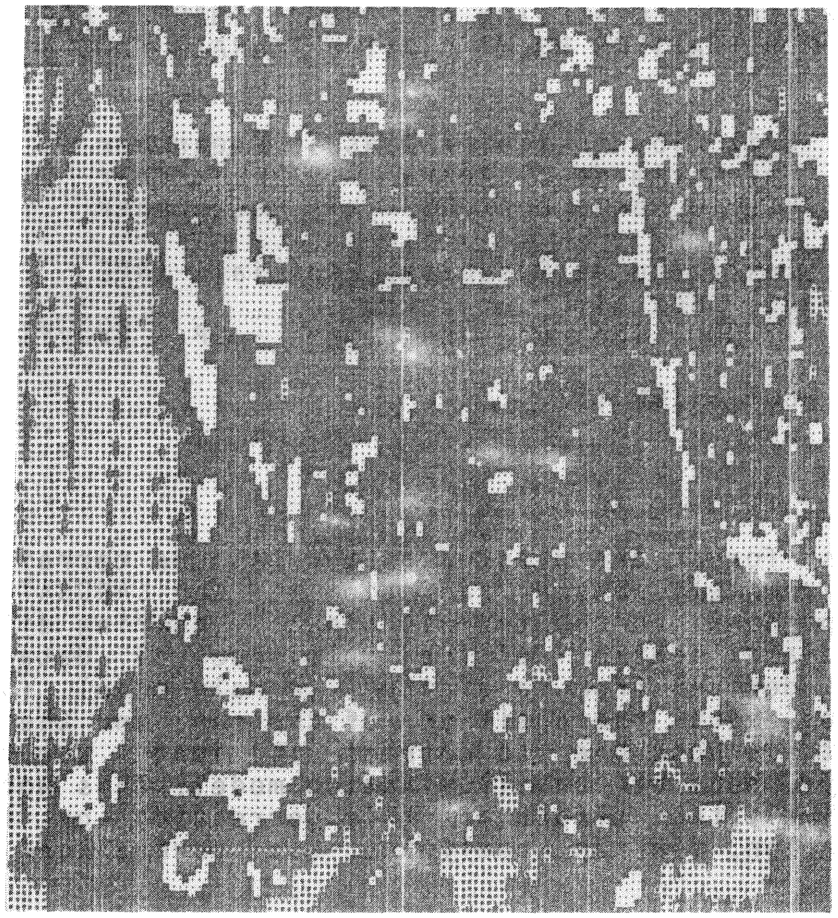

Figure 2 b - Segmented image: band 2 .