AN ATTRIBUTE–DRIVEN APPROACH FOR IMAGE REGISTRATION USING ROAD NETWORKS

Caixia Wang, Peggy Agouris, Anthony Stefanidis

Center for Earth Observing and Space Research
Department of Earth Systems and Geoinformation Sciences – (cwangg; pagouris; astefani)@ gmu.edu
George Mason University, Fairfax, VA 22030, USA

Commission IV, WG IV/2

KEY WORDS: Georeferencing, Matching, Imagery, GIS, Feature, Transformation

ABSTRACT:

Geospatial analysis is becoming increasingly dependent on the integration of data from heterogeneous sources. In this paper, we present an automated, feature-based approach for geometric co-registration using networks of roads (or other similar features). This approach is based on a graph matching scheme that models networks as graphs with embedded invariant attributes. The main advantages of our approach reside in its ability of using both geometric and topological attributes to reduce the ambiguity in search space for inexact matching as well as its invariance to translation, rotation and scale differences (through the use of appropriate attributes). Furthermore, our approach requires no user-defined threshold to justify local matches. Using the information derived from this matching process, the registration of two datasets can be accomplished.

1. INTRODUCTION

Geospatial analysis is becoming increasingly dependent on the integration of data from heterogeneous sources. The geometric co-registration of these datasets still remains a challenging and crucial task, especially given the emergence of novel data capturing approaches, like the use of unmanned aerial vehicles (UAVs) to capture long image sequences. In this context, registration may refer either to the registration of images to images, to generate for example long mosaics, or to the registration of images to maps, to identify their orientation parameters. This registration problem becomes increasingly complex when we consider differences in coverage, scale, and resolution as corresponding objects in two datasets may also differ to a certain extent.

Road networks usually are common features in areas of interest. Unlike points or point-like features e.g. manhole covers (Drewniok and Rohr, 1996) or building corners (Rohr, 2001), road networks contain inherently substantial semantic information in their structure (e.g. topology and geometry), and thus are considered robust entities for matching in our approach based on graph matching. A great deal of effort has been devoted to graph matching by the computer vision community. In the work of Barrow and Popplestone (1971), relational graph matching was first studied where a relational graph is designed to represent scene structure for matching. After that, it has been widely adopted and developed for matching problems. Two major approaches can be identified. One involves the construction of structural graph model where geometric attributes of components are not taken into consider. Matching techniques are developed solely based on structure pattern, like the graph and sub-graph isomorphism approaches (Shapiro and Haralick, 1985; Pellilo, 1999; Bunke, 1999; Jain et. al., 2002). The major drawbacks in these graph-theoretical methods are their computational complexity and inability to handle inexact matching due to noise or corruptions in the graph. Later works in Wilson and Hancock (1997) and Luo & Hancock (2001) exemplify some enhancements based on pure structural graph model. The second approach to the problem appreciates the measurements of network components and represents networks as attributed relational graphs. Matching techniques are developed to compute graph similarity based on these measurements and network relational structure, such as relaxation labelling algorithm (Rosenfeld et. al., 1976; Li,1992; Gautama & Borgharaef, 2005), information theory principles (Shi and Malik, 1998), Markov Random Field method (Li, 1994). In these approaches, invariant measurements of network components are essential for the matching as they can reduce ambiguities in local similarity and the corresponding searching space. But due to different scope of computer vision applications (e.g. face recognition, content-based image retrieval) research has addressed geometric and topological attributes of the network in a rather limited manner, focusing instead more on performance metrics (e.g. faster convergence).

In this work, we develop an efficient algorithm of inexact graph matching using invariant attributes (geometry and topology) included in geographic networks and is based on the relaxation labelling introduced by Hummel and Zucker (1983). The challenges we are facing include the computational complexity of matching network components (i.e. junctions and polygons), as well as errors in feature extraction due to the presence of noise in scenes, like building-induced shadows and occlusions. In this paper, the utilization of point networks and revised relaxation labelling provides the ability to utilize structures and geometric attributes derived from the network to improve the matching algorithm and thus achieve relatively efficient computation. The process is fully automatic in terms of no input needed from users. These unique advantages serve both as the motivation for our work and constitute the main contributions of this paper.

The remainder of the paper is organized as follows: Section 2 describes the formal representation of road networks in terms of attributed relational graph. The attributes developed for relaxation matching are described in Section 3. In Section 4, our revised relaxation labelling algorithm for matching is described in detail. Experimental results are presented in Section 5. Finally, Section 6 presents conclusion and outlines our future work.

2. NETWORK PREPARATION

Geographic features (road networks) extracted from both data sets are first transformed into graph structure as input to our approach: extracted intersections are modelled as vertices in the graph, while road segments between intersections are modelled as straight edges in the graph. The detection of intersections is not a topic addressed by this paper, as this is a well-researched topic in photogrammetry and computer vision. We assume that road intersections have been detected in both datasets being registered. Figure 1 exemplifies the transformation of the road networks in an image.

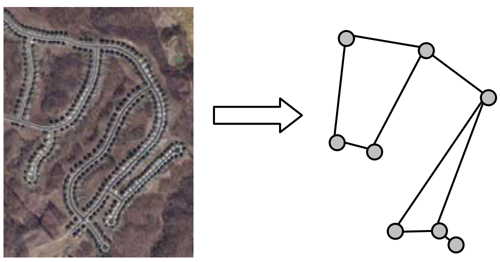

Figure 1. Graph representation of road networks on imagery

For the sake of clarity, we name the graph from image space as G_d and the one from corresponding object space as G_m. Correspondingly, V_1^d is a vertex of G_d , E_1^d an edge of G_d , V_1^m a vertex of G_m , E_1^m an edge of G_m …

3. FORMALIZATION OF INVARIANT ATTRIBUTES

Invariant attributes are essential for matching as they can reduce ambiguities in local similarity and the corresponding search space. Developing invariant attributes, however, is a non-trivial issue. In one hand, as the involved imagery and GIS datasets may differ in terms of resolution, scale, coverage, and orientation in general, the conjugate features may also differ to a certain extent. On the other hand, as road networks usually involve high volume of data, it is important to develop attributes that require less computational efforts. In this section, we introduce attributes derived from the geometry and topology of road networks, which are invariant to translations, rotations and scale changes.

We start with connectivity attribute represented by the formal adjacency matrix (noted as *A*), which can be used to model the topological structure of road networks.

Definition 1. *If there is at least one single road segment connecting road intersections i and j, i is said being connected to j (or vice versa). The entry for ij in the adjacency matrix A is of value 1. Otherwise, it is 0.*

The adjacency matrix can be derived from the graph. The entry values of the matrix correspond to the existence of edges between corresponding vertices, i.e. value 1 describes at least one edge, while value 0 represents no edge. By definition adjacency matrix is invariant with respect to translation, rotation, and even scale variations between the image and the corresponding geospatial dataset.

Typically Euclidean distance is an important measurement of the geometry. It is invariant to translations and rotations, but not to scale changes. In order to overcome this problem we use the relative distance between road intersections as a node-linking attribute (instead of Euclidean distance). Relative distance is defined as:

$$\hat{D}_{ij} = \frac{D_{ij}}{(D_{ij} + D_{it}) * 0.5} \tag{1}$$

where i, j, t = three road intersections
$\hat{D}_{ij}$ = relative distance between i and j
D_{ij} = euclidean distance between i and j
D_{it} = euclidean distance between i and t

A third attribute (basic loop attribute) can be derived from adjacency matrix. It is used to model higher network topological structures, specifically the formation of closed loops in it. In the case of networks the closed loops are of triangle, quadrangular or higher forms, and accordingly the basic loop is defined as:

Definition 2. *If vertex V_i has two adjacent (connected) vertices, each of which also has one common adjacent vertex other than V_i , V_i has one quadrangle associated to it.*

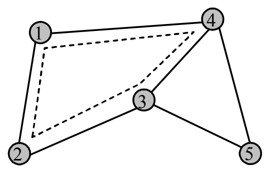

Figure 2. The quadrangle in networks

This is exemplified in Figure 2. V_1 has two adjacent vertices V_2 and V_4. Both V_2 and V_4 are adjacent to V_3. Thus, V_1 is associated to one quadrangle formed by V_1 , V_2, V_3, V_4. Same as V_4 , V_3 and V_2. As mentioned above, the property can easily be extended to more complex, polygonal loops, if so desired.

4. MATCHING TWO ROAD NETWORKS

Accordingly, the road network is defined through the graph embedded two topological (connectivity and basic loop) and one geometric attribute (relative distance). The reader can easily understand that additional attributes may also be used as needed. This type of graph is termed attributed graph here (and, similarly, attributed network). Using the above notations for these two networks, our aim in matching is to optimally correspond (label) nodes V_i^d in graph G_d to those in graph G_m satisfying certain matching criteria. The road network matching is formulated as a graph-labeling problem.

Based on relaxation labelling, the matching process iteratively re-labels the data nodes with model nodes by changing their corresponding weights. The weights are optimized according to their local geometric and topological similarity. After each iteration, the global matching (i.e. global compatibility) is measured. The process reaches an optimal matching when the global compatibility measurement becomes unchanged or varies to a limited threshold. We details the matching process in following subsections.

4.1 Local Similarity

Once we have constructed the attributed graphs from two networks we proceed with their optimal matching. Given V_k^m from G^m as the current label of V_i^d in G^d, the goodness of such mapping ($V_i^d \rightarrow V_k^m$) can be measured through a modified version of the exponential function (Li, 1992). Our aim is to iteratively re-label the nodes of the data graph with the model graph so as to optimize a global compatibility measured by the structures and geometries of matched nodes. The goodness of the local fit can be measured with H (V_i^d, V_k^m):

a) If not both V_i^d and V_k^m are associated with the basic loop

$$H(V_i^d, V_k^m) = \exp(-\Sigma \min | \hat{D}_{i,\{s,\alpha\}}{}^d - \hat{D}_{k,\{t,\tau\}}{}^m |) \tag{2}$$

Where i, s, α = road intersections in the dataset to be registered, where s & α are connected with i

k, t, τ = road intersections in model dataset, where t & τ are connected with k

$\hat{D}_{i(s,\alpha)}{}^d$ = sum of relative distances of intersections i, s and i, α

$\hat{D}_{k(t,\tau)}{}^m$ = sum of relative distances of intersections k, t and k, τ

b) If both V_i^d and V_k^m are associated with the basic loop

$$H(V_i^d, V_k^m) = \exp(-\Sigma \min | \hat{D}_{i,\{s,\alpha\}}{}^d - \hat{D}_{k,\{t,\tau\}}{}^m |) + \exp(-| \hat{D}_{(s,\alpha),j}{}^d - \hat{D}_{(t,\tau),\kappa}{}^m |) \tag{3}$$

Where i, s, α, j = road intersections in the dataset To be registered and form the basic loop of i

k, t, τ, κ = road intersections in model dataset and form the basic loop of k

$\hat{D}_{(s,\alpha),j}{}^d$ = sum of relative distances of intersections s, j and α, j

$\hat{D}_{(t,\tau),\kappa}{}^m$ = sum of relative distances of intersections t, κ and τ, κ

Use Figure 3 as an example. If we consider labelling V_2^m for V_1^d, V_1^d has two connected vertices V_2^d and V_3^d, which also both connect with vertex V_4^d. At the same time, V_2^m also has two connected vertices V_1^m and V_5^m both connecting with vertex V_4^m. In this case, H should be measured with the formula (3). If V_k^m has more than two adjacent vertices as V_1^m, we choose the two vertices that minimize the power value in function H.

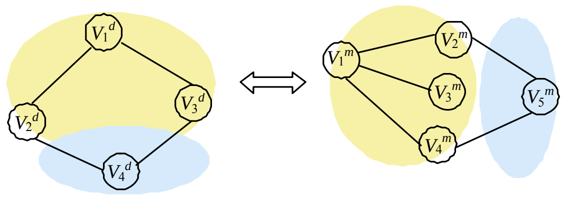

Figure 3. Vertices with inexact degrees

The novel feature of this local consistency measure H is its compound structure, which distinguishes it from many alternatives in the literature. Specifically, the geometry and topology for measuring local mapping goodness is formed by nodes both directly connected (marked with yellow circles in Figure 3) and indirectly connected (marked with blue circles in Figure 3) with the mapping nodes. The underlying advantages with these two measurements is that the constructed H function will not be affected by the presence of noise (i.e. the additional link V_3^m in Figure 3) and the ambiguity will be reduced as low as possible. Similarly, the presence of noise (i.e. additional links) in V^d would not affect our matching.

4.2 Global Compatibility

With function H, the local difference between V_k^m and V_i^d under the minimal relative distance constraint is mapped into a similarity measure for assigning V_k^m to V_i^d. As the continuous relaxation labeling framework, weighted values other than logical assertions (1 or 0) are attached to all possible assignments for each vertex in G^d. The weight (denoted by $p_i(\lambda)$) with which label V_λ^m is assigned to vertex V_i^d belongs to [0,1]. In addition, the sum of the weights for all possible assignments to any vertex should be equal to 1. Let Θ be all available assignments with V^m to V^d. The global compatibility function can be formed as:

$$\Lambda(\Theta | d, m) = \sum_{i,j} \sum_{k,\kappa} H_{ij}(V_k^m, V_\kappa^m) p_i(V_k^m) p_j(V_\kappa^m) \tag{4}$$

Thus, the optimal labeling of V^d with V^m will be the assignment that maximizes the above function. We use the gradient ascent algorithm, which iteratively computes the length and direction of the update vector to update the weight p such that the global compatibility function Λ will increase with each updating of p. The iteration terminates when the algorithm converges, generally producing an optimal labeling (or matching). Interested readers are referred to (Hummel and Zucker, 1983) for additional details.

5. EXPERIMENTS

We tested proposed approach in two experiments in order to demonstrate its performance and robustness. All tests are implemented in MATLAB environment.

5.1 Test 1

The two road networks used in this experiment are detected respectively from a map, typically having larger coverage, and from an image with smaller coverage. They are shown on the left in Figure 4, with their corresponding graphs on the right. The two networks reflect typical registration conditions, whereby an image and a corresponding map may differ substantially in terms of translations, rotations, and scale changes. It should also be noted that a link (between nodes a and e) in the map network does not exist in the image network. This introduces inexact matching in the two networks, but only in their structure.

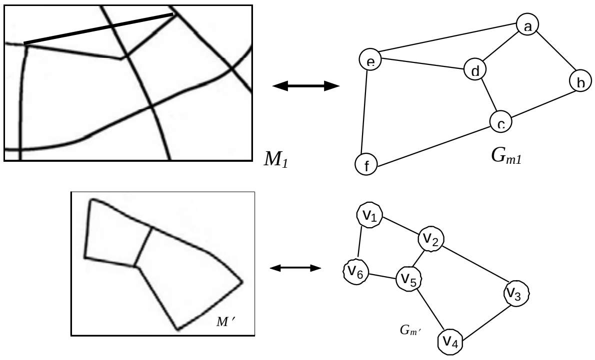

Figure 4. Detected networks and their graphs

Figure 5 illustrates the convergence of the global compatibility function under successive iterations until a maximum value is reached. The result using all three attributes is shown by the thinner curve (top) and its global compatibility increases faster and converges earlier than when using two attributes (connectivity and relative distance) only. The run time for this experiment is 1.2218 seconds (with two attributes) and 1.4821 seconds (with all three attributes).

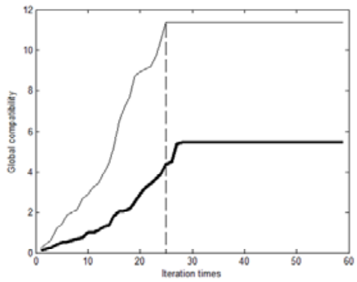

Figure 5. Comparison under inexact matching

The matching result is summarized in Table 1. It can be easily seen that all nodes were matched correctly despite differences in orientation (rotation, shift, and scales) between the two networks, or even differences in their actual structure (the presence of the a-e link).

	a	*b*	*c*	*d*	*e*	*f*
Matching result	V_6	V_1	V_2	V_5	V_4	V_3

Table 1. Matching result

5.2 Test 2

In this test, we examine the robustness of our approach in exact matching. M_2 in Figure 6 and M' in Figure 4 are two detected networks used in this experiment. It should be noted that M_2 has 9 intersections and 13 edges, while M' only has 6 intersections and 7 edges. These two datasets vary not only in structures like the example in Test 1, but also in nodes of the graph.

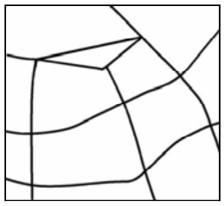

Figure 6. Detected network M_2 from the map

In addition, as shown in Figure 7, there are four components in M_2 marked with colors that have same topological pattern as M'. Thus, topological attributes only would produce multiple results.

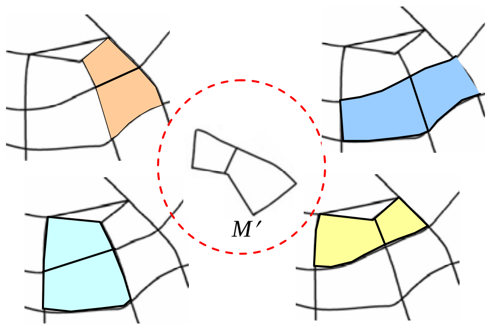

Figure 7. Topologically similar components

The convergence of the global compatibility is shown in Figure 8. Similar as Figure 5, the result using all three attributes is shown by the red curve (top) and its global compatibility increases faster and converges earlier than when using two attributes only. In this test, the matching starts to converge after 30 iterations, slower than in Test 1 as we have relatively complex networks for matching.

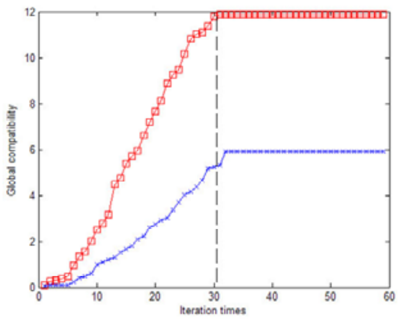

Figure 8. Global compatibility vs. iteration times

The matching result is graphically described in Figure 9. Despite the topological similarity problems shown in Figure 7, intersections in M' are correctly mapped to M_2.

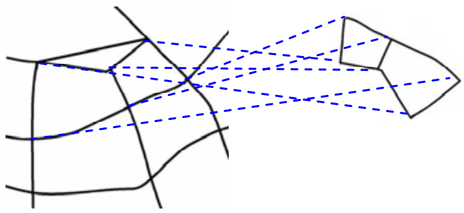

Figure 9. Matching result

6. CONCLUSION AND FUTURE WORK

This paper introduced a novel matching approach to the georegistration problem based on graph matching. It offers the ability to utilize information about the topology and geometry of a network to establish correspondence. The ability to utilize both allows us to reduce the ambiguity of local consistency, especially when inexact matching takes place. Furthermore, the approach does not require user input, other than detecting road intersections through image processing. Thus our approach offers a robust and general solution to the image-to-x registration problem using networks.

Future work will further investigate additional attributes to give rise to invariant description of patterns in networks. It will also include an extension of the proposed approach to more complex networks.

ACKNOWLEDGEMENT

This work was supported by the National Geospatial-Intelligence Agency through NURI grant NMA 401-02-1-2008.

REFERENCES

Barrow, H. G. and Popplestone, R. J., 1971. Relational descriptions in picture processing. *Machine Intelligence*, 6, Edinburgh.

Bunke, H., 1999. Error correcting graph matching: On the influence of the underlying cost function. IEEE Trans. PAMI, 21(9), pp. 917–922.

Drewniok, C. and Rohr, K., 1996. Automatic exterior orientation of aerial images in urban environments. In: Proceedings of ISPRS, Vienna, Austria, Vol. XXXI, Part B3, pp. 146-152.

Gautama S., & Borgharaef, A., 2005. Detecting change in road networks using continuous relaxation labelling. In: ISPRS Hannover Workshop 2005, Hannover.

Hummel, R.A. and Zucker, S.W., 1983. On the foundations of relaxation labeling processes. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-5(3), pp. 267-287.

Jain, B. J. and Wysotzki, F., 2002. Fast winner-takes-all networks for the maximum clique problem. In: Proceedings of the 25th Annual German Conference on AI: Advances in Artificial Intelligence, pp.163-173.

Li, S. Z., 1992. Matching: Invariant to translations, rotations and scale changes. Pattern Recognition, 25(6), pp. 583-594.

Li, S. Z., 1994. A markov random field model for object matching under contextual constraints". In Proceedings of the IEEE Computer Vision and Pattern Recognition Conference (CVPR), Seattle, Washington, pp. 866–pp.869.

Luo B. & Hancock, E. R., 2001. Structural graph matching using the EM algorithm and singular value decomposition. IEEE Trans. PAMI, 23(10), pp. 1120-1136.

Pelillo, M., 1999. Replicator equations, maximal cliques, and graph isomorphism. Neural Computation, 11(9), pp. 1933-1955.

Rohr, K., 2001. Landmark-based image analysis: using geometric and intensity models. Kluwer Academic Publisher.

Rosenfeld, A., Hummel, R. and Zucker, S., 1976. Scene labeling by relaxation operations. IEEE Trans. Syst. Man. Cybern., 6, pp. 420-433.

Shapiro, L.G. and Haralick, R.M., 1985. A metric for comparing relational descriptions. IEEE Trans. PAMI, 7(1):90–94.

Shi, J. and Malik, J., 1998. Self inducing relational distance and its application to image segmentation. Lecture Notes in Computer Science, 1406, pp. 528.

Wilson, R.C. & Hancock, E. R., 1997. Structural matching by discrete relaxation. IEEE PAMI, 19(6), pp. 634-648.