RELAXATION MATCHING FOR GEOREGISTRATION OF AERIAL AND SATELLITE IMAGERY

Caixia Wang[1] *Anthony Stefanidis*[2] *Peggy Agouris*[2]

[1]Department of Spatial Information Science and Engineering
University of Maine, Orono, ME, 04469-5711, USA

[2]Center for Earth Observing and Space Research
Department of Earth Systems and Geoinformation Sciences
George Mason University, Fairfax, VA 22030, USA

caixia@spatial.maine.edu {astefani, pagouris}@gmu.edu

ABSTRACT

In this paper, we present a novel approach for image-to-map registration through graph matching. In our graph-based matching approach road networks are presented as graphs, and registration is performed through the optimal mapping between two such graphs. A global compatibility function is formulated to measure the overall goodness of correspondence between two graphs, and is optimized through the use of continuous relaxation labeling. The main advantages of our approach lie on its invariance to translation, rotation and scale differences (through the use of appropriate attributes), as well as on its use of network structure to reduce the ambiguity in search space for inexact matching. Furthermore, our approach requires no user-defined threshold to justify local matches. In this paper we present the theoretical background behind our approach, and experimental results to demonstrate its performance.

***Index Terms*—** Graph Matching, Imagery, GIS, Feature

1. INTRODUCTION

Recent advances in commercial satellites and space-/air-borne sensors result in the availability of substantial amounts of high-resolution imagery. The georegistration of this imagery, i.e. identifying its location and orientation in space, is a fundamental operation for its subsequent exploitation in a variety of geospatial applications. Georegistration typically involves the identification of the same entities (e.g. points or lines) in the image as in a geospatial database (e.g. a map) depicting the same area.

Point-based georegistration has long been the most popular approach, primarily because point features (e.g. manhole covers [1], building corners [2]) are relatively easy to detect. However, points contain minimal semantic information, and as such point matching tends to be an error-prone process. Thus, there have been efforts to use more complex features for matching. For example, Schickler [3] used 3D wireframe building models for georegistration, an approach though that required very good approximations of the orientation parameters. [4], [5], and [6] used polygon features representing land-use classes to reconstruct the absolute orientation of aerial imagery. Road networks contain inherently substantial semantic information in their structure (e.g. their topology and geometry), and thus are considered robust matching entities.

This work uses point networks (such as roads and road networks) as matching primitives and is based on the relaxation labeling introduced by Hummel and Zucker [7]. 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. Although we have substantial efforts in the computer vision community addressing image registration as a graph matching problem ([8]; [9]; [10]), the geometry and topology of the network have not received enough attention. In this paper, the utilization of point networks and revised relaxation labeling provides the ability to utilize not only point information, which is relatively easy to detect, but also additional structures and 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 introduces the formal abstraction of road networks by using attributed graphs. The attributes that are constructed for relaxation matching are described in Section 3. In Section 4, our revised relaxation labeling algorithm for matching is described in detail. Experimental results are presented in Section 5. Finally, Section 6 presents some conclusion and outlines future work.

2. GRAPH REPRESENTATION

Road networks can be represented as graphs, with road intersections as graph vertices. Intersection detection 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 the image to be registered, and the corresponding geospatial database (i.e. a corresponding map, or even another georegistered image), and represent networks as attributed graphs. In an attributed graph a set of attributes is used to express relationships among its vertices (e.g. links between them) and properties of such relationships (e.g. distance or orientation, or even curvature of the segments linking two nodes).

The image space network can be represented as $G_d = (V_d, R_d, A_d)$. In this notation, $V_d = \{V_1^d, V_2^d, \ldots V_r^d\}$ is the set of r vertices of road intersections; $R_d = \{D_{ij}^d \mid (i,j) \in V_d^2\}$ is a set of relative distance relations, linking vertex pairs defined over $V_d^2 = V_d \times V_d$. The other attribute is tabulated in the adjacency matrix A_d, also defined over $V_d^2 = V_d \times V_d$, with entries of 1 (or 0) when a connection exists (resp. does not exist) between the two corresponding two nodes. Otherwise, when two nodes are not linked, the corresponding adjacency matrix entries are 0s. More formal definitions of those attributes are provided in the next section.

Thus, the road network is defined in this manner through a sequence of vertices, and attributes among these nodes: adjacency, distance, etc. The reader can easily understand that additional attributes may also be used as needed. Similarly, the corresponding object space network can also be defined as $G_m = (V_m, R_m, A_m)$. Using the above notations for these two networks, our aim in matching is to optimally correspond (label) nodes $V_d = \{V_1^d, V_2^d, \ldots, V_r^d\}$ in graph G_d to those from the set $V_m = \{V_1^m, V_2^m, \ldots, V_s^m\}$ in graph G_m satisfying certain matching criteria.

3. PROPERTY FORMALIZATION

With above representation, road networks from image and object space can be represented with graph structures associated with attributes. In this section, we will further give formal definitions of such attributes.

We start with adjacency, which allows the graph to model the topological structure of road networks.

Definition 1. *If there exists one road segment that directly connects road intersections i and j, i is considered to be adjacent to j and this property is represented as an edge between corresponding graph vertices V_i and V_j. The entry for ij in the adjacency matrix A is of value 1. Otherwise, it is 0, i.e.:*

$$A_{ij}^d = \begin{cases} 1 & \text{if } V_i^d \text{ is adjacent to } V_j^d \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

By definition adjacency is invariant with respect to translation, rotation, and even scale variations between the image and the corresponding geospatial dataset. Euclidean distance on the other hand is invariant to translations and rotations, but not to scale changes. In order to overcome this problem we use the relative distance as a node-linking attribute (instead of Euclidean distance). Relative distance is defined as:

$$\hat{D}_{ij} = \frac{D_{ij}}{(D_{ij} + D_{it})/2} \tag{2}$$

where D_{ij} is the Euclidean distance between vertices V_i and V_j, V_j and V_t are two vertices adjacent to current vertex V_i, that is A_{ij}=1 and A_{it}=1. As vertices in the graphs denote road intersections, every vertex will have at least two adjacent vertices. In the case of more than two adjacent to current vertex V_i, V_j and V_t in the relative distance are selected randomly from all vertices adjacent to V_i.

A third attribute (loop attribute) can also be formed related to vertices. It is used to model higher network topological structures, and specifically the formation of closed loops in it. In the case of road networks the closed loops are predominantly of quadrangular form, and accordingly this property is defined as:

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

As mentioned above, the property can easily be extended to more complex, polygonal loops, if so desired.

4. OPTIMAL MAPPING ESTIMATION

Network matching becomes a matching of attributed graphs, and we proceed using a relaxation labeling approach. 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 attributes of matched nodes.

Given V_k^m from G_m as the current label of V_i^d in G_d, let $\{s,\alpha\}$ be any two adjacent vertices of V_i^d and $\{t,\tau\}$ be any two adjacent vertices of V_k^m. The goodness of the local fit can be measured with relative distance; if the two adjacent vertices of current node have a common adjacent node (represented by j and κ separately in two graphs) other than the current node, a second relative distance is then applied:

$$H(V_i^d, V_k^m) = \begin{cases} \exp(-\sum \min \left| \hat{D}_{i,\{s,\alpha\}}{}^d - \hat{D}_{k,\{t,\tau\}}{}^m \right| & \text{if no common adjacent node} \\ \exp(-\sum \min \left| \hat{D}_{i,\{s,\alpha\}}{}^d - \hat{D}_{k,\{t,\tau\}}{}^m \right|) + \exp(-\left| \hat{D}_{ij}{}^d - \hat{D}_{k\kappa}{}^m \right|) & \text{otherwise} \end{cases} \tag{3}$$

Use Figure 1 as an example. If we consider labeling V_2^m for V_1^d, V_1^d has two adjacent vertices V_2^d and V_3^d both connecting with vertex V_4^d. At the same time, V_2^m also has two adjacent vertices V_1^m and V_5^m that connect with vertex V_4^m. In this case, H should be measured with the second form. 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.

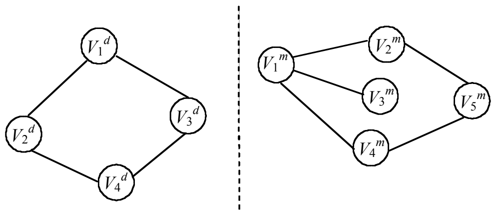

Figure 1. Vertices with inexact degrees

The novel feature of this local consistency measure H is its compound exponential structure, which distinguishes it from many alternatives in the literature. 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 1) 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.

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 with which label V_λ^m is assigned to vertex V_i^d is denoted by $p_i(\lambda)$ and satisfies:

$$0 \le p_i(\lambda) \le 1, \text{ for } \forall i \in V_d, \forall \lambda \in V_m \tag{4}$$

and

$$\sum_{\lambda = V_1^m}^{V_q^m} p_i(\lambda) = 1, \text{ for } \forall i \in V_d \tag{5}$$

Let Θ be all available assignments with V_m to V_d. The global compatibility function can be formed as:

$$\Lambda(\Theta \mid d, m) = \sum_{i,j} \sum_{k,\kappa} H_{ij}(V_k^m, V_\kappa^m) p_i(V_k^m) p_j(V_\kappa^m) \tag{6}$$

Thus, the optimal labeling of G_d with G_m will be the one (Θ*) that maximizes the above function:

$$\Lambda(\Theta^*) = \max(\Lambda) \tag{7}$$

We use the gradient ascent algorithm, which iteratively computes the length and direction of the update vector to update 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 unambiguous labeling (or matching). Interested readers are referred to [7] for additional details.

5. EXPERIMENTAL RESULTS

We implemented our graph matching approach in Matlab. In order to demonstrate the performance of proposed approach, we apply the relaxation algorithm to find the correspondences in road networks from a satellite image and a map. The two detected road networks used in this experiment are marked by M and M' below (Figure 2), where M represents the attributed graph built from the road network in the map; M' represents the attributed graph built from the road network from the satellite image and was created by subjecting M to arbitrary rotation, translation and scale changes. Thus, the two networks reflect typical registration conditions, whereby an image and a corresponding map may differ substantially in terms of these conditions. It should also be noted that we have introduced in the map network a link (between nodes a and e) that does not exist in the image network. The reason for this is explained in the following paragraphs.

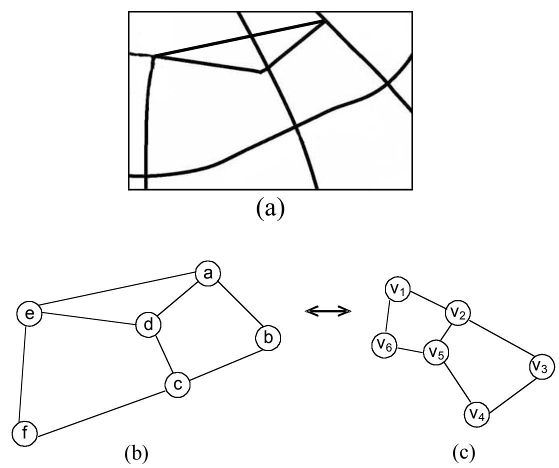

Figure 2. Experiment data: (a) road network in the map; (b) attributed graph M abstracted from the map; (c) attributed graph M' from the synthetic satellite image.

To comprehensively test the performance of our approach, we do the following two experiments:

1) *Exact matching* by assuming the link between vertex a and e does not exist in the map (or by adding a similar link between V_6 and V_4). Thus, the two graphs have the same structure. In Figure 3, the top line (linking circles) shows graph matching results using all three attributes, while the bottom curve (linking crosses) shows matching results using only two attributes: relative distance and adjacency (ignoring the loop attribute). It is easily seen that global compatibility increases faster when all three attributes are

considered, especially after tenth iteration. The run time is 0.9113 (two attributes) and 1.3219 (all three attributes).

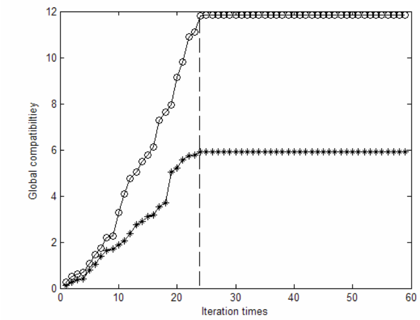

Figure 3. Comparison under exact matching

2) *Inexact matching* whereby the link between vertex *a* and *e* is allowed to participate in the map network, without having a corresponding link in the image network. Thus the two networks differ partially in terms of their structure. Matching results are shown in Figure 4. 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 only. The run time for this experiment is 1.2218 seconds (with two attributes) and 1.4821 seconds (with all three attributes).

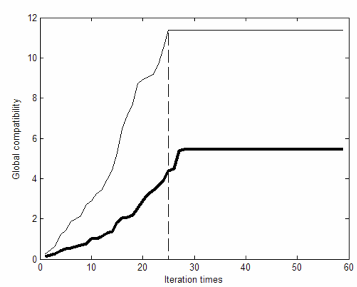

Figure 4: Comparison under inexact matching

model / data	a	b	c	d	e	f
v_1	0	1	0	0	0	0
v_2	0	0	1	0	0	0
v_3	0	0	0	0	0	1
v_4	0	0	0	0	1	0
v_5	0	0	0	1	0	0
v_6	1	0	0	0	0	0

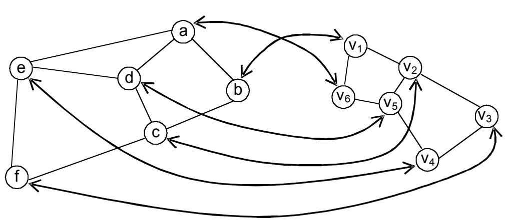

Figure 5. Road network matching result

The matching result is visualized in Figure 5, where the correspondences between nodes as a result of the network matching are shown by arrows. 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). The results in Figs. 3 and 4 also demonstrate the importance of the addition

6. CONCLUSION AND FUTURE WORK

This paper introduced a novel matching approach to the georegistration problem, which offers the ability to utilize information about the topology and geometry of a network in order 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.

ACKNOWLEDGEMENTS

This work was supported by the National Geospatial-Intelligence Agency through grant NMA 401-02-1-2008, and by the National Science Foundation through grant ITR-0121269. We also acknowledge CEOSR's VAccess Mid-Atlantic Geospatial Consortium, funded by NASA.

7. REFERENCES

[1] C. Drewniok & K. Rohr. Automatic exterior orientation of aerial images in urban environment, *Intern. Archives of Photogrammetry & Remote Sensing*, 30: 146-152, 1996.

[2] K. Rohr. Landmark-based image analysis: using geometric and intensity models. Kluwer, 2001.

[3] W. Schickler. Feature matching for outer orientation of single images using 3-D wireframe control points, *Intern. Archives of Photogrammetry & Remote Sensing*, Vol. 29, 1994.

[4] A. Morgado & I. Dowman. A procedure for automatic absolute orientation using aerial photographs and a map, *ISPRS Journ. of Photogrammetry & Remote Sensing*, 52: 169-182, 1997.

[5] I. Dowman. Automating image registration and absolute orientation: solutions and problems, *Photogrammetric Record*, vol. 16, pp. 5-18, 1998.

[6] H. Hild, N. Haala, & D. Fritsch. A strategy for automatic image to map registration, *Intern. Archives of Photogrammetry and Remote Sensing*, vol. 33, pp. 313-321, 2000.

[7] R. A. Hummel & S. W. Zucker. On the foundations of relaxation labeling processes. *IEEE PAMI*, 5(3): 267-287, 1983.

[8] R. C. Wilson & E. R. Hancock. Structural matching by discrete relaxation. *IEEE PAMI*, 19(6): 634-648, 1997.

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