Modeling Multi-object Spatial Relationships for Satellite Image Database Indexing and Retrieval

Grant Scott, Matt Klaric, and Chi-Ren Shyu

Department of Computer Science, University of Missouri-Columbia
{grantscott, mnkkrc, shyuc}@missouri.edu

Abstract. Geospatial information analysts are interested in spatial configurations of objects in satellite imagery and, more importantly, the ability to search a large-scale database of satellite images using spatial configurations as the query mechanism. In this paper we present a new method to model spatial relationships among sets of three or more objects in satellite images for scene indexing and retrieval by generating discrete spatial signatures. The proposed method is highly insensitive to scaling, rotation, and translation of the spatial configuration. Additionally, the method is efficient for use in real-time applications, such as online satellite image retrievals. Moreover, the number of objects in a spatial configuration has minimal effect on the efficiency of the method.

1 Introduction

Satellite images are playing an important role in many applications, such as environmental study and homeland security. For example, in the context of geospatial intelligence, query methods that provide selections of objects in a query image and retrieve images with similar spatial relationship among objects will greatly assist analysts to have deeper understanding of relevant geospatial information.

Traditional approaches in spatial indexing of image objects normally partition images into several bounding rectangles or spheres to describe the locations of the extracted objects. This approach includes R-tree [1], R^{+}-tree, R^{*}-tree [2], and bounding spheres SS-tree [3]. These indexing methods are designed specifically for the purpose of localizing objects of interest in an image when the locations of the extracted objects are invariant to scaling, translation, or rotation.

In addition to the traditional spatial indexing approaches, many CBIR researchers have made significant contributions to the modeling of spatial relationships for image retrieval [4][5][6]. The SaFe system developed by Smith and Chang [4] applied a 2-D string approach [7] to capture spatial relations, e.g., *adjacency, nearness, overlap*, and *surround*. Shyu and Matsakis [8] applied histogram of forces [9] to model spatial relationships between lesions and anatomical landmarks on medical images. A domain-independent technique presented by Natsev et al. [10] uses sliding windows within an image to capture the relationship among neighboring objects and extracts signatures from it. This method may not be robust enough to identify the variety of inter-object relationships

W.-K. Leow et al. (Eds.): CIVR 2005, LNCS 3568, pp. 247–256, 2005.

found in the geospatial domain because the sliding window technique may suffer from sensitivity to rotation. In [11], Matsakis et al. provide a method to determine affinity applied to a spatial object configuration to produce another view of the configurations.

Recently, many prominent content-based image retrieval (CBIR) approaches for satellite image databases have been implemented with certain degrees of success. The work presented by Datcu et al. [12] shows the results of breaking an image into regions for the purpose of classification, but not for the analysis of inter-object relationships. While, such techniques are useful for being able to identify general classes of inter-object relationships, they lack the ability to search for many specific relationships within geospatial imagery. A pair-wise, single-object, query-by-shape method, defined by Dell'Acqua [13], uses the point diffusion technique for efficient object comparisons in remote sensing images. Additionally, the work presented by Prasher and Zhou [14] highlights an efficient scheme for encoding the spatial relationships of objects. The paper demonstrates sensitivity to object translation and rotation. Bian and Xie [15] use geographic properties to model global object dependence. However, none of the aforementioned approaches address the issue of modeling multi-object spatial relationships to provide query methods that allow users to select a set of objects from a satellite image and retrieve database images with similar spatial configuration. In this paper, we provide a method to extract spatial information that is highly insensitive to rotation, scaling, and translation of configuration, with applicability in an efficient indexing structure for fast online retrievals.

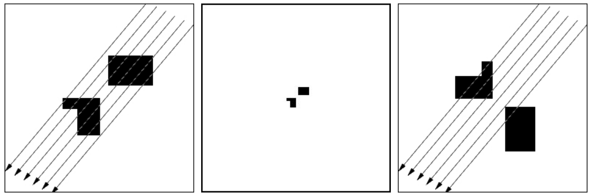
This paper is organized as follows. In Section 2 we briefly review the concepts of modeling the spatial relationship between two objects using histogram of forces, then detail our method of extending this to $n - tuple$ object configurations for spatial signature extraction. We provide results from experiments on collections of real satellite image object sets in Section 3. Finally, we offer some insights and future discussions in Section 4 along with our concluding remarks.

2 Generating Spatial Features of Multiple Objects

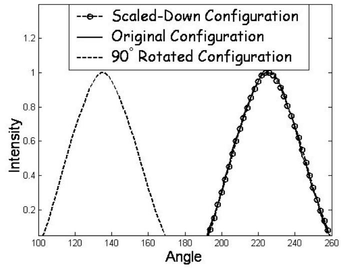
In this paper, we operate under the assumption that relevant objects can successfully be extracted from image scenes. The extraction and grouping of objects into configurations may be fully automated, require a human-in-the-loop, or be manually performed. This section details the concepts used to generate a spatial signature of an object configuration by extending the pair-wise determination of spatial relationships through application of histogram of forces.

2.1 Histogram of Forces

The histogram of forces, introduced by Matsakis and Wendling [9, 16], is a method for computing the spatial relationship between a pair of objects. A collection of parallel directional lines is conceptually rotated 360° through an image. Along each angle, Θ, parallel lines may cross the two objects, designated as referent and argument, to form longitudinal segments. The histogram of forces



(a) (b) (c)



(d)

Fig. 1. Histogram of Forces: For each angle, Θ, longitudinal raster segments of referent and argument objects are used to measure force between a pair of objects. (a) two objects with maximal forces when $\Theta = 226°$, (b) scaled-down objects, (c) after 90° rotation of the original objects, no force accumulated at 226°, (d) histogram of forces for (a)-(c)

is calculated by rotating around the centroid of the referent object, and measuring forces between the argument and referent. A measure of forces is calculated between the two objects from segments of the two objects which occupy the same parallel line. All of the parallel lines together provide the response for the given Θ component of the histogram of forces. Figure 1(a), depicts two objects that have a distinct spatial relationship, namely the rectangle is above and to the right of the other object. The resulting histogram of forces, Fig 1(d), peeks at angle Θ equal to 226°. The histogram peeks when the longitudinal segments exert maximal force. As the angle Θ changes, less of the parallel lines pass through both objects, thereby decreasing the intensity of response in the force histogram. A notable property of a normalized histogram of force is scaling invariance. This quality is demonstrated by the overlapping response peaked at 226° from Fig 1(a) and 1(b) depicted in Fig 1(d). Readers are encouraged to examine [16] for detailed treatment of the histogram of forces.

2.2 Multiple Object Configuration Problem

The spatial relationship between any two objects can be efficiently represented by histogram of forces. However, representing configurations of n objects using pair-wise spatial relationships introduces scalability concerns from an information indexing perspective. To extend histogram of forces to configurations of more than two objects, there are a few important considerations. The first issue to address for multiple object configurations is representing the spatial relationship independent of the number of objects. Some approaches, such as [17], encode the relationship between every pair of objects. However, as the number of objects in the spatial configuration increases, the lack of scalability becomes evident. For example, when comparing two k-object configurations $k(k-1)/2$ histograms are used to represent each configuration. Correlating sets of histograms between k-object configurations becomes a significant computational task equivalent to graph matching. Applying pair-wise histogram of forces is complicated by the possibility that an object may be the referent object during one analysis, but the argument object in a subsequent analysis. The second, and perhaps more challenging issue is to automatically generate signatures which are approximately rotationally invariant. Such signatures should be indexable and searchable without human intervention. Consider again Fig. 1(a) and 1(c); if this image was rotated any significant amount, such as 90°, the spatial relationship has an entirely new representation in the histogram of forces, depicted in Fig 1(d). Preferably, for a given spatial configuration of objects the signature generated should be approximately equal in the spatial feature space for any rotation, scaling, or translation.

One could possibly extend the F-signature concept proposed by Wending et al. [18] to obtain features from the union of multiple objects by computing histogram of forces from a set of objects to themselves. However, to our knowledge, there is still no automatic algorithm to make the F-signature indexable and searchable without rotating the signatures in increments of a certain angle during each comparison. At this moment, such a signature is still inapplicable to real-time large-scale satellite image database retrievals for spatial configuration queries.

2.3 Extraction of Spatial Signatures

Our approach to model spatial configurations for multiple objects is to develop spatial signatures which are insensitive to rotation, translation, and scaling. We devised a method to use a synthetic reference object to obtain a spatial signature that is unaffected by the ordered consideration of objects. This object is placed outside of the configuration, allowing it to simultaneously capture spatial features relative to each member of the spatial configuration. Given a reference object, A, and the spatial relationships to members of an object configuration, a portion of the spatial information describing the intra-object relationships is encoded. To maintain order invariance, we conceptually treat all members of the spatial configuration as a single disjoint object. This leads to a force histogram from the reference object that spans no more than 180°. If a fixed image position is utilized for the synthetic reference object position, configuration rotation and translation both drastically affect the spatial signature. To obtain a rotation insensitive feature set, we must simply place the reference object in a position relative to the spatial configuration. A natural approach to this task is the application of principal component analysis. In the case of our spatial configurations of image objects, we treat each pixel position that is a member of an object as a sample point in 2D space. This allows the calculation of the centroid of the spatial configuration. Given the centroid and the pixel samples, a covariance matrix (Σ) can be computed. Knowing that our spatial configurations always exist in the simple 2D coordinate plane, the eigenvalues and first eigenvector are

$$\lambda = \frac{tr(\Sigma) \pm \sqrt{tr(\Sigma)^2 - 4 * det(\Sigma)}}{2} \tag{1}$$

$$e_1 = \left(1, \frac{-\sigma_{01}}{\sigma_{11} - \lambda_{max}}\right) \tag{2}$$

where $tr(\Sigma)$ and $det(\Sigma)$ are the trace and determinant of the covariance matrix, respectively. Equation (2) represents the direction of the principal axis of the spatial configuration, in (y,x) order, where σ_{01} and σ_{11} are elements of Σ, and λ_{max} is the larger eigenvalue from (1). We position our rotation invariant reference object along this principal axis. The distance along the axis is determined by finding the radius of the smallest circle positioned at the configuration centroid and bounding the entire configuration. Figure 2 shows the placement of our reference object, A, outside the bounding circle.

In (2), the y component is always fixed to a positive value. This causes a reference object to rotate through a 180° arc along the bounding circle when the configuration rotates. To achieve full 360° rotational invariance, two reference objects must be used and the resultant features merged. Figure 2 depicts the placement of the dual reference objects. We construct our spatial signatures by calculating the histogram of forces, H_{+y} and H_{-y}, for each reference object against the object configuration. The histogram generated from each reference object is then aligned to the principal axis of the configuration. After this alignment, two windows up to 180°, W_{+y} and W_{-y}, centered at the principal axis are constructed from H_{+y} and H_{-y}. Each W is partitioned into F bins, and each bin generates a feature value which is the average response from H over that bin. In our experiments, we chose an F value of 20. This results in each feature

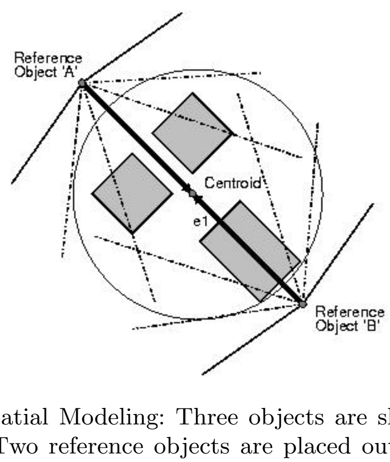


Fig. 2. Multi-Object Spatial Modeling: Three objects are shown with their centroid and principal axis, e_1. Two reference objects are placed outside the bounding circle along e_1, equidistant from the centroid

being an average response across 8°. From both reference objects we compute $2 * F$ histogram response features for a spatial configuration. To ensure that the features are rotationally insensitive, we order each bin, $i \in [1, F]$, from W_{+y} and W_{-y}, such that

$$S[i] = \max\{W_{+y}[i], W_{-y}[i]\} \tag{3}$$

$$S[i+F] = \min\{W_{+y}[i], W_{-y}[i]\} \tag{4}$$

where S is the spatial signature from the object configuration and $W_{+y}[i]$ and $W_{-y}[i]$ represent bin i in from W_{+y} and W_{-y}, respectively. As a final step, the spatial signature, $S[i]$, is normalized to $[0, 1]$, and a final feature is added to represent density of the objects within our earlier defined bounding circle.

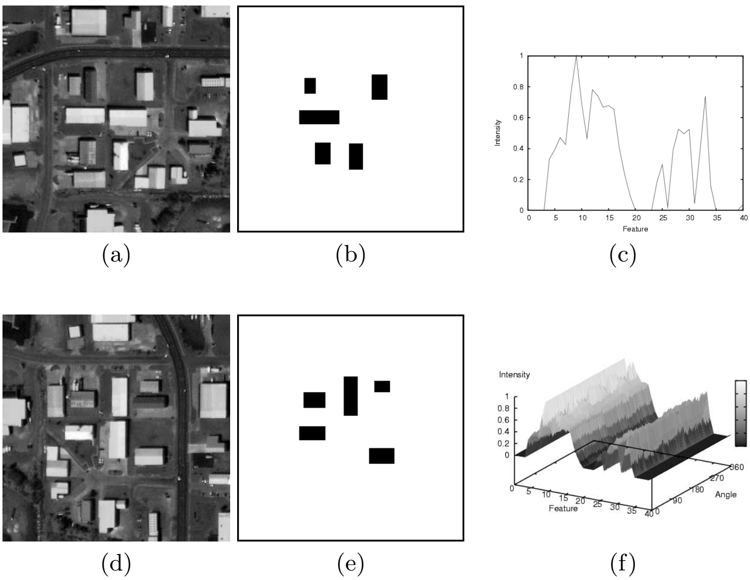


Fig. 3. Rotation Insensitive Multi-Object Spatial Modeling: (a) represents an original panchromatic IKONOS image; (b) the spatial configuration of interest, extracted from (a); (c) the spatial signature extracted from (b); (d) the panchromatic image rotated 90° clock-wise; (e) the spatial configuration of interest, extracted from (d); (f) a surface plot of the spatial signatures generated as the configuration is rotated through 360°

3 Experimental Results

The algorithms presented above were applied to sets of objects identified from several scenes of satellite imagery at 1-m resolution. From these images, groups of three and four objects were analyzed using our presented algorithm, yielding features representing spatial relationships among objects for more than 10,000 groups of objects. Although the number of configurations analyzed may be very

Table 1. Average Recall of ***Rotated*** Configurations: Ten object configurations were rotated between 0 and 360 degrees at 5 degree increments for a total of 720 sets. The values shown indicate the average recall at rank n in the results. Recall at rank n is calculated as the percent of the expected configurations correctly returned in the top n results. Recall of ***Scaled*** Configurations: Ten object configurations were scaled to ten different image sizes for a total of 100 sets. The values shown indicate the recall at rank n in the results

	Rank	1	2	3	4	5	10
Recall %	Rotation	82	91	97	98	98	100
	Scale	82	97	99	100	-	-

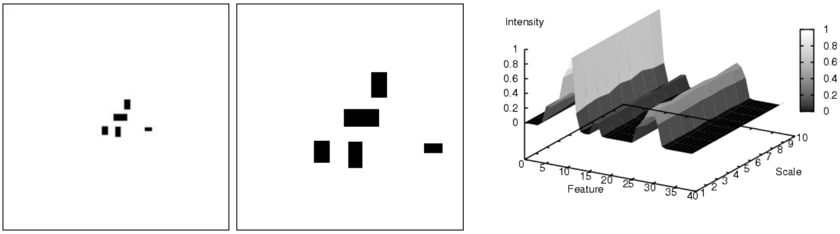
large, the time required to extract features from a configuration is well under one second for a 256x256 pixel image on a typical Pentium 4 computer.

Evaluation of the spatial configurations extracted by our approach is a very subjective task. To address this issue, the queries selected for evaluation were morphed from their original configuration using object scaling, rotation and translation. For a given query using a morphed configuration, we expect the original configuration as the top ranked result. To demonstrate that our approach has the ability to detect a wide variety of spatial relationships among objects, we require an equally varied set of test data for evaluation. For example, some of the relationships include linear configurations, triangular layouts, L-shaped patterns, etc.

The first test seeks to verify the claim that the algorithm presented is insensitive to rotation. This is measured by generating test queries resulting from rotating several selected sets of objects. Ten object sets were chosen and each was rotated between 0 and 360 degrees at 5 degree intervals. The resulting configurations were then analyzed using our algorithm and used as a query in our indexing system. The results shown in Table 1 validate the claim that the presented method is insensitive to rotation. Theoretically, we expect spatial signatures to be rotationally invariant, however, due to the raster arrangement of image pixels, small variations occur. In our collection many arrangements have high similarity in spatial configuration; from our observations, in queries where the expected configuration was not returned at rank one, configurations with highly similar configurations were returned.

The second experiment serves to confirm the scaling insensitivity of the algorithm. In this experiment, the image dimensions were varied from 100x100 to 1000x1000 pixels. Again, each of these resulting images was analyzed and fed into our information retrieval system as a query. Once again, the results shown in Table 1 validate the claim that the algorithm is scale insensitive. Figure 4(a) and (b) depicts a single object configuration at two scales. As seen in the surface plot in Fig. 4(c), the spatial signature is nearly identical at each scale analyzed for this configuration.

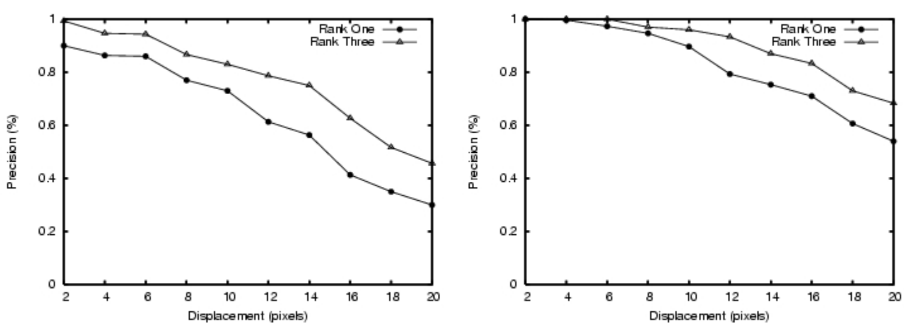
The final experiment translated a single object of the configuration by varying distances. These translations are performed to evaluate the robustness of our approach with regard to subtle changes in spatial configuration. In this test,



(a) (b) (c)

Fig. 4. Scale Insensitive Multi-Object Spatial Modeling: (a) A spatial configuration of five objects scaled to 100x100 pixels, (b) original 256x256 pixel image, and (c) the surface plot of the spatial signature across 10 image sizes from 100x100 to 1000x1000 pixel image scenes

only one object is translated at a time – this is due to the fact that translating multiple objects at the same time in different directions may dramatically alter the spatial relationships among the objects. Again, at this stage ten sets of objects were used for evaluation. In the first iteration of this stage, each object is translated a distance of two pixels; distance is measured by the Manhattan distance measure. This is repeated a total of ten times for a total number of 30 morphed configurations. For the next iteration, the distance of translation is increased by two pixels, continuing up to a total displacement of 20 pixels. This increase in translation distance allows for the similarity of signature generated to be analyzed as the original spatial configuration becomes increasingly distorted. As the amount of object displacement increases, the difference of the spatial



(a) (b)

Fig. 5. Precision of Object Translated Configurations: Ten object configurations were morphed by translating a single object by a variable displacement. The degree of displacement varied from 2 to 20 pixels. The numbers shown indicate the precision at rank 1 and rank 3. (a) Precision plots of three object configurations as a single object of the configuration is displaced by increasing amounts, (b) the equivalent evaluation of four object configurations

signature between the original and morphed configuration increases. This trend is depicted in Fig. 5.

4 Discussion and Conclusion

We propose a method for extracting signatures which represent spatial configurations of multiple objects. Our method is efficient and robust, making it applicable for real-time retrieval systems. The efficiency of our algorithm is not bound by the number of objects which compose the spatial configuration. We achieve a high degree of insensitivity to rotation of the spatial configuration due to our use of the principal components to place reference objects. As the features of our spatial signature are derived from the histogram of forces, they inherit an image scaling invariance property as well. It is noteworthy to mention that as number of objects in the configuration increases, the effect of translating a single object decreases, as shown in Fig 5. Since our spatial signature is a discrete feature vector, we plan to apply the Entropy Balanced Statistical k-d tree [19]. This will allow very large data sets of spatial configurations to be indexed for real-time satellite image retrievals. Further research will explore coupling object-based information with our spatial signatures to further refine searches.

5 Acknowledgement

This project is currently supported by the National Geospatial-Intelligence Agency University Research Initiatives(NURI) under grant number HM1582-04-1-2028. The authors would like to thank Drs. Pascal Matsakis and Ozy Sjahputera for histogram of forces source code and fruitful discussions.

References

1. Guttman, A.: R-Trees: A dynamic index structure for spatial searching. *Proc. of ACM SIGMOD*, (1984) 47–57
2. Beckmann et al.: The R^*-tree: An efficient and robust access method for points and rectangles. *Proc. of ACM SIGMOD*, (1990) 322–331
3. White, D. A., Jain, R.: Similarity indexing with the ss-tree. *Proc. 12th Intl. Conf. on Data Engineering*, (1996) 516–523
4. Smith, J.R.: Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression. *Ph.D. Thesis, Columbia University*, (1997)
5. Chu, W.W.,Hsu, C.C., Cardenas, A.F., Taira, R.K.: A knowledge-based image retrieval with spatial and temporal constructs. *IEEE Transactions on Knowledge and Data Engineering*, 10(6) (1998) 872–888
6. Li, J., Wang, J.Z., Wiederhold, G.: IRM: integrated region matching for image retrieval. *Proc. of the 8th ACM international conference on multimedia*, Los Angeles, CA, (2000) 147–156
7. Chang, S.K., Shi, Q.Y., Yan, C.Y.: Iconic indexing by 2-D strings. *IEEE Transactions on Pattern Anal. and Machine Intell.*, 9(3) (1987) 413–428

8. Shyu, C.R., Matsakis, P.: Spatial lesion indexing for medical image databases using force histograms. Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition Vol. 2, Kauai, Hawaii, December (2001) 603–608
9. Matsakis, P.: *Relations spatiales structurelles et interpretation d'images*, Ph. D. Dissertation, Institut de Recherche en Informatique de Toulouse, France, (1998)
10. Natsev, A., Rastogi, R., Shim, K.: WALRUS: A Similarity Retrieval Algorithm for Image Databases. *IEEE Trans. Knowledge and Data Engineering*, 16 (2004) 301–316
11. Matsakis, P., Keller, J.M., Sjahputera, O., Marjamaa, J.: The use of force histograms for affine-invariant relative position description. *IEEE Transactions on Pattern Anal. and Machine Intell.*, 26(1) (2004) 1–18
12. Datcu, M., et al.: Information Mining in Remote Sensing Image Archives: System Concepts. *IEEE Trans. on Geosci. Remote Scensing*, 41 December (2003) 2923–2936
13. Dell'Acqua, F., Gamba, P.: Query-by-Shape in Meteorological Image Archives Using the Point Diffusion Technique. *IEEE Trans. on Geosci. Remote Scensing*, 39 September (2001) 1834–1843
14. Prasher, S., Zhou, X.: Efficient Update and Retrieval of objects in a multiresolution geospatial database. *Proc. 15th Int. Conf. on Scientific and Statistical Database Management*, July (2003) 193–201
15. Bian, L., Xie, Z.: A spatial dependence approach to retrieving industrial complexes from digital images. *The Professional Geographer*, 56(3) (2004) 381–393
16. Matsakis, P., Wendling, L.: A New Way to Represent the Relative Position between Areal Objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(7) (1999) 634–643
17. Smith, W.F., Lam, C.P., Chen, X., Maxville, V.: Heuristics for Image Retrieval Using Spatial Configurations. *Proc. of VIIth Digital Image Computing: Techniques and Applications*, December (2003) 909–918
18. Wendling, L., Tabbone, S., Matsakis, P.: Fast and robust recognition of orbit and sinus drawings using histogram of forces. *Pattern Recognition Letters*, 23 (2002) 1687–1693
19. Scott, G., Shyu, C.R.: EBS k-d Tree: An Entropy Balanced Statistical k-d Tree for Image Databases with Ground-Truth Labels. *Proceedings of the International Conference of Image and Video Retrieval, Lecture Notes in Computer Science*, 2728 (2003) 467–476