An Integrated Approach to Retrieving Industrial Building Complexes from High-Resolution Images

Ling Bian, Jonghoon Lee

Department of Geography, State University of New York at Buffalo,
Amherst, NY 14261, USA - lbian@geog.buffalo.edu, jl227@buffalo.edu

KEY WORDS: Content Based Image Retrieval, Discrete Urban Objects, Semivariogram, Wavelet Transform, Spatial Continuity and Discontinuity

ABSTRACT:

This paper reports results of retrieving industrial building complexes from high-resolution images. Content-based image retrieval uses a sample image of specific content to retrieve images in a database that contain similar contents. The success of geographic image retrieval relies on the effective representation of geographic entities. Discrete urban objects, most typically man made structures, exhibit spatial continuity internally and identifiable boundaries externally. A semivariogram and a wavelet method are used to represent the spatial continuity and discontinuity properties associated with industrial complexes, respectively. The two methods are used individually and in combination for the retrieval. Results show that the integrated method notably improves retrieval precision over either one method when used individually, while the overall performance is encouraging for all three methods.

1. INTRODUCTION

The new generation, high-resolution satellite images, such as IKONOS, present both opportunities and challenges to urban remote sensing. Due to the high spatial resolution of these images and the spatial scale and arrangement of urban entities, a number of research issues arise. These include, for example, the urban entities discernable from these images, the properties important to identify these urban entities, and the methods appropriate to represent these properties in order to identify these entities (Barnsley et al. 2000).

This paper reports results of retrieving industrial building complexes from high-resolution images. The retrieval is based on spatial continuity and discontinuity properties associated with discrete urban objects. Two methods, a semivariogram method and a wavelet method, are used to represent the spatial continuity and discontinuity properties, respectively. The two methods were applied individually and in combination for the intended retrieval. Before a detailed explanation of the two methods, the nature of image retrieval, discernable geographic entities, and study area are first discussed below.

1.1 Image Retrieval

Image retrieval uses a sample image of specific content to retrieve images in a database that contain similar contents. Image retrieval is in many ways similar to text retrieval, such as that on the Internet. However, unlike text retrieval that uses a text sample to retrieval texts that are identical to the sample, image retrieval uses an image sample to search for images that are similar, rather than identical, to the sample. This presents a number of challenges for image retrieval, and the foremost is how to represent contents of an image.

From the perspective of geographic entity identification, two types of entities are of concern, continuous fields and discrete objects. Continuous fields occupy an extended area in images and normally exhibit rich texture pattern. General land cover types, such as forests, water surfaces, and residential areas, are typical examples of these fields. Texture-based indices have been most successful for their representation (Haralick, 1979; Manjunath and Ma, 1996; Carr and Miranda, 1998; Zhu and Yang, 1998; Sheikholeslami et al., 1999; Aksoy and Haralick, 2000). Discrete objects, most typically man-made structures, may not cover a sufficiently large area to reveal a texture pattern or they may not contain much texture information at all. Rather, they display spatial continuity internally and spatial discontinuity externally. The internal continuity indicates the body of an object, represented by a consistency in tone. The external discontinuity indicates boundaries between the object and background, represented by strong spectral contrasts. We use a semivariogram method and a wavelet method to represent these spatial continuity and discontinuity properties, respectively, through numeric indices.

Operation wise, image retrieval consists of two stages. The first stage establishes the representation of properties of geographic entities, expressed as a vector of numeric indices. These indices are computed for all database images. The second stage retrieves images from the database in response to sample images supplied by users. For each given sample image, the vector of indices is computed and subsequently compared with that of each database image. The database images are then ranked according to the similarity between them and the sample image. In this study, the similarity is computed using a measurement developed by Huang et al. (1999):

$$RL = \sum \frac{[\,|\,V_i(I_1)\ - V_i(I_2)\,|\,]}{[1 + V_i(I_1) + V_i(I_2)\,]} \qquad (1)$$

where RL is the relative similarity, expressed as a normalized difference. $V_i(I_1)$ and $V_i(I_2)$) are ith index in the index vector V for a sample images I_1 and a database image I_2, respectively. Those images that are highly ranked are retrieved. To evaluate the retrieval performance for this study, the retrieved images are compared with a master map that delineates all industrial complexes in the study area. The retrieval performance is expressed by precision, which is the ratio of the number of correctly retrieved images over the total number of retrieved images.

1.2 Study Area and Data

The image database used in this study contains 120 digital aerial photographs of the City of Buffalo, NY, at a map scale of 1:6,000. These photographs are tiled into 5635 images of 128x128-pixeleach. The size of 128x128 pixels is selected to include major parts of a complex, if not the entire complex, in an image. The tiling is designed such that each adjacent pair, in both the vertical and the horizontal directions, overlaps by 50%. This is to ensure the capture of a complex within an image while minimizing the computation burden. The industrial buildings were initially established for heavy industry operations. Although most of the operations have left the area, the building complexes remain as prominent features in the city landscape. These discrete urban objects represent a type of geographic entity that is much less studied than the general land cover types. Eight sample images are used to represent the building complexes for the intended retrieval (Figure 1).

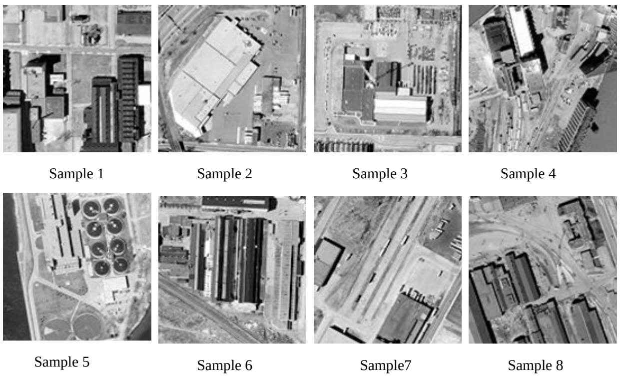

Figure 1. The eight sample images of traditional industrial complexes in the study area.

2. SEMIVARIOGRAM, WAVELET, AND INTEGRATED METHOD

2.1 Semivariogram Method

Semivariogram has been commonly used to describe spatial continuity of geographic phenomena. A semivariogram can be expressed as:

$$\gamma(\mathrm{h}) = \frac{1}{2\mathrm{N}} \sum \left(Z_x - Z_{x+h} \right)^2 \qquad (2)$$

Where γ(h) is semivariance. Z_x and Z_{x+h} are values of two locations that are separated by a lag distance h. N is the total number of location pairs being compared. The semivariance is computed for an increasing lag distance h. The resultant series of semivariances can be plotted against the corresponding lag distances to form a semivariogram. The semivariance tends to increase with the lag distance up to a certain lag distance. Beyond this distance, called range, the semivariance no longer responds to the increasing lag distance. This range distance is found sensitive to the size of dominant features in an image (Lam and Quattrochi, 1992). The semivariogram may vary in different directions and this anisotropic nature of spatial continuity can help detect the shape of an object. In addition, the amplitude of semivariances reflects the spectral contrast between an object and its background. It is argued that the basic form and anisotropy of semivariogram can help identify geographic entities (Carr and Miranda, 1998; Bian and Xie, 2004).

A number of indices are used in this study to represent the form and anisotropy of a semivariogram, according to methods described in Bian and Xie (2004). Semivariograms at eight sets of 16 directions are computed, with an angular resolution of 360/16 degrees, to represent the anisotropy of semivariograms. For a given direction, eight semivariances at lag distances of 1, 3, 6, 10, 15, 25, 40, 60 pixels, respectively, are used to capture the form of a semivariogram. The eight directional semivariograms are sorted according to their degree of spatial continuity. A direction of a greater degree of spatial continuity is usually associated with lower semivariance values than other directions. The purpose of sorting is to compare objects according to their direction of spatial continuity and independently of their azimuthal orientation in an image. The aforementioned procedure results in a vector of 64 indices for spatial continuity, eight semivariances for each of the eight directions. The 64-index vector of each sample image is compared with that of each database image to determine the similarity between them. Subsequently the top 50 images in response to each sample image are retrieved. Figure 2 displays the retrieval precision of the top 10 retrieved images in response to each sample image.

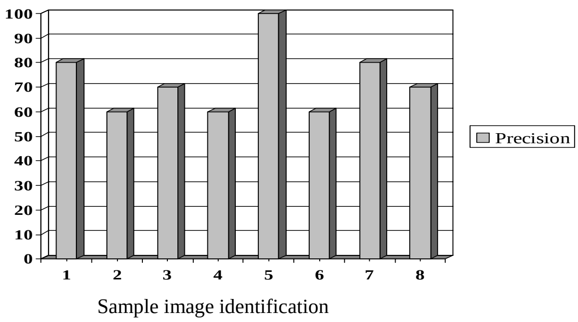

Figure 2. The retrieval precision of the top 10 images in response to each sample image using the wavelet method.

2.2 Wavelet Method

The wavelet transform primarily performs a filter function to an image. For a simple Harr wavelet transform used in this study, a one-level transform produces an approximation sub-image and three detail sub-images. The approximation sub-image represents the major trend of the original image. For geographic images, the detail sub-images contain mostly edges in horizontal, vertical, and diagonal directions, respectively. This filter process can be applied again to the approximation sub-image, resulting in four additional, next level sub-images. Thus, wavelet transform can represent two sets of information, trend and details, at multiple scales. The information in sub-images is usually summarized by statistical indices to represent the underlying properties of geographic entities (Manjunath and Ma, 1996; Sheikholeslami et al., 1999; Bian, 2003). In this study, we use the edge information in detail sub-images at multiple scales to represent the spatial discontinuity property associated with industrial building complexes.

The edge information is represented according to methods described in Bian (2003) and summarized as follows. The sub-images of wavelet are derived from an average filter. That is, every (non-overlapping) pair of adjacent pixels is averaged and the average values are recorded in the approximation sub-image. The differences between the original and the average values are recorded in the detail sub-images. Three statistics, mean, standard deviation, and number of edge pixels are computed for each detail sub-image. The former two indices are expected to represent the magnitude and range of spectral contrast, respectively, relevant to edges. The extraction of edge pixels involves a parameter, ranging between 0 and 1, to control the amount of edge pixels. In this study, a value of 0.7 is chosen to capture the few but prominent edges characteristic of industrial building complexes. Three levels of wavelet transform are derived to best distinguish industrial complexes from other urban objects. In addition to three indices for each of the three detail sub-images at each of the three levels, the mean and standard deviation are also computed for the approximation sub-image at the third level. These lead to 29 indices for spatial discontinuity. Figure 3 displays the retrieval precision of the top 10 images in response to each sample image using the 29-index vector.

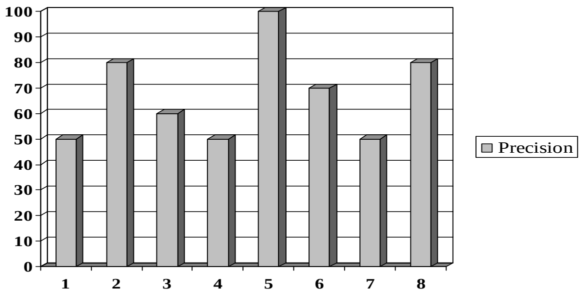

Figure 3. The retrieval precision of the top 10 images in response to each sample image using the semivariogram method.

2.3 Integrated Method

The motivation to integrate the semivariogram and wavelet methods is to devise a comprehensive representation for geographic entities. This is because both spatial continuity and discontinuity properties are intrinsic to geographic entities and they coexist. Neither approach alone can effectively represent the other kind of information. Several approaches to integration have been reported in the literature. A common approach, the index-level integration, is used in this study. This approach integrates two or more index vectors to form a hybrid index vector in order to improve the retrieval performance. Because one index vector usually represents one property, this approach assumes that all properties involved are equally important in the representation of image contents. Each index vector thus carries an equal weight in the integration. This approach has been used in many successful image retrieval systems, although mostly for non-geographic images (Flickner et al., 1995; Bach et al., 1996; Pentland et al., 1996). In this study, the 64-index vector for spatial continuity and the 29-index vector for spatial discontinuity are integrated to form a 93-index vector for retrieving industrial building complexes. In order to compare the integrated method with the two individual methods, the cumulative retrieval precision of the top 50 images, averaged over the eight sample images, is computed for each method (Figure 4).

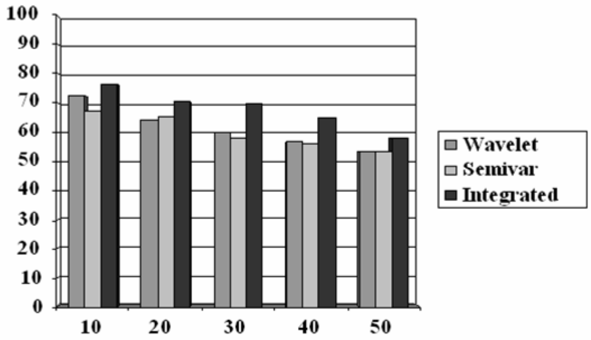

Figure 4. The cumulative, average retrieval precision of the top 50 images for the wavelet, semivariogram, and integrated method.

3. DISCUSSION AND CONCLUSIONS

First, the precision values shown in Figure 4 for all three methods are promising in comparison to other reports on geographic image retrieval (Manjunath and Ma, 1996; Carr and Miranda, 1998; Bian, 2003; Bian and Xie, 2004). It seems that the two sets of indices that represent spatial continuity and discontinuity, respectively, are adequately robust to represent the essential properties of the target objects. Second, as indicated in Figure 4, the cumulative retrieval precision decreases steadily with the increasing number of retrieved images for all three methods. This shows the effectiveness of the chosen indices because higher ranked images (top n images) are expected to better match to the sample images (higher precision). Third and lastly, the integrated method consistently improves the retrieval precision over the two individual methods by a notable margin. This result is encouraging given the rather simple integration method used in this study. Between the two individual methods, the wavelet method performs slightly better than the semivariogram method for industrial building complexes.

REFERENCE

Aksoy, S., and Haralick, R.M. 2000. Using Texture in Image Similarity and Retrieval. *Texture Analysis in Machine Vision* (M. Pietikainen, editor), World Scientific, Hackensack, NJ, pp.129-149.

Bach, J.C., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R. and Shu, C. 1996. The Virage image search engine: An open framework for image management. *Proceedings of Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases IV*, February 1996, San Jose, California (SPIE: The International Society for Optical Engineering, Bellingham, Washington), pp.76-87.

Barnsley, M.J., Moller-Jensen, L., and Barr, S.L., 2000. Inferring urban land use by spatial and structural pattern recognition, In Donnay J., M.J. Barnsley, and P.A. Longley (editors). *Remote Sensing and Urban Analysis*, London: Taylor and Francis, 115-144.

Bian, L., 2003. Retrieving urban objects using a wavelet transform approach. *Photogrammetric Engineering and Remote Sensing*, 69(2): 133-141.

Bian, L., and Xie, Z. 2004. A spatial dependence approach to retrieving industrial complexes from digital images. *Professional Geographer*, 56(3): 381-393.

Carr, J.R., and Miranda, F.P. 1998. Semivariogram in comparison to the co-occurrence matrix for classification of image texture. *IEEE Transactions on Geoscience and Remote Sensing*, 36(6): 1945-1952

Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P. 1995. Query by image and video content: the QBIC system. *IEEE Computer*, 28(9): 23-32.

Haralick, R.M., 1979. Statistical and structural approaches to texture. *Photogrammetric Engineering and Remote Sensing*, 49(1): 55-64.

Huang, J., Kumar, R., Mitra, M., Zhu, W. and Zabih, R.1999. Spatial color indexing and applications. *International Journal of Computer Vision*, 35(3): 245-268.

Lam, N.S.N., and Quattrochi, D. 1992. On the issues of scale, resolution, and fractal analysis in the mapping sciences. *Professional Geographer*, 44(1): 88-97.

Manjunath, B.S., and Ma, W.Y.1996. Texture features for browsing and retrieval of image data. IEEE Transactions on *Pattern Analysis and Machine Intelligence*, 18(8): 837-841.

Pentland, A., Picard, R.W. and Scarloff, S. 1996. Photobook: Content based manipulation of image databases. *International Journal of Computer Vision*, 18(3): 233-254.

Sheikholeslami, G., Zhang, A. and Bian, L. 1999. A multi-resolution content-based retrieval approach for geographic images. *GeoInformatica*, 3(2): 109-139.

Zhu, C., and X. Yang, 1998. Study of remote sensing image texture analysis and classification using wavelet. *International Journal of Remote Sensing*, 19(16): 3197-3203.