Adaptive Image Classification for Aerial Photo Image Retrieval

Sung Wook Baik [1] and Ran Baik [2]

[1] College of Electronics and Information Engineering, Sejong University,
Seoul 143-747, KOREA
sbaik@sejong.ac.kr
[2] Department of Computer Engineering, Honam University,
Gwangju 506-090, KOREA
baik@honam.ac.kr

Abstract. The paper presents a content based image retrieval approach with adaptive and intelligent image classification through on-line model modification. It supports geographical image retrieval over digitized historical aerial photographs in a digital library. Since the historical aerial photographs are gray-scaled and low-resolution images, image retrieval is achieved on the basis of texture feature extraction. Feature extraction methods for geographical image retrieval are Gabor spectral filtering, Laws' energy filtering, and Wavelet transformation, which are all the most widely used in image classification and segmentation. Adaptive image classification supports effective content based image retrieval through composite classifier models dealing with multi-modal feature distribution. The image retrieval methods presented in the paper are evaluated over a test bed of 184 aerial photographs. The experimental results also show the performance of different feature extraction methods for each image retrieval method.

1 Introduction

A digital archive of aerial photography is very useful to environmental scientists as well as business and governmental agencies for the purpose of environmental evaluation, land development planning, land use analysis, and so on. As an example, a collection of historical aerial image photographs can be used for chronologically tracking urban and rural development when it has been digitalized and archived for automatic access into a digital library. The content-based image retrieval in digital libraries helps to relieve the tedious work of manually finding the geographical region of interest. Content-based image retrieval requires the integration of image processing and information retrieval technologies. This is the retrieval of images on the basis of features automatically derived from the images themselves. Texture, color and shape are used the most widely in most researches to describe features in the image. However, the retrieval of the historical aerial image photographs is based only on texture features because they are gray-scaled and low-resolution images. Therefore, more ro-

G.I. Webb and Xinghuo Yu (Eds.): AI 2004, LNAI 3339, pp. 132–139, 2004.

bust feature extraction methods are required to allow effective retrieval results. The feature extraction is described in the next section.

In low resolution aerial images, we cannot apprehend the appearance of a certain objects in detail. Therefore, we need to use regions with a collection of tiny and complicated structures--such as man-made features including buildings, roads, parking lots, airports and bridges--in order to deal with them as texture motifs for image classification/segmentation. We can also regard the shapeless regions of natural resources such as forests, rivers and oceans as texture motifs.

This work presents a texture-based geographical image retrieval system with adaptive and intelligent image classification through on-line model modification. It provides geographical image retrieval over a test bed of 184 aerial photographs ranging from 350 to 650 Kbytes in size.

2 Texture Feature Extraction for Aerial Images Retrieval

Directionality, coarseness, and regularity of patterns appearing on an aerial image are represented by texture. To represent geographical features for aerial image retrieval, there are three popular texture feature extraction methods such as Gabor spectral filtering [1], Laws' energy filtering [2], and Wavelet Transformation [3], all of which have been widely used for various classification and image segmentation tasks.

Gabor filters are useful for dealing with the texture characterized by local frequency and orientation information. Gabor filters are obtained through a systematic mathematical approach. A Gabor function consists of a sinusoidal plane of particular frequency and orientation modulated by a two-dimensional Gaussian envelope. A two-dimensional Gabor filter is given by:

$$G(x, y) = \exp[\frac{1}{2}(\frac{x}{\sigma_x^2} + \frac{y}{\sigma_y^2})]\cos(\frac{2\pi x}{n_0} + \alpha) \qquad (1)$$

By orienting the sinusoid at an angle α and changing the frequency n_0, many Gabor filtering sets can be obtained. An example of a set of eight Gabor filters is decided with different parameter values (n_0 = 2.82 and 5.66 pixels/cycle and orientations α = 0°, 45°, 90°, and 135°).

Table 1. Laws' Filter Weights and Specifications

Specification	Weight of filter
L5 (Level)	(1,4,6,4,1)
E5 (Edge)	(-1,-2,0,2,1)
S5 (Spot)	(-1,0,2,0,-1)
R5 (Ripple)	(1,-4,6,-4,1)
W5 (Wave)	(-1,2,0,-2,1)

Laws' convolution kernels based on five dimensional vectors are used as an energy filter bank. It consists of 25 filters (Table 1), which can be derived from their weights. Convolving and transposing each other produces various square masks of 25 filters. Each filter is 5x5 matrices and is designed as 5x5 windows. All Laws' masks are directional and illuminant tilt sensitive except L5L5.

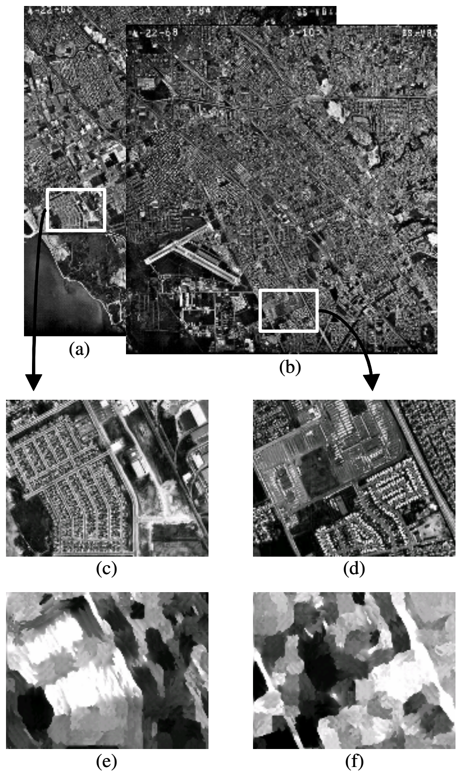

Fig. 1. a) A large aerial image including a query image. b) A large aerial image including a retrieved image. c) A query image. d) A retrieved image. e) Texture features of the query image in (c). f) Texture features of the retrieved image in (d)

A texture feature extraction algorithm based on the wavelet transform provides a non redundant signal representation with accurate reconstruction capability and forms a precise and uniform framework for signal analysis at different scales [4]. The pyramidal wavelet transform is used because of its non data redundancy and less complexity. Basically, in the pyramidal wavelet transform, the original image is decomposed into four sub-images, which are one approximation (LL) and three details (LH, HL, HH) frequency components at each level. The HH, LH, HL and LL sub-images represents diagonal details (higher frequencies in both directions, corners), vertical higher frequencies (horizontal edges), horizontal higher frequencies (vertical edges) and lowest frequencies, respectively [5]. The decomposition procedures are performed repeatedly on an approximation component at each level, and hence 3n+1 numbers of sub-images are produced for ‘n’ level decompositions. Thus, many wavelet transform sub-images can be achieved from different levels, and the variance of each sub-image is used as a texture feature [6]. However, the most significant information appears in middle frequency regions for texture images [7], LH and HL sub-images are selected from each decomposition level to compute the channel variances of a feature image. Since there is no criterion to determine the decomposition level that yields the best discriminations, it is necessary to define the desired (optimal) level. In practice, deeper level decompositions could not contain significant information and will give unreliable data.

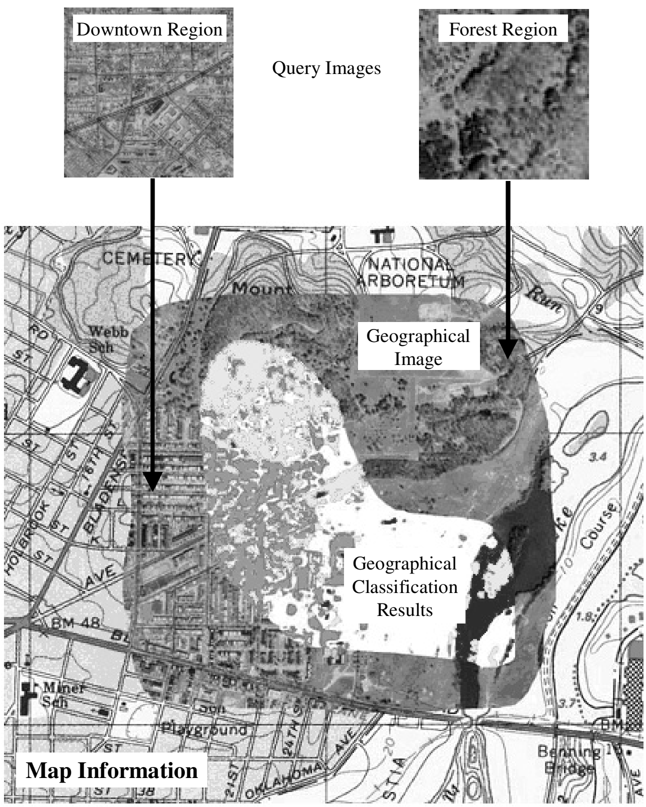

Fig.2. An example of feature classification by RBF models

3 Image Retrieval Using Adaptive Image Classification

This section describes how we adopted a Radial Basis Function based neural network classifier for image retrieval. Radial Basis Function classifiers have been used to model image feature distributions for a variety of research objectives such as image classification and segmentation. In previous researches, we can refer to some methods for modeling image feature data for image retrieval. One method [8] is to build a hybrid neural network for clustering texture patterns in the feature space. In its training phase, feature distributions are partitioned into several clusters through the Kohonen feature map. And then the captured clusters become associated with class labels using a winner-takes-all representation, and class boundaries are finally decided using a learning vector quantization scheme. Another method [9, 10] is to model texture feature data with the Gaussian mixture model, which consists of several Gaussian distribution components corresponding to texture clusters.

A modified Radial Basis Function classier (RBF) [11], with Gaussian distribution as a basis, was chosen for texture data modeling and classification. This is a well-known classifier widely used in pattern recognition and suited for engineering applications. Its well-defined mathematical model allows for further modifications and on-line manipulation with its structure and parameters. It can be easily implemented as a neural network [12]. The RBF classifier models a complex multi-modal data distribution through its decomposition into multiple independent Gaussians. The model can be dynamically adjusted by changing mode parameters over time. Data classification provides a class membership along with a confidence measure of that membership. Two types of model modification [13] are introduced for this work; 1) Self node modification that does not change/shift the node center and adjusts function spread through the feature space, 2) Node Generation that creates a new node due to the increase in the multi-modality of data distribution.

A RBF function F_r consists of a set of basis functions that form localized decision regions. Overlapping local regions formed by simple basis functions can create a complex distribution. For Gaussian distribution, as a basis function, each region is represented by its center and width corresponding to a mean vector and a covariance matrix (μ,Σ). For a multi-modal distribution of a class r, a RBF can be formed through the following linear combination of these basis functions;

$$F_r(X)\ w_0 + \sum_i w_i f_{ri}(X) \tag{2}$$

where: w_i is the trainable weight vector (for $i = 0,\ldots, N_r$); r is the class membership number; N_r is the number of nodes (basis functions) in class r ; and

$$f(X) = \exp[-1/2(X-\mu)^T \Sigma^{-1}(X-\mu)] \tag{3}$$

Each group of nodes corresponds to a different class. The combination of nodes is weighted. Each node is a Gaussian function with a trainable mean vector and a co-

variance matrix. Classification decision yields a class r of the highest $F_r(X)$ value for a sample vector X.

The image retrieval algorithm using adaptive image classification is summarized as follows:

1. Choose the best (primary) filter from a filter bank, which represents a salient feature (Fig. 1(e)) within the query image (Fig. 1(c)). We can obtain a segmented and homogeneous region with the salient feature represented by the filter if the salient feature partially dominates the query image.
2. Find similar images (Fig. 1(d)) to the query image with a threshold for similarity measurement if a dominant region (Fig. 1(f)) in each image is detected when the image is convolved by the primary filter of the query image. When the threshold is low, we can retrieve many candidate images. A collection of these candidate images is an intermediary retrieval result obtained by the primary filter.
3. Select some filters from the bank to complement the primary filter according to a feature reference table in order to analyze the query image in detail. These filters are called a secondary filter set, with which it is possible to represent a variety of features with regard to the segmented homogeneous region.
4. Collect the feature sample vectors from the segmented region according to the results obtained by the convolution with the secondary filter set.
5. Model feature sample vectors in the RBF classifier.
6. Match the candidate images in the intermediary result with the RBF classifier models. Fig. 2 shows an example of feature classification by RBF models. A collection of the matched images is the result obtained by the simple RBF classification method.
7. Modify the RBF classifier to retrieve more images through model modification. Model modification is achieved through the combination of the current models and the feature distributions of the images matched in step 6.
8. Match the candidate images in the intermediary result with the RBF classifier models once again. A collection of the matched images is the result obtained by composite RBF models.

4 Experiments

Experimental data are provided by the UC Berkeley Library Web [14]. They are 184 aerial photographs of the San Francisco Bay area, California, flown in April, 1968 by the U.S. Geological Survey. The scale of the originals is 1:30,000. Each photograph image has the size of approximately 1300 X 1500 pixels with 256 grey-level (the size and resolution of each image are little different from each other). It is cut into about 195 (13 X 15) overlapped sub-images (texture tiles) of size 200 X 200. A test bed for image retrieval has about 35,880 (184 X 195) texture tiles. For evaluation purposes, 30 types of visually similar patterns are provided by human observers. From 10 to 50 texture tiles are also selected for each pattern according to human visual experiences and indexed for retrieval performance evaluation, during which the rest of the images not selected are also used together with the indexed images.

We evaluate the performance of our image retrieval method with different feature extraction methods: 1) Gabor filter bank, 2) Law's energy filter bank, and 3) Wavelet transform filter bank. The Gabor and Law's banks consist of 48 Gabor filters (12 orientations and 4 scales in the frequency domain) and 25 filters, respectively. The Wavelet bank has 24 wavelet filters generated by using three Daubechies wavelets (db1, db2, db3) and five biorthogonal wavelets (bior1.3, bior2.4, bior3.7, bior4.4, bior5.5) at one, two and three scale decomposition.

Table 2 summarizes the performance of the aerial image retrieval methods according to several experimental results.

Table 2. The performance of the aerial image retrieval methods

	Gabor		Law		Wavelet	
	Recall	Precision	Recall	Precision	Recall	Precision
Primary Filter	75%	23.5%	71.4%	20%	69.2%	21%
Simple RBF Model	68%	55.5%	63%	52%	61.6%	53.7%
Composite RBF Model	81.5%	54.5%	78.3%	51.6%	76.2%	51.3%

5 Conclusion

This paper proposed a texture-based geographical image retrieval system with adaptive and intelligent image classification through on-line model modification. Through extensive performance comparisons under several experiments with different feature extraction methods and different retrieval methods, we show that the adaptive image classification improved the average recall-precision rates on aerial images over the simple image retrieval approach. We also show that the Gabor bank outperforms the other banks (Law's and Wavelet banks).

Acknowledgement. This work was supported by a Korea Research Foundation Grant (KRF-2003-003-D00407).

References

1. L. Chen, G. Lu and D. Zhang, Effects of Different Gabor Filter Parameters on Image Retrieval by Texture, Proceedings of the 10th International Multimedia Modeling Conference, pp. 273-278, 2004.
2. A. Gasteratos, P. Zafeiridis, I. Andreadis, An Intelligent System for Aerial Image Retrieval and Classification, LNCS, Vol. 3025, pp. 63-71, 2004.
3. B. Zhang, C. I. Tomai and A. Zhang, An Adaptive Texture Image Retrieval System Using Wavelets, Proceeding of the ICARCV International Conference, Vol. 3, pp. 1210-1215, 2002.
4. M. Unser, Texture classification and segmentation using wavelet frames, IEEE Transactions on Image Processing, Vol. 4, Issue. 11, pp. 1549-1560, 1995.

5. S. Mallat, Multifrequency channel decompositions of images and wavelet models, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 37, Issue. 12, pp. 2091-2110, 1989.
6. C. Chen, Filtering methods for texture discrimination, Pattern Recognition Letters, Vol. 20, pp. 783-790, 1999.
7. T. Chang and C. Kuo, A wavelet transform approach to texture analysis, Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 4, pp. 661-664, 1992.
8. B. Zhu, M. Ransey and H. Chen, Creating a Large-Scale Content-Based Airphoto Image Digital Library, IEEE Transactions on Image Processing, Vol. 9, Issue. 1, pp. 163-167, 2000.
9. S. Bhagavathy, S. Newsam and B. S. Manjunath , Modeling Object Classes in Aerial Image Using Texture Motifs, Proceedings of Pattern Recognition 16th International Conference, Vol. 2, pp. 981-984, 2002.
10. C. Carson, M. Thomas, S. Belongie, J. M. Jellerstein, and J. Malik, Blobworld: a System for Region-based Image Indexing and Retrieval, Proceedings of the third International Conference on Visual Information Systems, pp. 509-516, 1999.
11. S. Theodoridis, and K. Koutroumbas, Pattern Recognition, Academic Press, 1999.
12. S. Haykin, Neural Networks, Prentice Hall, 2nd edition, 1999.
13. S. W. Baik and P. Pachowicz, On-Line Model Modification Methodology for Adaptive Texture Recognition, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 32, Issue. 7, 2002.
14. http://sunsite.berkeley.edu/AerialPhotos/vbzj.html#index