Shape Similarities Measures Based on Two-level ARG for the Retrieval of Remote Sensing Images

Ming Li, Hua Qin, and Weidong Sun
Department of Electronic Engineering
Tsinghua University, Beijing, China,
Emails: liming2000@mails.tsinghua.edu.cn; qinhua00@mails.tsinghua.edu.cn; wdsun@tsinghua.edu.cn.

Abstract
Content-based image retrieval is widely investigated in image database area. However, work on content-based retrieval for remote sensing images (RSI) is seldom reported in literature. In this paper, a new strategy for content-based remote sensing images retrieval based on shape similarities measures has been proposed, which includes the shape description of each object and spatial relational representation between objects based on the two-level attributed relational graph (ARG). We have also showed that Zernike moments are very effective in describing shape features of each region, and the two-level ARG is efficient for spatial relational information of each disjoint region of an object in RSI.

As one of the applications, we have applied this method in our network-oriented multiple satellites images management and online distribution system, and the cloud objects of NOAA AVHRR images are used for verifications. An actual case for the descriptions of a cloud object with Zernike moments, the representations of relations with a two-level ARG and the shape similarity measure results in the database is given.

1 Introduction
With both scientific and commercial applications, the role of content-based image retrieval is quickly expanding, not only to provide the user with a more friendly interface, but also to find ways to browse large databases more naturally.

To promote the researches on the environment, climate and global changes, and to provide the information distribution and retrieval of RSI via internet, a multiple satellites (GMS, NOAA, and FY1C) images receiving, processing, archiving and management system is developed in our laboratory. Through this network-oriented multiple satellites images management and online distribution service, the users can retrieve their desirable satellite images by the following ways. 1) based on the time and satellite type. 2) based on the geographical regions by geographical range selections on the map or user’s inputted latitude/longitude ranges. 3) based on the features coverage percentage, area, and location, here we regard the features as cloud, ice/snow, vegetation, water. 4) based on the shape similarities including feature shapes and spatial relation matching evaluations by user’s selecting image as a sample to be matched. For this service, shape is an important feature to be considered. In particular, the shape of the objects appearing in remote sensing images might deserve more attention than it actually received so far. Many data analyses are based on shape (road extraction, cloud, ice/snow, water, and vegetation analysis) and many true targets or useless data sets may be recognized by looking at their shapes. We thus expect significant improvements in many situations by adding shape features to the set of search features. The problem is which kind of shape description and spatial relational representation methods of remote sensing images could be useful for our retrieval system.

Shape similarity evaluation has been widely analyzed in computer vision to recognize objects in a scene, and in image databases to retrieve images with the same subject. There are two types of shape descriptors: contour-based shape descriptors and region-based shape descriptors. Contour-based shape descriptors may not be suitable for complex shapes that consist of several disjoint regions such as cloud, vegetation, snow/ice, and etc. Moreover, since they are often based on a single contour, most contour-based shape descriptors, such as Fourier descriptors, chain code-based descriptors, are not appropriate for describing shapes consisting of several separated regions.

Region-based shape descriptors, such as moments, are more reliable for shapes that have complex boundaries, because they rely not only on the contour pixels but also on all pixels constituting the shapes. The drawback of regular moments is that there exists redundant information in the moments since the bases are not orthogonal and high-order moments are sensitive to noise [10].

However, Zernike moment descriptor has many desirable properties such as rotation invariance, robustness to noise, expression efficiency, fast computation and multi-level representation for describing the various shapes of patterns. With a proper normalization method, scale and translation invariance can also be achieved [8], [10]. Therefore Zernike moment descriptor is very suitable for our system requirement.

Due to Zernike moments, we can easily describe each separated region of an object such as cloud, vegetation, and etc. But an object consists of several disjoint regions, and the spatial relation among regions in remote sensing images is very important information for recognition and interpretation tasks. So we should consider spatial arrangements for each disjoint region of an object. Attributed relational graphs have been widely used to represent the spatial relation between object models. The attributed relational graph has been defined by Fu...
and Tsai [11]. It has been used for representing objects or scenes and studies extensively in image understanding. Based on matching of two ARG’s, similarities between two objects or two scenes can be measured. Since the matching is a process of pairing nodes of two ARG’s that may be similar or different greatly in structure, many researchers have proposed various methods for graph matching. Some of them are called structural matching or inexact matching. Here, we especially defined the two-level ARG for remote sensing images. The difference with the classical ARG is that we define the area and Zernike moments vector as each node attribute and the distances between nodes as each edge attribute. The area of each node (such as a cloud region) is the first low level information, describing the primitive property of nodes, and the distance is also low level information for spatial structure between nodes. The Zernike moments vector is the second high level information, which represents the detailed content of nodes.

The organization of this paper is as follows. In section II, preprocessing for remote sensing images is discussed. A brief review of the Zernike moments and shape description will be the topic of Section III. Spatial relational representation by two-level attributed relational graph, shape similarities measures between two two-level ARG’s and graph matching process are presented in Section IV. In section V experiments have been conducted as specified by NOAA-14 AVHRR images, and the conclusions of our study are given.

2 Preprocessing for Remote Sensing Images

Shape pattern is an important visual feature of remote sensing images especially in classifying some specific thematic information, such as clouds, hurricanes. To realize shape feature extraction, we should preprocess remote sensing images. Five steps are included in this procedure. 1) There are potentially many sources of geometric distortion of RSI and their effects are severe. Therefore these geometric distortions must be corrected for NOAA, GMS, and FY1C images. There are two techniques that can be used to correct geometric distortion present in RSI. One is to model the nature and magnitude of the sources of distortion and use these models to establish correction formulae. The second approach depends on establishing mathematical relationships between the pixels in an image and the corresponding coordinates of those points on the ground (via a map). Here we choose the latter approach for geometric corrections of NOAA, and FY1C images. 2) For the following normalization of Zernike moments, region area, and distance between regions, these images must be mapped to the fixed geographical region. In our system, we select the geographical ranges (65-145E, 5-55N) as our observed region for NOAA and FY1C images. Because GMS images have the fixed observed regions in the given time, we do not map GMS images. 3) To extract shape features in RSI, we should perform the thematic classification. There are so many visual features extraction methods used in remote sensing image research works, such as ISODATA, K-mean value clustering and supervised classification techniques. But the strength of supervised classification based upon the maximum likelihood procedure is that it minimizes classification error for classes that are distributed in multivariate normal fashion. Moreover, it can label data relatively quickly. Its major drawback lies in the need to have delineated unimodal spectral classes beforehand. This, however, is a task that can be handled using clustering, based on a representative subset of image data. Used for this task, unsupervised classification performs the valuable function of identifying the existence of all spectral classes, yet it is not expected to perform the entire classification. Here, we make classifications of GMS, NOAA, and FY1C images using a hybrid supervised/unsupervised method [2]. In this method, we choose cloud, snow/ice, water, and vegetation as our interesting objects (classes), and carry through this classification. 4) After this thematic classification, we will gain the interesting objects, which may be small neighborhoods or even single pixels, we should perform region growing algorithms (Young and Fu 1986) to let small region grow into a bigger region by merging its neighboring regions if the neighboring regions have some properties as the small region. To extract shape information of the interesting objects, we should perform region segmentation techniques to segment each interesting regions such as cloud objects, etc. 5) Those regions which area are bigger than a given threshold or the top three bigger regions will be used in shape similarity measures. We save these images to the binary images.

3 Shape Description

There are many methods used for image shape analysis, such as chain code, polygonal approximation, moment, Fourier descriptor, etc. We may choose many approaches including a variety of moments such as geometrical moment and Zernike moment [3], boundary-based analysis via Fourier descriptors [4], autoregressive model [5], syntactic method [6], and so on. A flexible recognition system must be able to recognize an object shape regardless of its rotation, scale and translation variance. For image shape analysis, three fundamental issues relating to their usefulness must be considered. They include (1) sensitivity to image noise, (2) aspects of information redundancy and (3) capacity for image representation. The invariant properties of moments of 2D and 3D shapes have received considerable attention in recent years. One important advantage of using moment for image shape analysis is that it is not sensitive to image noise. Several invariant moments include geometrical moment, Legendre moment, Zernike moment, pseudo-Zernike moment, rotational moment and complex moment. Geometrical moment is the most common and simplest moment. But in many cases it is not efficient in representing image features.

In order to meet our system requirement of the remote sensing images based on the shape similarity, the shape descriptor should have enough discriminating power and immunity to noise. In addition, the descriptor should be invariant to scale and rotation, not to mention the computation efficiency. The Zernike moment descriptor has such desirable properties such as rotation invariance, scale and translation invariance, robustness to noise, expression efficiency, fast computation and multi-level representation for describing the global shape of a pattern electively.
3.1 The Definition of Zernike Moments

Zernike introduced a set of complex polynomials which form a complete orthogonal over the interior of the unit circle. The form of these polynomials \( V_{nm}(x, y) \) is:

\[
V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta)
\]

where \( n \) is a positive integer or zero, \( m \) positive and negative integers subject to constraints \( n - |m| \) even and \( |m| \leq n \), \( \rho \) is the length of vector from origin to \( f(x, y) \) pixel, \( \theta \) is the angle between vector \( p \) and \( x \) axis in count-clockwise direction, \( R_{nm}(\rho) \) is the radial polynomial defined as

\[
R_{nm}(\rho) = \sum_{s=0}^{\lfloor (n-m)/2 \rfloor} (-1)^s (n-s)! \rho^{n-2s} s!(n+|m|)! s!((n-|m|)/2 + s)!
\]

Note that \( R_{n, -m}(\rho) = R_{n, m}(\rho) \).

These polynomials are orthogonal and satisfy:

\[
\int_{x^2+y^2 \leq 1} [V_{nm}(x, y)]^* V_{pq}(x, y) dx dy = \pi \delta_{np} \delta_{mq}
\]

with \( \delta_{ab} = \begin{cases} 1 & a = b \\ 0 & \text{otherwise} \end{cases} \)

Zernike moments are the projection of the image function onto these orthogonal basis functions. The Zernike moment of order \( n \) with repetition \( m \) for a continuous image function \( f(x, y) \), after vanishing outside the unit circle is

\[
A_{nm} = \frac{n+1}{\pi} \int_{x^2+y^2 \leq 1} f(x, y) V_{nm}^*(\rho, \theta) dx dy
\]

For a digital image, the integrals are replaced by summations as follows:

\[
A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(\rho, \theta), x^2 + y^2 \leq 1
\]

To compute the Zernike moments of a given image, the center of the image is taken as the origin and the pixel coordinates are mapped to the range of the unit circle. Those pixels falling outside the unit circle are not used in the computation.

The process of shape descriptions by Zernike moments for a remote sensing image is demonstrated in Fig. 1. Through the previous preprocessing procedures for RSI, we may gain the binary images for each interesting region of objects. Since the Zernike moments are defined over a unit circle, the radius \( R \) of a circle is determined to enclose the shape completely from the centroid of the binarized shape in the image to the outer most pixel of the shape. To achieve scale and translation uniformity, the regular moments (i.e. \( \beta \)) of each image are utilized. A normalized image function \( g(x, y) \) can be normalized with respect to scale and translation by transforming it into \( g(x, y) \), where

\[
g(x, y) = \frac{f(x/a + \bar{x}, y/a + \bar{y})}{f(0)}
\]

with \( (\bar{x}, \bar{y}) \) being the centroid of \( f(x, y) \) and \( a = \sqrt{\beta/|A_{00}|} \), with \( \beta \) a predetermined value. This normalization step [8] allows the scale invariance for the descriptor. Zernike moments are then extracted from the normalized image, and the magnitudes of Zernike Moments of all orders can be composed of a vector for the shape descriptors. The total number of moments used in the shape descriptor was determined experimentally. Zernike basis functions are computed offline and stored in database. Moreover, Zernike moments are computed using the fast computation method [10]. In our system, because \( |A_{00}| \) as a constant value is equal to \( \beta/\pi \), Zernike moments are normalized by \( |A_{00}| \).

3.2 Shape Description by Zernike Moments

In remote sensing images, clouds, one kind of cover-types, vary continually with time, and they are one of the active uncertainty factors. The shape and area of cloud objects often change, and cloud objects often have complex contours and many disjoint regions. Moreover, Zernike moment have been proven to be superior to moment functions such as geometric moments in terms of their feature representation capabilities and robustness in the presence of image quantization error and noise [7] generally even though it is more complicated than the geometrical moment. Therefore Zernike moments are very suitable for the shape descriptions of cloud objects and other objects in RSI. Although we focus on the discussion of the cloud objects in this paper, the methods for cloud objects are also fit for other objects such as vegetation, snow/ice, etc.

The spatial relational arrangements of objects in remote sensing images constitute an important visual feature for similarity evaluation tasks. The attributed relational graph is widely used as a straightforward representation of structural patterns. The nodes of the graph represent pattern primitives whereas the
A valid pattern primitive is just a subset of \( L_v \) in which each attribute appears only once and each fi may take values \( T_i = \{ t_{ij} | j=1,2,\ldots,J_i' \} \). \( L_a \) is a subset of \( L_v \) in which each attribute appears only once and \( \Pi \) denotes the set of all those valid pattern primitives. Thus, each node will be represented by an element of \( \Pi \).

Similarly, for the edge, we have the attribute set \( F = \{ f_i | i=1 \} \) in which each fi may take values \( T_i = \{ t_{ij} | j=1,2,\ldots,J_i' \} \). \( L_e \) is a subset of \( L_v \) in which each attribute appears only once and \( \Theta \) denotes the set of all those valid pattern primitives. Thus, each node will be represented by an element of \( \Theta \).

4.2 Representing a Cloud Object with a Two-level ARG

Cloud objects are the active uncertainty factors in RSI. So we choose a cloud object as a sample. Fig.2 (a) is the original NOAA-14 AVHRR image. Fig.2 (b) illustrates a cloud object extracted from the original image using section II and III methods. Its graph representation is shown in Fig.2 (c), and Table 1 is the attribute table of nodes and edges, where d(i,j) denotes the distance between node-i and node-j, and the name of each cloud object in Fig.2 (c) denotes the corresponding regions in Fig.2 (b).

For the cloud object of Fig.2 (b), one simple geometrical shape named “area” is used as component primitives. It seems reasonable here that area attribute of each node is an important factor considered in remote sensing images. To compute the area, we can sum up pixel numbers of each node in the corresponding projection map and normalize it by the total pixels in the fixed observation region. Anther shape attribute of each node is Zernike moments discussed in section III.

![Image](Figure 2)

**An example of representing a cloud object with a two-level ARG**

**Table 1**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Node-1</th>
<th>Node-2</th>
<th>Node-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1: Area</td>
<td>0.1437</td>
<td>0.0398</td>
<td>0.0356</td>
</tr>
<tr>
<td>Zernike</td>
<td>A20: 0.5776</td>
<td>A20: 0.7532</td>
<td>A20: 0.5575</td>
</tr>
<tr>
<td>Moments</td>
<td>A20: 0.1530</td>
<td>A20: 0.0647</td>
<td>A20: 0.1782</td>
</tr>
<tr>
<td>vector</td>
<td>A10: 0.0788</td>
<td>A10: 0.0210</td>
<td>A10: 0.1072</td>
</tr>
<tr>
<td>(six orders)</td>
<td>A10: 0.0488</td>
<td>A10: 0.0200</td>
<td>A10: 0.0464</td>
</tr>
<tr>
<td></td>
<td>A62: 0.3783</td>
<td>A62: 0.7977</td>
<td>A62: 0.3951</td>
</tr>
<tr>
<td></td>
<td>A60: 0.3701</td>
<td>A60: 0.2085</td>
<td>A60: 0.3962</td>
</tr>
<tr>
<td></td>
<td>A64: 0.0732</td>
<td>A64: 0.2020</td>
<td>A64: 0.1082</td>
</tr>
<tr>
<td></td>
<td>A66: 0.1693</td>
<td>A66: 0.0733</td>
<td>A66: 0.1867</td>
</tr>
<tr>
<td></td>
<td>A63: 0.1158</td>
<td>A63: 0.0835</td>
<td>A63: 0.0685</td>
</tr>
<tr>
<td></td>
<td>A65: 0.0489</td>
<td>A65: 0.0108</td>
<td>A65: 0.0544</td>
</tr>
<tr>
<td></td>
<td>A66: 0.1623</td>
<td>A66: 0.5700</td>
<td>A66: 0.2741</td>
</tr>
<tr>
<td></td>
<td>A64: 0.4073</td>
<td>A64: 0.3287</td>
<td>A64: 0.4022</td>
</tr>
<tr>
<td></td>
<td>A65: 0.1755</td>
<td>A65: 0.0675</td>
<td>A65: 0.2608</td>
</tr>
<tr>
<td></td>
<td>A66: 0.0482</td>
<td>A66: 0.0099</td>
<td>A66: 0.0746</td>
</tr>
<tr>
<td>Distance</td>
<td>d(1,2): 0.5158</td>
<td>d(1,2): 0.5158</td>
<td>d(1,2): 0.4893</td>
</tr>
<tr>
<td></td>
<td>d(1,3): 0.4893</td>
<td>d(1,3): 0.3632</td>
<td>d(1,3): 0.3632</td>
</tr>
</tbody>
</table>

4.3 Similarities between Two Two-level ARG’s

With a two-level ARG for a cloud object in Fig.2, any node \( n_i \) can be represented by a vector of \( x_i \), the edge between \( n_i \) and \( n_j \) by a vector of \( d_{ij} \), where \( i \) and \( j \) represent the \( i \)th and the \( j \)th nodes, respectively.

For two two-level ARG’s such \( G_i \) and \( G_j \), we can give a similarity distance between the two nodes and two edges as follows:

\[
D_{ni} = W_{x_i} \cdot |x_i| [1] - x_i [1] / |W_{x_i} + W_{d} \cdot \left( \sum_{n=2}^{n} (x_i [s] - x_i [s])^2 \right)^{1/2}
\]

\[
D_{nj} = |d_i^{0,j} - d_j^{0,j}|
\]

1303
Where $D_n$ and $D_e$ is a unified distance ($0 \leq D_n \leq 1, 0 \leq D_e \leq 1$). $W_a$ and $W_z$ is properly selected weights for area and Zernike moments attributes respectively. So a similarity between two two-level ARG’s can be defined directly with

$$Sim_{cloud}(G_r, G_z) = 1-(W_n \cdot D_n + W_e \cdot D_e)$$

Where $W_n$ and $W_e$ are appropriately selected weights for nodes and edges attributes respectively. In particular, when $W_z$ is equal to zero, we think this situation is the first level (coarse) similarity for RSI.

In the same way, when we research on the vegetation, water, snow/ice, and other changing objects, $Sim_{veg}, Sim_{water}$ and so on can be defined as such methods.

All in all, based on the importance of each objects, we can define the total similarity for the two remote sensing images:

$$Sim = 1-(W_c \cdot Sim_{cloud} + W_v \cdot Sim_{veg} + W_w \cdot Sim_{water} + W_s \cdot Sim_{snow})$$

$W_c, W_v, W_w$ and $W_s$ is properly selected weights for cloud, vegetation, water, and snow/ice objects respectively. Users can choose these weights to satisfy their requirement.

4.4 Algorithm for Graph Matching

A similarity measure means to compute the similarity between two attributed relational graphs. The similarity measure is a matching function, and gives the degree of matching or similarity for a given pair of two objects represented by two-level ARG’s. In order to evaluate the similarity between graphs, we must use an efficient graph matching algorithm, error-correcting subgraph isomorphism algorithm [11], algorithms reported by Ullman [13] will suffer from combinatorial explosion when the size of the graphs becomes large. So we choose the fast graph matching algorithm [12] to compute the similarity between graphs. This efficient algorithm significantly reduces computation time compared with the classical algorithms, without any significant loss in matching results. The fast algorithm [12] is designed to find a graph isomorphism when both graphs have the same number of nodes and a subgraph isomorphism when one has fewer nodes than the other. The complexity of this algorithm depends on the number of phases $K$, where the value of $K$ ranges from 1 to the minimum of the numbers of nodes in the two graphs to be matched. In matching procedures of nodes, for each valid mapping, we can compute the similarity distances of nodes and edges using (8), (9), (10).

5 Experiments and Conclusion

The sample graph in Fig.2 (b) was regarded as a sample image to be matched in our database, which has many NOAA-14 images received from Dec. 2000, we perform our proposed methods for each image, and save these results and parameters to our database. In this experiment, we chose 6 orders of Zernike moments for each region shape, and set $W_n = W_a = W_z = W_e = 0.5$. Through the previous processing, finally we compute the similarity distances between the sample graph and other NOAA images. The main six types of images retrieved from the database are shown in Fig.3.
shape and area of each node and structural pattern between
nodes are similar to those of the sample graph. In Fig. 3 (b), if
this graph is rotated with 90 degrees clockwise, we can find out
that it is relatively similar to the sample. From Fig. 3 (e) to Fig. 3
(f), the similarities are very low, and we may easily make out
the differences of spatial relation between nodes, shape features
and the area of each corresponding node with the sample graph.
With the similarity value continually lower, the differences
become comparatively larger in visual features and vice versa.
So this method is relatively satisfactory for the retrieval of
remote sensing images based on shape similarities measures.

Our experiment proved sufficient coherence of retrieved results
when Zernike moments are used to describe shape features and
the two-level ARG is used to represent relations among those
objects. We can say that this technique is promising for the
retrieval of content-based remote sensing images. Now this
approach has been used for our network-oriented multiple
satellites images management and online distribution system.
On the other hand, much work is still required to make this
method faster and more precise.

References

[1] T. Y. Young and K. S. Fu, Handbook of pattern recognition and
Aided Analysis of Landsat-1 MSS Data: A Comparison of Three
Approaches including A Modified Clustering Approach.
Information Note 072475,” Laboratory for Applications of
information processing and analyzing system,” Engineering
Application Artificial Intelligence, vol.7, no. 6, pp. 593-606,
1994.
boundary analysis: representation and reconstruction,” IEEE
comparative study and new results,” Pattern Recognition, vol.
[8] Alireza Knotanzad and YawHua Hong, ” Invariant Image
Recognition by Zernike Moments, ” IEEE Trans. PAMI, vol.12,
moments,” IEEE Trans. Pattern Analysis, Machine Intelligence,
algorithm for fast computation of Zernike moments,” Pattern
SMC, vol. 9, no. 12, December 1979.
Graph Matching,” in Proc. 16th International Conf. Pattern