Design of Content-Based Retrieval System in Remote Sensing Image Database

LI Feng  ZENG Zhiming  HU Yanfeng  FU Kun

ABSTRACT  To retrieve the object region efficaciously from massive remote sensing image database, a model for content-based retrieval of remote sensing image is given according to the characters of remote sensing image application firstly, and then the algorithm adopted for feature extraction and multidimensional indexing, and relevance feedback by this model are analyzed in detail. Finally, the contents intending to be researched about this model are proposed.

KEYWORDS  content-based retrieval; remote sensing image; image database; feature extraction; object region

Introduction

Content-based image retrieval (CBIR) technology was proposed in 1990s and it is a image retrieval technology using image vision contents such as color, texture, shape, spatial relationship, not using image notation to search images. It resolves some traditional image retrieval problems, for example, manual notations for images bring users a large amount of workload and inaccurate subjective description. After more than one decade, it has been developed as content-based vision information retrieval technology including image information and video information. Great progress has been made in theory and applications. A large number of papers and literatures about theory research and technology application have been published. A lot of CBIR systems have been developed for business and science research. Great deals of international conferences have been held by IEEE and SPIE organization.

At present, CBIR technology obtains successful applications in face reorganization fields, fingerprint reorganization fields, medical image database fields, trademark registration fields, etc., such as QBIC system of IBM Corporation, Photobook system of MIT Media Laboratory and Virage system of Virage Corporation. It is difficult to apply these systems in massive remote sensing image searching because remote sensing image has many features including various data types, a mass of data, different resolution scales and different data sources, which restrict the application of CBIR technology in remote sensing image field. In order to change the current situation, we must resolve some problems as follows.

1) Storing massive remote sensing image data of TB level.
2) Designing reasonable physical and logical pattern of remote sensing image database.
3) Adopting adaptive image feature extraction algorithms.
4) Adopting efficient high-dimensional indexing structure for efficient search.
5) Designing efficient and reasonable content-based searching engine of massive remote sensing image database.

Received on April 18, 2006.
LI Feng, Ph. D, Institute of Electronics of CAS, 19 West Beishuan Road, Beijing 100080, China.
E-mail: lifeng@opt.ac.cn
1 Design architecture of content-based remote sensing image

For practical applications, users are often interested in the partial region or targets, such as military targets, public targets and ground resource targets in remote sensing image instead of the entire image. For example, the small scale important targets and regions of remote sensing image arrest more attention than the entire remote sensing image in application. We bring forward a remote sensing image searching model in which some important target and regions are cut out by image processing tools firstly. These important targets and regions image slice are stored as BLOB data, their archive information are stored as attribute data secondly. We build the relationship between the image slices of these important targets and regions and original remote sensing image finally.

These image slice features of important targets and regions extracted by color, texture, shape, spatial relationship, etc. are stored in feature database. In order to promote the efficiency of feature extraction algorithm, we often adopt parallel computation technology to use computer resource in network fully. The operating efficiency of total system is also promoted.

Efficient indexing technology is a key factor for applying the content-based image retrieval in massive image database successfully. Indexing technology developed from traditional database and has been applied in content-based image retrieval field subsequently. For remote sensing image application, we need to use parallel computation to reduce dimension number of feature space and match feature among images in order to construct efficient massive retrieval engine for content-based remote sensing image. Fig. 1 shows an architecture frame of content-based remote sensing image.

This model is composed of following four parts: remote sensing image pre-processing, remote sensing image database, remote sensing image searching engine, human-computer interaction (HCI).

Remote sensing image pre-processing part selects and cuts out important targets and regions of interest (ROI) of remote sensing image as new image. It combines these slice images with original image and extracts these slice images vision feature and construct multidimensional indexing structures. Remote sensing image database part consists of feature database and image database. The vision feature data extracted by many kinds of algorithms such as texture co-occurrence matrix representation, wavelet outline descriptor, etc. is stored in feature database. The ROI image data and metadata of remote
sensing image is stored in image database. Searching engine is the core of this architecture. It provides efficient and accurate retrieval result. HCI part provides users with the searching interface by which users commit searching request using example searching and sketch searching method. Reference feedback technology is introduced in this part to promote the recall and precision of retrieval.

2 Feature extraction

Feature extraction algorithm is the basis of content-based image retrieval. Important targets and regions in remote sensing image represent diverse contents because of diverse resolution, diverse phase and diverse imaging sensors. Their color, texture, shape feature can be extracted as these targets and regions image feature. Color feature is insensitive to image size and direction and ineffective to image rotation and translation. But it is not good to express local feature of image. Texture feature has the rotation invariance and the anti-noise capacity. But the texture feature extracted by different image with diverse resolution has a very large difference. Shape feature has high complexities for design and computation, so there is a need to extract a large number of features to express shape by modeling. The dimension of the feature vector of shape feature is too high to index. According to the character of remote sensing application, users pay close attention to the texture feature and shape feature. Texture feature has the rotation invariance and the anti-noise capacity. But the texture feature extracted by different image with diverse resolution has a very large difference. Shape feature has high complexities for design and computation, so there is a need to extract a large number of features to express shape by modeling.

2.1 Texture feature extraction

Texture feature is an important feature of image. The main texture feature algorithms used for image retrieval are Tamura method, MRSAR (multi-resolution simultaneous auto-regressive method), Gabor transform method, PWT (pyramid wavelet transform) method, TWT (tree wavelet transform) method, co-occurrence matrix representation method, edge histogram descriptor method, and so forth. Edge histogram descriptor is one of texture descriptor recommended by MPEG-7 international standard. Image texture descriptor is computed in either space domain or transform domain. We can also compute it using random field model.

MPEG-7 standard classify edge histogram descriptor into image texture descriptor. In fact it expresses five kinds of edge sub-image which are 0° edge, 45°edge, 90°edge, 135°edge and non-directional edge. They are expressed as follows.

\[
E = \begin{bmatrix} E_{00} & E_{45} & E_{90} & E_{135} \end{bmatrix}
\]

Edge histogram descriptor is computed according to following steps.

1) Image is divided into 16 (4 × 4) sub-images. These sub-images are divided into a series of image blocks and each block contains 4 sub-blocks.

2) Each sub-image should be checked whether 5 kinds of edges exists respectively using 5 kinds of edge detection operators as above.

3) The edge histogram of sub-image is computed respectively and each edge histogram contains 5 histogram bins.

4) The histogram bins of 16 sub-images is integrated to become edge histogram of entire image which contains 80 histogram bins.

5) Edge histogram acquired as above acts as image texture feature to retrieve image.

2.2 Shape feature extraction

Shape is also an important feature of object. It is very difficult to acquire shape feature of object automatically so that the image retrieval based on shape is applied usually to retrieve the image whose objects are recognized easily. Shape can be described by some global and local feature...
such as area, fractal dimension, eccentricity, shape moment, curvature and so on. The main shape description and analysis methods\[1\] are region invariance moment, Fourier descriptor, wavelet boundary descriptor\[4\], autoregression model\[2\], wavelet modulus maxima\[4\], edge histogram and so on.

It is the semantic feature that the shape of target in image. The description of object shape is more complicated than color description and texture description. The object’s shapes in image which acquired with diverse angle of vision have a very large difference. In order to match target shape accurately, there is need to solve some complicated problems such as scale invariance and rotation invariance, etc. These factors are very important to image retrieval.

The shape feature extraction based-on region invariance moment is defined as follows: for image \( f(x,y) \), if it is continuous in subsections and it has finite nonzero point in \( XY \) plane, then we can prove that its moments of every order exist.

Define the \( p+q \) order moment as

\[
m_{pq} = \sum_x \sum_y x^p y^q f(x,y)
\]

It is proved that \( f(x,y) \) and \( m_{pq} \) are one and only for each other. The central moment of the \( p+q \) order is

\[
m_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y)
\]

where \( (\bar{x},\bar{y}) \) is centroid coordinate, \( \bar{x} = m_{10}/m_{00} \), \( \bar{y} = m_{01}/m_{00} \). The normalized central moment is

\[
\eta_{pq} = \frac{m_{pq}}{m_{00}}
\]

\[
r = \frac{p+q}{2} + 1, p + q = 2,3,\cdots
\]

where

\[
\begin{align*}
\phi_1 &= \eta_{00} + \eta_{11} \\
\phi_2 &= (\eta_{10} - \eta_{01})^2 + 4\eta_{11} \\
\phi_3 &= (\eta_{00} - 3\eta_{11})^2 + (3\eta_{10} - 3\eta_{01})^2 \\
\phi_4 &= (\eta_{11} + \eta_{00} + \eta_{10} + \eta_{01})^2 \\
\phi_5 &= (\eta_{00} - 3\eta_{11} + \eta_{01} + \eta_{10})[\eta_{00} + \eta_{11}^2 - 3(\eta_{10} + \eta_{01})^2 + (3\eta_{10} - 3\eta_{01})(\eta_{01} + \eta_{10})] \cdot [3(\eta_{00} + \eta_{11})^2 - (\eta_{10} + \eta_{01})^2] \\
\phi_6 &= (\eta_{00} - \eta_{11})[\eta_{00} + \eta_{11}^2 - (\eta_{01} + \eta_{10})^2] + 4\eta_{11}(\eta_{01} + \eta_{10})(\eta_{01} + \eta_{10})
\end{align*}
\]

Above seven moments invariant to translation, rotation, and size can be calculated by combining normalized 2-order and 3-order central moments. Then they will be used for searching feature vectors describing object shapes.

3 High dimensional indexing

The feature vector of remote sensing image acquired by feature extraction algorithms as above has a high dimension with \( 10^2 \) level generally. As a result, matching feature vector need complex computation and reduce retrieval efficiency subsequently. The feature vector dimension must be reduced firstly and effective multi-dimension indexing algorithm must be proposed secondly.

The mapping from high dimension to low dimension named as dimension reduction. There are two widely used approaches: KLT transform\[4\] and column-wise clustering method\[7\]. KLT and its variation have been used in many areas such as face recognition, eigen-images, principal component analysis and information analysis, etc. Faloutsos and Lin propose KLT fast approximate algorithm\[1\] to reduce the feature dimensions. The research shows that the dimension of most actual dataset (vision feature vector) can be reduced and can not be degenerated on retrieval effect. Chandrasekar and his fellows propose SVD method\[3\] which is effective and steady on numeric data during KLT transform. Image retrieval system is a dynamic system and the new images are added in image collective continuously. This method has a capability to process index updating dynamically. Besides KLT, clustering is the other one powerful tool to reduce dimensions and it is applied into pattern recognition, voice analysis and information retrieval widely. Similar objects (template, signal and document) are collected to recognize or classify. This kind of clustering is named as row-wise. At the same time, the column-wise
can also reduce dimensions.

The high dimensional indexing technology appears in the middle of 1970s. The main methods are k-d tree and quadtree but they can not meet performance demand. To meet the demands of space indexing from GIS and CAD, Guttman proposes R tree indexing structure and some improved algorithms of R tree, such as R+ tree, R * tree, are developed subsequently. Traditional tree indexing technology is proposed for searching on relation database instead of image database retrieval. At present, the focus is that how to identify and improve indexing technology suitable for expressing high-dimension feature vector. Clustering and neural network are two kinds of promising methods.

4 Relevance feedback

Feedback is a kind of usual adjustment technology which adapts to the user's demands and promotes retrieval precision in content-based vision information retrieval. The most usual feedback method is relevance feedback. It is a instructional studying technology and its entire process is as follows.

1) According to the similar matrices and the sample images, the features of images in database, a sort list is generated on basis of similarity degree. More similar they are in feature space, more forward their sequences are.

2) Users select a group of positive-feedback images coincident to retrieval content and a group of negative-feedback images coincident to retrieval content, and submit them to retrieval system.

3) Retrieval system re-optimize similarity matrices according to submitting information from users, then the first step, submit the new result to users finally.

The typical iterative formula to realize that is:

\[ Q_{i+1} = aQ_i + \beta \left( \frac{1}{N_p} \sum_{D_p} D_i \right) - \gamma \left( \frac{1}{N_n} \sum_{D_n} D_i \right) \]

where \( a, \beta, \gamma \) are weighted constants; \( Q_i \) is the searching position of \( i \) iteration; \( D_i \) is the feature vector; \( D_p \) is the positive-feedback image collective; \( D_n \) is the negative-feedback image collective; \( N_p \) is the number of the similar sample in positive-feedback images; \( N_n \) is the number of the similar sample in negative-feedback images.

The retrieval system utilizes the interaction information from users effectively and promotes the precision and the efficiency by optimizing retrieval vectors and similarity matrices. So using relevance feedback technology in image retrieval will provide more information for retrieval system and improve the retrieval precision.

REFERENCES

[1] Yong Rui (1999) Image retrieval: current techniques, promising direction and open issues [J]. *Journal of Visual Communication and Image Representation*, 10(3):39-62

[2] Gudivada V N, Raghava V V (1995) Content based image retrieval systems[J]. *IEEE Computer*, 28(9):18-22

[3] Ma Weiyng, Zhang Hongjiang (1998) Benchmarking of image feature for content-based retrieval[C]. The Thirty-Second Asilomar Conference on Signals, System and Computer, California, USA

[4] Mallat S, Hwang W L (1992) Singularity detection and processing with wavelets[J]. *IEEE IT*, 38(2):617-643

[5] Zhang Minjin (2003) Content-based visual information retrieval[M]. Beijing: Science Press (in Chinese)

[6] Ng R T, Sedighian A (1996) Evaluating multidimensional indexing structures for images transformed by principal component analysis[C]. SPIE Conference, San Jose, California

[7] Salton G, McGill M J (1983) Introduction to modern information retrieval[M]. New York: McGraw-Hill

[8] Faloutsos C, Lin K (1995) Fastmap: a fast algorithms for indexing, data mining and visualization traditional and multimedia database[C]. ACM SIGMOD, San Jose, California

[9] Chandrasekaran S, Manjunath B S, Wang Yuanfang, et al. (1997) An eigenspace update algorithm for image analysis[J]. *Graphical Models and Image Processing*, 59(5):321-332

[10] Guttman A (1984) R-tree: a dynamic index structure for spatial searching[C]. ACM SIGMOD, Boston