First Steps Towards an Interactive Image Retrieval System for Geoscientific Images

F. Boschetti and S. B. Cho

CSIRO, Exploration and Mining, E-mail: fabio.boschetti@csiro.au
Department of Computer Science, Yonsei University, E-mail: sbcho@csai.yonsei.ac.kr

Abstract

We present the first steps towards the development of an interactive system for the retrieval of geoscientific images. The images do not need to be assigned specific keywords in order to be retrieved. The search for a suitable image is performed by interacting with the user who ranks various candidate images according to the similarity with a target. A Genetic Algorithm employs the user's ranking to converge towards the desired image. The system allows for the retrieval of images even when the target is not defined exactly, as well as to modify the target during the search.

1. Introduction

As digital image libraries grow rapidly in size, content-based image retrieval has attracted the attention of researchers in several fields. Content-based image retrieval offers two main advantages compared to the traditional keyword-based approach. First, images in the database do not need to be assigned keywords, a process which is not only is time consuming, but also problem dependent and to some extent subjective. Even within a relatively narrow field of application (geology for example), keywords effective for a specific task (such as geomorphology) may not be effective for a different one (structural geology). Secondly, keywords may not allow accurate discrimination of the images.

Several working content-based systems have already been developed, including QBIC system of IBM (Niblack et al., 1993), Chabot of UC Berkeley (Ogel and Stonebraker, 1995), QVE (Hirata, and Kato, 1992), Photobook of MIT, and Image Surfer of Interpix Software. A particularly interesting approach has been developed in order to allow for human intuition and emotion in the retrieval process (Takagi, 2001). The system provides the user with the capability to express subjective judgment via the use of interactive evolutionary computation (Lee and Cho, 1998).

There are various reasons why a system for interactive retrieval of scientific images could be useful. First, one may want to search for a specific feature inside an image (a specific geological or geophysical signature, or a specific part of a fossil or a plant, for example). This clearly cannot be achieved by keywords, since it is not possible to precisely label all features inside an image. Moreover, some features do not have regular shapes, and accordingly traditional pattern matching algorithms may not succeed. In this case we can envisage subdividing images into smaller windows and applying the search to these sub-images. Alternatively, we can envisage a scientist having to classify an image arising from applications outside his/her particular expertise. Imagine collecting the petal of a flower and trying to classify the flower by using an archive of flower images. The only way to find the proper flower would be to scan all images until the closest petal is found. Our system is aimed at mimicking the action of querying an expert by providing information like 'my petal is similar to figure 4 and 8, and very different from figure 12 and 2'. Such information alone is used by our search engine to narrow the options and hopefully find the desired image.

Finally, we can envisage the situation in which the target of the search is not exactly defined, or may vary during the search due to hints provided by the retrieved images themselves. This happens frequently in geosciences: often geologists are after visual clues to spark possible novel interpretations. For all these purposes an interactive approach would be beneficial.

2. Background

2.1 Interactive Genetic Algorithm

A Genetic Algorithm (GA) is an optimization technique inspired by biological evolution. It works by creating a population of candidate solutions and iteratively improving them by applying operators equivalent to biological selection, crossover and mutation. Because of their widespread use we refer the reader to some standard GA literature (Goldberg, 1989). Details about the specific GA implementation used in this paper can be found in Lee and Cho (2001).

Standard GAs, like most other optimization algorithms, use a numerical evaluation of the quality of a candidate solution to drive the optimization process. In order to extend the use of GAs to problems for which a solution cannot be expressed in numerical terms, Interactive Genetic Algorithms (IGA) have been developed. This GA variant employs the subjective choice of the user to measure the quality of a solution. This allows dealing with applications in which human intuition or emotions are of crucial importance. Traditionally IGAs have been applied to fields such as computer graphics and art, while more recently they have proven successful in engineering and geoscientific problems, such as geomechanical modeling (Boschetti and Moresi, 2001). An exhaustive review of IGA implementation and application can be found in Takagi (2001).

2.2 Interactive Image Retrieval System

In this work we employ an interactive image retrieval system developed by Cho (2000).

Basically, the system attempts to model the user expectations by applying genetic operators to the user choices. The system is based on three main components:

1) a coding algorithm to extract the main features in an image thereby reducing the dimensionality of the search problem.
2) a user interface, allowing the user to view some images from the data base and rank them

according to his/her judgment;

3) a GA to generate a new set of candidate images according to the user ranking;

The three components are described in the following sections, with particular emphasis on the first one, which has been the main area of investigation in this work.

2.3 Image Coding

Two image coding algorithms have been implemented and tested in this work: Discrete Wavelet Transform (DWT) and local multi-point statistics (MPS).

Discrete Wavelet Transform. A wavelet transform (WT) attempts to extract information about both the size and the location of the features in a signal (Graps, 1995). In recent years WT has been applied to a wide range of applications, from signal processing to the solution of differential equations, to algebra. In the signal-processing world, WT is today seen as an alternative, and at times as a complement, to using Fourier Transform. Whereas the basis functions of Fourier transform are sinusoids with infinite support, the wavelet basis are oscillatory functions whose behavior is reasonably localized both in space and scale. The result of a WT thus gives information about both the frequency content and the spatial location of the main features in a signal. Basically, any oscillatory function with zero mean and appropriate scaling behavior can be employed as a wavelet basis.

For applications in which a compact representation is needed, it can be advantageous to use an orthogonal wavelet basis. Families of orthogonal wavelet basis are described in the literature (see for example Daubechies, 1988). In this work we employed the Haar basis, which is particularly popular because of its ease of implementation.

A standard two-dimensional Haar wavelet decomposition of an image is implemented by pyramid algorithm (Edwards, 1991). The data are passed through two convolution functions, each of which creates an output stream that is half the length of the original input. The low pass filter accounts for the average of the signal within the scaling window, while the high pass filter accounts for the difference between the original signal and the average. This approach (as well as similar ones based on different wavelet basis) is particularly useful for image coding since most of the information of an image gets stored into a limited number of coefficients. The remaining majority of coefficients with small magnitude can be set to zero, with only minor loss of information in the reconstructed image.

Allowing for a greater loss of information, a very small number of coefficients have been shown to be sufficient for a broad classification of images into different classes (Jacobs, 1995). This is the rationale behind the use of WT for coding in our interactive image retrieval system.

As shown in Jacobs' work, storing 40~50 largest-magnitude coefficients in each color band works best as discriminatory criteria for image classification. Moreover, truncating the coefficients appears to even improve the discrimination power of the metric. Therefore, we only store the sign information of coefficient values into our representation.

Local multi-point statistics. The local multi-point statistic comes from scanning an image with a small pixel window. At each window position, the template lies over (say) $m \times n$ pixels. We interpret the values in the pixels under the template as the coordinates of a point in "mn dimensional state space". Thus, each template position defines a point in the state space. Note that this is not a tiling process as the template positions overlap. When the entire image is scanned, a 'cloud' of state space points is obtained, the configuration and density of which contains information about local textures. Note that this method is a higher dimensional probabilistic extension of the idea underlying nonlinear analyses of time series (Packard, 1980).

In our work, after scanning each image and generating the cloud of points in the state space, we randomly select 50 points as representative of the image structure. This represents the coding with the retrieval system.

2.4 User Interface.

The user interface developed for this work is shown in Figure 2. It allows the user to view a predetermine number of images (equivalent to the population size in the GA) and to rank them according the user expectations, by use of scroll-down menus. The user ranking is then fed to the GA for the generation of the next set of images.

2.5 IGA implementation

The GA has been implemented with a population size of 12 individuals. After the DWT or the MPS have been applied to an image, and a fixed number of coefficients have been selected, these are stored in the GA chromosome. We employ one point crossover by selecting a point and swapping a part of chromosome depending on its position. More details about the GA implementation can be found in Cho (2000).

3 Image Retrieval System for geoscientific images

The entire system is constructed as shown in Fig. 3.

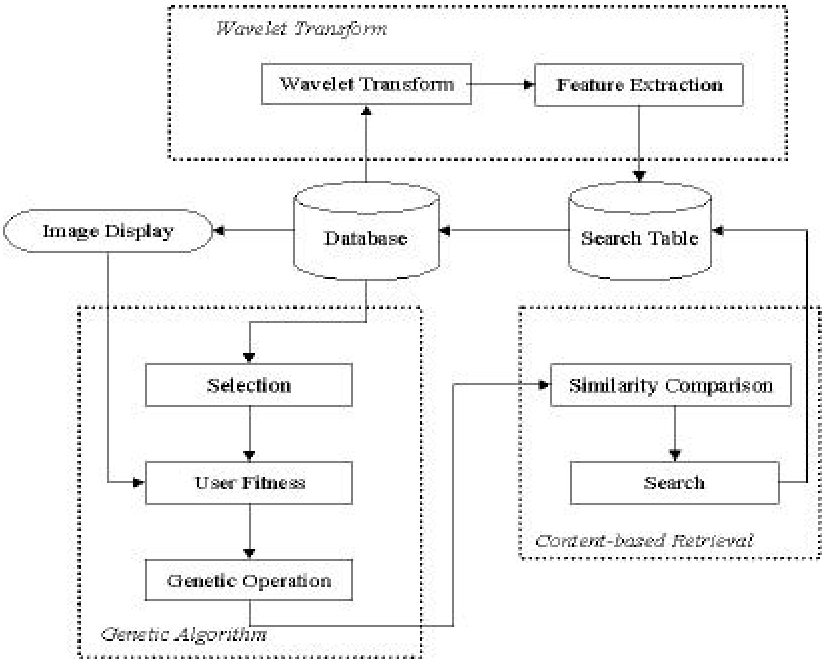

Figure 1. System structure.

The interactive image retrieval system works in the following way.

1) In the preprocessing step, image coding (either DWT or MPS) is performed for every image in the database, and the compressed representations are stored in a search table.
2) The GA randomly generates a set of images.
3) The user ranks the images according to their resemblance to the image he/she is attempting to retrieve;
4) The genetic operators are applied to the selected images parameters (codings) and a set of new image codings is generated.
5) Because the set of images in the database is limited, the image codings generated by the genetic operators most likely do not correspond to any specific image in the database. The images to be displayed in the next generation are chosen by looking for the images in the data base whose codings are closer to the ones generated by the GA.
6) Steps 3-5 are iterated until the user is satisfied with the retrieved image.

4 Geological database.

The efficiency of the DWT and MPS coding in the image retrieval system has been tested on a database of geological images. The images have been scanned from “The Microstructures of Deformed Rocks, by Lionel Weiss, 1963. They represent photographs of rock structures characterizing the signatures of different geological processes. They include various patterns of folds and fractures at various scales. A sample of the images can be seen in Figure 2. It should be immediately clear that most information in the images is contained at the fine scales and that the general appearance of the images is quite similar. This makes this particular database an extremely hard test for the discriminatory power of our system.

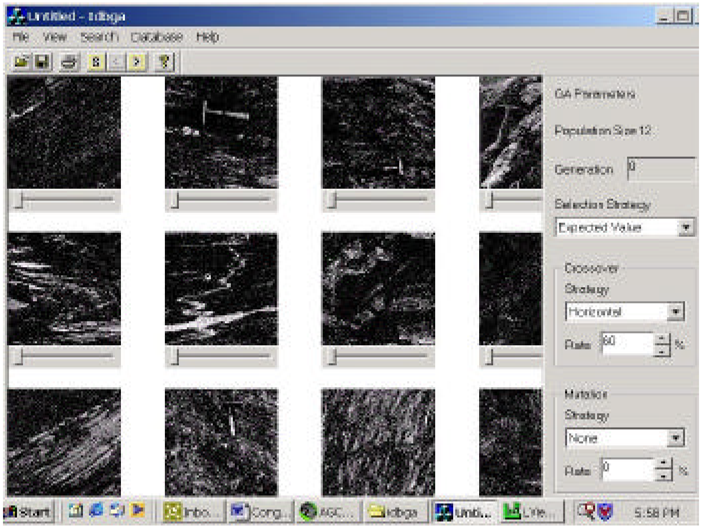

Figure 2. User interface and sample of the rock images.

5 Comparison of coding efficiency.

The crucial factor in determining the performance of the retrieval system is the efficiency of the coding algorithm to discriminate and classify the images. In order to evaluate the potential of both DWT and MPS in this geological application we asked 3 experts to classify the rock images according the geological criteria. The images were classified into 8 categories. After this classification was obtained we tested both DWT and MPS results to see whether their coding respected the experts’ classification. This is achieved by checking the intra-class variance against the interclass variance. The results are presented in the following tables.

Table 1. Statistical results for the WT coding.

Category name (df)	Mean (in class)	Variance (in class)	Mean (out of class)	Variance (out of class)
Category1(39/273)	20.235	0.989	21.093	0.012
Category2(62/250)	19.944	0.237	19.388	0.098
Category3(20/292)	22.985	1.221	24.932	0.290
Category4(13/299)	23.273	1.983	23.192	0.372

Category5(119/193)	22.746	0.389	22.893	0.187
Category6(14/298	21.084	0.983	25.939	0.098
Category7(25/287)	21.073	1.297	21.221	0.072
Category8(20/292)	18.876	0.692	17.219	0.153

Table 2. Statisticalresults for the MPS coding.

Category name (df)	Mean (in class)	Variance (in class)	Mean (out of class)	Variance (out of class)
Category1(39/273)	39.13	1.234	32.35	0.876
Category2(62/250)	42.86	0.992	38.92	1.156
Category3(20/292)	45.91	0.171	26.89	0.821
Category4(13/299)	37.12	0.783	41.34	0.472
Category5(119/193)	32.15	0.649	29.19	0.927
Category6(14/298	44.34	1.921	41.97	0.862
Category7(25/287)	35.98	0.831	29.97	0.163
Category8(20/292)	18.876	0.692	17.219	0.153

The results show that WT works better at discriminating the rock images categories than the MPS. However, none of the two coding strategies is particularly effective (WT discriminates 3 categories among 8, while MPS only 1). Clearly this is the area where further research needs to be focused. In particular the following issues need to be addressed:

1) How many wavelet coefficients needs to be stored to best discriminate the rock classes?
2) Should the coefficients be chosen at specific scales?
3) How many points in the MPS strategies should be chosen?
4) Should they be chosen randomly or at specific spatial location (center of the sample could?)
5) Is there any other feature extraction method that could outperform the above methods?

6 Concluding Remarks

Two concepts have been explored in this work. First is the idea of using the user's subjective judgment to search scientific database of images. The rationale is that, despite the search for scientific images being usually performed by deterministic factors, at times a database may be so large than an exhaustive search or an exhaustive labeling of the images may be impossible. Another rationale, probably the most important, is that at times the user may be uncertain of what criteria to use to discriminate an image, or he/she may be searching a database of images related to scientific subject outside his/her particular expertise.

The second concept investigated is whether a pseudo deterministic approach (wavelet transform followed by the selection of the largest magnitude coefficients) or a fully statistical one (random selection of points from a cloud describing the internal structure of an image) gives the best coding strategy for the discrimination and classification of images.

Our initial results on a very challenging database of images suggest that the development of interactive image retrieval in scientific images is worth pursuing. Further investigation is however necessary to determine the best coding algorithm for image classification. Establishing the necessary minimum number of wavelet coefficients necessary and the optimum number of points in the statistical cloud of image template samples for the image representation will be our next research step. Other image processing techniques, including specific feature extraction will also be tested.

References

F. Boschetti and L. Moresi, "I*nteractive Inversion in geosciences*", 2001, *Geophysics,* 64, 1226-1235.

Daubechies, I., 1988. "Orthonormal bases of compactly supported wavelets", Commun. Pre Appl.Math., 41, 909-996.

T. Edwards, "Discrete wavelet transform: Theory and implementation," *Technical Report*, Stanford University, September 1991.

A.Graps, "An introduction to wavelets," *IEEE Computational Science and Engineering*, vol. 2, no. 2, pp. 50-61, 1995.

D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company Inc. 1989.

K. Hirata, and T. Kato, "Query by visual example: Content based image retrieval," *Advances in Database Technology*, pp.56-61, 1992.

C.E. Jacobs, A. Findkelstein, and D.H. Salesin, "Fast multiresolution image querying," *Proc. of SIGGRAPH 95*, 1995.

J.-Y. Lee, and S.-B. Cho, "Interactive genetic algorithm for content-based image retrieval," *Proc. of Asian Fuzzy Systems Symposium (AFSS'98),* pp.479-484, Masan, Korea, June 1998.

J.-Y. Lee and S.-B. Cho, "Wavelet encoding for interactive genetic algorithm in emotional image retrieval," C. Hsu (Ed), Advanced Signal Processing Technology by Soft Computing, pp. 225~239, World Scientific Pub., 2001.

W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D.Pekovic, P. Yanker, C. Faloutsos, and G. Taubin, “The QBIC project : Querying images by content using color, texture, and shape,” *Storage and Retrieval for Image and Video Databases*, pp.173-187, SPIE, 1993.

V.E. Ogel, and M. Stonebraker, “Chatbot: Retrieval from a relational database of images,” *IEEE Computer*, pp.40-48, September 1995.

Packard, N.H., Crutchfield, J.P., Farmer, J.D., and Shaw, R.S., 1980, Geometry from a time series: Phys. Rev. Lett., **45**, 712-6.

H. Takagi, 2001, “Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation” *Proceedings of the IEEE, 89, 1275-1296*