The Third AFSS(1998), 479-484

Interactive Genetic Algorithm for Content-based Image Retrieval*

Joo-Young LEE and Sung-Bae CHO

Department of Computer Science, Yonsei University

134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, Korea

Tel:+82-2-361-2720 Fax:+82-2-365-2579 Email : [jylee,sbcho]@candy.yonsei.ac.kr

Abstracts

As technology in computer hardware and software advances, efficient information retrieval from multimedia database gets highly demanded. Recently, it has been actively exploited to retrieve information based on the stored contents. However, most of the methods emphasize on the points which are far from human intuition or emotion. In order to overcome this shortcoming, this paper attempts to apply interactive genetic algorithm to content-based image retrieval. A preliminary result with subjective test shows the usefulness of this approach.

Keywords : Interactive genetic algorithm, content-based image retrieval, wavelet transform

1. Introduction

As the power of computer improves, computer users request the facility for information retrieval from multimedia database. In particular, they have interests in the methods that allow to retrieve information based on the stored contents. Several working systems have already been developed: QBIC system of IBM, QVE (Query by Visual Example) of Hirata and Kato and so on. Since these systems extract the features, such as color histogram, texture and some shape features, from user by query, they perform better than the conventional methods based on keyword matching. However, most of the previous systems are based on the engineering approach that has little relevancy to human intuition and emotion.

Due to the lack of expression capability, the user gets some difficulty to retrieve the images that he/she wants. Image retrieval based on human intuition could overcome the difficulty, and play an important role of reducing psychological burden to retrieve the images for which the user cannot think of the concrete query.

In this paper, we present interactive genetic algorithm that can perform content-based image retrieval using human intuition. In interactive genetic algorithm, each individual in population is evaluated by the user and the next generation is produced based on the evaluated value, by which the system might incorporate human preference into the process of image retrieval. Also, the experiments with a small database indicate that the interactive genetic algorithm can be useful in content based image retrieval.

2. Related Works

In this section, we will briefly introduce related works on the content-based image retrieval and interactive genetic algorithm.

2.1. Content-based Image Retrieval

The previous research can be divided into two classes, depending on whether the object retrieved can be clearly expressed or not. The case that can be clearly expressed is further divided into three categories. The first one assigns the contents based on the objects of the image or on their relationship, the second is based on the features of the image, and the third is by image shape approximation.

When target image cannot be clearly expressed, not only management and combination of keywords, but also well designed user interface to be able to interpret user's intention are important. Methods which use knowledge database and queries flexibly to extract user's intention have been proposed. These methods allow flexible image retrieval, but still cannot reflect

* This work has been supported in part by a grant no. 981-0919-099-2 from the Korea Science and Engineering Foundation (KOSEF).

user's emotion or feeling.

2.2. Interactive Genetic Algorithm

Interactive genetic algorithm uses user's judgment as a fitness. It has not been studied broadly so far. But a property which user determines fitness allows to adapt interactive genetic algorithm in the field of the graphics or the art.

For example, Caldwell and Johnston applied it to tracking criminal suspect[5], this system produces montages by the fitness based on the face of criminal suspect by witness's account. Baker implemented line drawing system based on the user's aesthetic criteria using interactive genetic algorithm. The criteria of this system are attraction, amusement, control and so on. However, it was difficult to make specific shape to want exactly, because the size of population was small. To solve the problem, he tried to expand search space by increasing mutation rate.

3. Discrete Wavelet Transform

We use two dimensional wavelet transform to extract features from images and to use the result of the transform for a search. Wavelet transform allows for very good image approximation with just a few coefficients. This property has been exploited for lossy image compression.

```
// one-dimensional decomposition
Proc DecomposeArray(A:array[0..h-1] of color)
  A = A / √h
  while h > 1 do
  h = h / 2
  for i=0 to h-1 do
    A'[i] = ( A[2i] + A[2i+1] ) / √2
    A'[h+i] = ( A[2i] - A[2i+1] ) / √2
  end for
  A = A'
end while
end proc

// entire r×r image decomposition
Proc DecomposeImage(T:array[0..r-1,0..r-1] of color)
for row = 1 to r do
   DecomposeArray(T[row,0..r])
   end for
for col = 1 to r do
   DecomposeArray(T[0..r,col])
end for
end proc
```

Fig. 1 Pyramid algorithm.

The coefficients of a wavelet transform provide information that is independent of original image resolution. Thus, a wavelet based scheme allows the resolution of the query and the target to be effectively decoupled [4]. Especially, we have no need to maintain detailed informations to reconstruct completely.

In particular, Haar wavelet transform is selected, because it can be easily implemented and fast. Fig. 1 shows pyramid algorithm that provides computational efficiency in implementing wavelet transform [3].

This algorithm operates on a finite set of N input data. These data are passed through two convolution function, each of which create output stream that is half the length of the original input. One half of the output is produced by the low pass filter, and the other half is produced by the high pass filter. Low pass output contains information content and is used as an input of filtering in the next step. A high pass output contains the difference between the true input and reconstructed input. In general, high-order wavelets tent to put more information into the low-pass output. Wavelet based compression is caused by making the high-pass output to be nearly zero. A standard two dimensional Haar wavelet decomposition of an image is very simple to code. It involves one dimensional decomposition on each row of the image followed by a one dimensional decomposition on the column of the result.

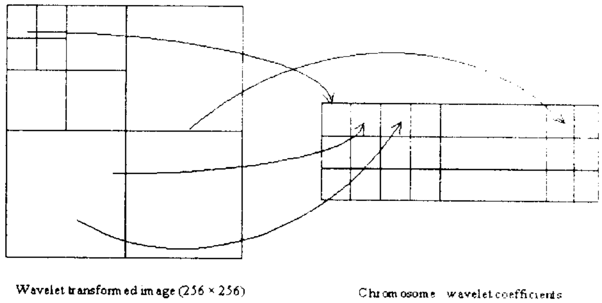


Fig 2. Chromosome using wavelet coefficients

The $r \times r$ matrix, T, through the above procedure has the average color of the image in entry $T[0,0]$ and wavelet coefficients in the other entries of T. We can reconstruct original image without loss using this information, but because there is no need to maintain information to search, we extract the largest m coefficients. The Jacob's work shows that storing the 40~50 largest-magnitude coefficients in each color works best and truncating the coefficients appears to improve the discriminatory power of the metric [4]. We only store the information to determine whether positive or negative coefficient value.

The Jacob's work compares feature of extracted data with features in the database and stores the degree of similarity between two features. Here, the average color value and the wavelet coefficients are evaluated by each criteria. An image of the highest magnitude value is provided as a result of the search. To evaluate similarity between query and target images, we use following equation.

$$\| Q,T \| = w_{0,0} \, | Q[0,0] - T[0,0] | + \sum_{i,j} w_{i,j} \, | \tilde{Q}[i,j] - \tilde{T}[i,j] |$$

Here, let us think of Q[i,j] and T[i,j] as representing a single color channel of wavelet decomposition of the query and target image. Let Q[0,0] and T[0,0] be the overall average intensity of that color channel.

4. Interactive Genetic Algorithm

Genetic algorithm is a model of machine learning derived by the procedure of evolution in nature. This is performed by creating the population of individuals that are represented by chromosome. The chromosome is the string that can be thought as the human gene. The individuals in the population go through the evolution. This procedure takes the model that is similar to the evolutionary procedure that different individuals compete for resources in the environment. Some are better than others and they are more likely to survive and propagate their genetic material.

```
t := 0;
init-population P (t);
evaluate P (t);
while not done do
    t := t + 1;
    P':= select-parents P (t);
    recombine P' (t);
    mutate P' (t);
    evaluate P' (t);
    P := survive P, P' (t);
od
```

Fig. 3 Simple genetic algorithm.

The algorithm starts with initial population. In the beginning, fitness value of each individual is evaluated to determine how appropriate the individual is for the given problem. In general, individuals in an initial population are randomly generated. Next, two individuals that have relatively high fitness value are selected from the population. Selected individuals can be regarded as a parents. New individuals called the children are created by recombinating chromosomes of parents. Here, a selection and/or a crossover operator are used. Mutation is what replaces existing gene code with randomly generated code. It performs a function that puts potentially good genes or properties which may be lost during applying genetic operation into a population.

Interactive genetic algorithm is a genetic algorithm that adopts user's choice as a fitness, when fitness function cannot be exactly determined. The property allows to develop system based on human intuition or emotion. Therefore, in this paper, system obtains fitness value of color and shape of wanted image from a user, and it is used to select good individual. Fig. 3 shows the procedure of a simple genetic algorithm.

4.1. Chromosome Representation

The information of features is represented by the following data structure.

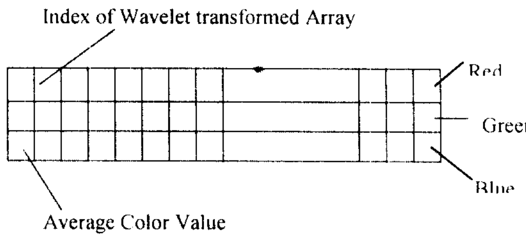


Fig. 4 Representation of chromosome.

4.2. Genetic Operators

The size of population is 12, the fitness values for shape and color are obtained from user. The strategy of selection is governed by expected frequency of each individual, and one point horizontal and vertical crossover are used. Mutation is not adopted.

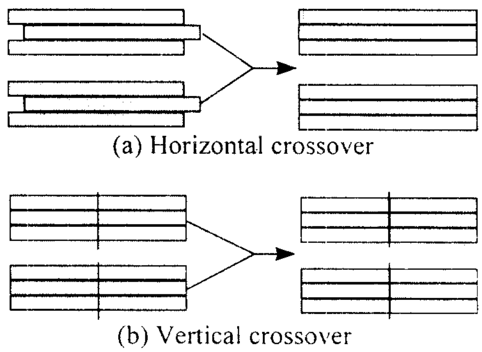


Fig. 5 Horizontal / vertical crossover operations.

4.3. Implementation

The user interface in our system is shown in Fig 6. Initial population consists of randomly selected images. A user determine color and shape fitnesses of images that are most similar to what he/she wants, then the system creates the next generation using the genetic algorithm and displays new and more similar images. This procedure is repeated until the user finds a wanted image. If the result of the next generation is worse, the system allows to back to the previous generation and different fitness can be specified. The user can increase or decrease a weight of a color. Mutation is not considered in the present implementation, because the literature reports that applying mutation to a small sized population is of little use and we are in that case. Fig. 7 shows searched images of cheerful, gloomy, and cool impressions.

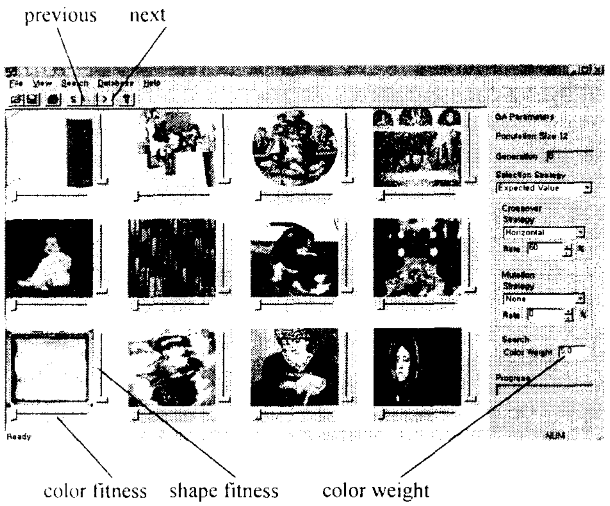
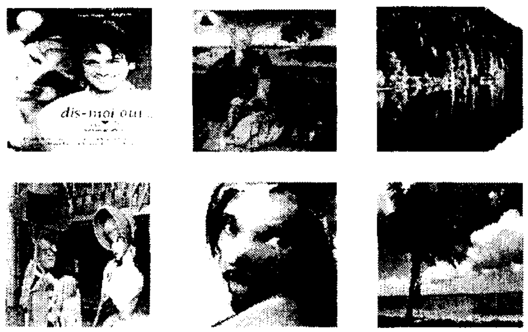


Fig. 6 The user interface.



(a) cheerful (b) gloomy (c)cool

Fig. 7 Example images obtained by the interactive genetic algorithm with respect to three emotional queries.

5. Experimental Results

5.1. Environments

The system is written in Microsoft Visual C++. It runs on Pentium PC. A searching table is constructed by a batch job over the 256×256 JPEG image. The size of image database is 170. To evaluate a performance of this system, we requested 10 graduated students to search cheerful, gloomy, and cool images and to answer the questions given below: 1) How similar the result image is what the user wants and 2) How long it takes to find.

5.2. Result Analysis

The first question is to testify whether a content-based image retrieval is efficient, we tested satisfaction with images and search times over cheerful, gloomy, and cool images. The result of test is shown in table 1 and table 2.

In this test, the 70 percents of the subjects were satisfied with finding cheerful and cool images, however, in case of gloomy image, most of the subjects were unsatisfied. A major cause is that the size of database is very small, so database does not contain images as many as users are satisfied. But the retrieval of the cool image shows that a content-based image retrieval method by human intuition or emotion can have good performance if large database is given.

In the second question, we limited the number of trial to 5. Table 2 is shown that a half of subjects could find a satisfied image before limitation, and the rest of them take five trials.

Table 1. Response statistics from 10 users for the three queries.

	cheerful	gloomy	cool
Very good	2	1	3
Good	3	0	4
Normal	2	3	0
Bad	0	2	2
Very bad	3	4	1
Total	10	10	10

Table 2. Number of trials to obtain the satisfied images for the three queries.

Trials	cheerful	gloomy	cool
1	2	0	2
2	1	2	1
3	1	1	2
4	1	3	2
5	2	2	3
over 5	3	2	0
total	10	10	10
average	4	4	3

Table 3. Comparison of Number of trials until the target samples were found.

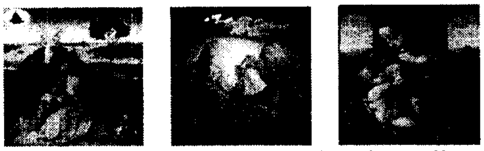
Samples	Without crossover	With crossover
1	2	1
2	over 10	5
3	over 10	over 10
4	1	1
5	over 10	3

(a) sample 1 (b) sample 2 (c) sample 3

(d) sample 4 (e) sample 5

Fig. 8 Samples used to test usefulness of crossover.

parent 1 parent 2



(a) The children crossovered horizontally.

(b) The children crossovered vertically.

Fig. 9 Example images showing the usefulness of crossover.

The second test is to inspect usefulness of crossover. Test condition is divided into two cases. One is that crossover is used, the other is that it is not used. Fig. 8 shows tested sample images. The test starts with initial population and does not use back tracking. The result is shown in table 3.

The sample 2 could not obtain a result when crossover was not used. Sample 1 and sample 2 could obtain results using only color evaluation, because they contain many part of black color, but database has a few of black images. Sample 3 is complex shape, and database contains many similar images, so it could not obtain results in all cases. In spite of relatively simple shape, sample 5 could obtain a result when crossover is not used. Fig. 9 shows example images of crossover.

6. Concluding Remarks

This paper has proposed an approach that searches an image using human intuition and emotion. This method allows to search not only explicitly expressed image, but also abstract image such as "cheerful impression image", "gloomy impression image", and so on.

Applications applied interactive genetic algorithm take an advantage of obtaining a solution that user's preference is reflected. However, the result is very subjective, so the evaluation of an efficiency and of the result is difficult. Takagi attempts to apply the subjective test based on a paired comparison to evaluate performance [1]. This paper tested efficiency using statistical method.

As you know from the shown result, a performance of a system using interactive genetic algorithm is influenced by database size [1][2]. Therefore at first, we need to obtain enough database and to get a optimized problem solution by designing a better crossover operator and adopting a mutation operator. But the larger size of a database is, the more storage and search time are needed.

Conclusively, it remains many things to do for system improvement. However, a approach using interactive genetic algorithm based on human intuition and emotion allows to develop a system for not only effective image retrieval, but also supporting a human initiative spirit.

References

[1] K. Aoki, H. Takagi and N. Fujimura, "Interactive GA-based design support system for lighting design in computer graphics," *Proc. Int. Conf. Soft Computing*, pp.533-536, 1996.

[2] H. Takagi and K. Ohya, "Discrete fitness values for improving the human interface in an interactive GA", *IEEE Computer Magazine*, Vol.28 No.9, pp109-112, 1996.

[3] T. Edwards, "Discrete wavelet transforms : Theory and implementation," *Standford University TR*, pp.1-9, September 1991.

[4] E. Jacobs, A. Findkelstein and D. H. Salesin, "Fast multiresolution image querying," *Proc. SIGGRAGH 95*, 1995.

[5] J.-H. Lee, S.-B. Cho and Y.-C. Choy, "Design and implementation of content based image retrieval system using interactive genetic algorithm", *Proc. Korea Information Science Socie Conf.*, Vol. 24, No. 1, pp 341-344, 1996.

[6] C.Caldwell and V. S. Johnston, "Tracking a criminal suspect through 'face-space' with a genetic algorithm". *4th Int'l Conf. On Genetic Algorithms*, pp.416-421,1991.

[7] E. Baker and M Seltzer, "Evolving line drawings," *Proc. Graphics Interface '94*, pp 91-100, 1994.