

Retrieval of Butterfly from Its Sketched Image on Internet

Kayo Suzuki[†], Mitsuru Nagao[†], Yoshifumi Shimodaira^{*}, and Hiroaki Ikeda[†]

Two types of image retrieval systems were constructed to compare the image retrieval using keywords to that using a sketched image. The processing time for the image retrieval using keywords was 1.0 second although that using a sketched image was 18.3 seconds. For one trial of the image retrieval, the success rate using keywords was only 18% although that using a sketched image was as high as 77%. Based on this experiment, the image retrieval using a sketched image reduced the number of data bytes sent through the internet line by a factor of one-eighth, keeping the access time unchanged.

1. Introduction

Nowadays, many users are waiting for high performance image retrieval systems for multimedia applications, and many works have been accomplished for establishing image retrieval systems^{[1]-[4]}. For instance, Xia Wan, et al., C.Oriji, et al., and Y. Sakai, et al. proposed their own image retrieval systems^{[1]-[3]}. In addition, Donald A. Adjero, et al. described requirements and issues for the multimedia database management^[4]. Requirements for the performance of the multimedia database management system (DBMS) are categorized in the literature [4]. Based on the issues which were argued in the literatures [1] through [4], the performance of the image retrieval system was measured on the image database system of butterflies to retrieve the desired images of those in the present studies. The success rate of the image retrieval using keywords was low, i.e., 18% even though the processing time for the image retrieval was short, and the success rate of the image retrieval using a sketched image was high enough, i.e., 77% even though the processing time for the image retrieval was long. These experimental results suggest that the tradeoff between the success rate and processing time of the image retrieval needs to be established.

The purpose of this study is to compare the image retrieval using keywords and to that using a sketched image so as to establish an ideal image retrieval system. In order to test the ideal image retrieval system, the image retrieval systems were experimentally constructed. One of the systems

constructed is such that the user retrieves the desired images from the image databases storing 78 species of butterfly using a set of 21 keywords on the basis of the system concept of using keywords^[5]. The other is such that the user retrieves the desired images from the image databases storing 78 species of butterfly using the image sketched by the user on the basis of the system concept of using an image^[6]. The performance of each image retrieval system was measured and evaluated in both processing time and the success rate so as to obtain an ideal image retrieval system.

2. Comparison Between Image Retrievals Using Keywords and Sketched Image

2.1 Image Retrieval Using Keywords on Database System of Butterflies

The outline of the image retrieval using keywords on the image database system is as shown in Figure 1. A sample image table is stored in an image database. The sample image table lists a set of keywords and sample images. A set of 21 keywords is attached to the respective sample images for 78 species of butterfly.

When the user retrieves the desired images of butterflies in the image database, the user specifies the keywords (See (1) in Figure 1.) and sends the specified keywords from the user terminal to the image database (See (2) in Figure 1.). The specified keywords are compared with this set of keywords which are attached to the respective sample images listed on the sample image table stored in the image database (See (3) in Figure 1.). The sample images are selected (See (4) in Figure 1.), and then displayed on the user terminal as the result of the image retrieval (See (5) in Figure 1.).

[†] Hamamatsu Research Center, Telecommunications Advancement Organization of Japan.

^{*} Shizuoka University.

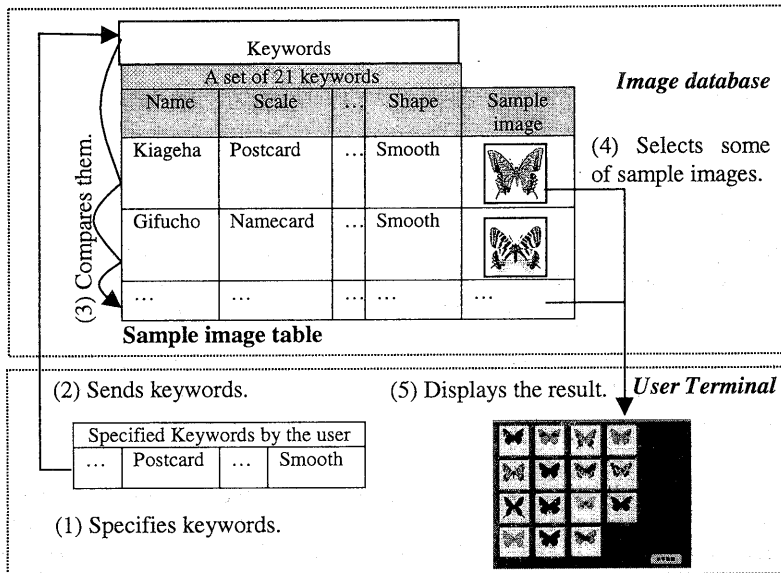


Figure 1. Outline of the image retrieval using keywords.

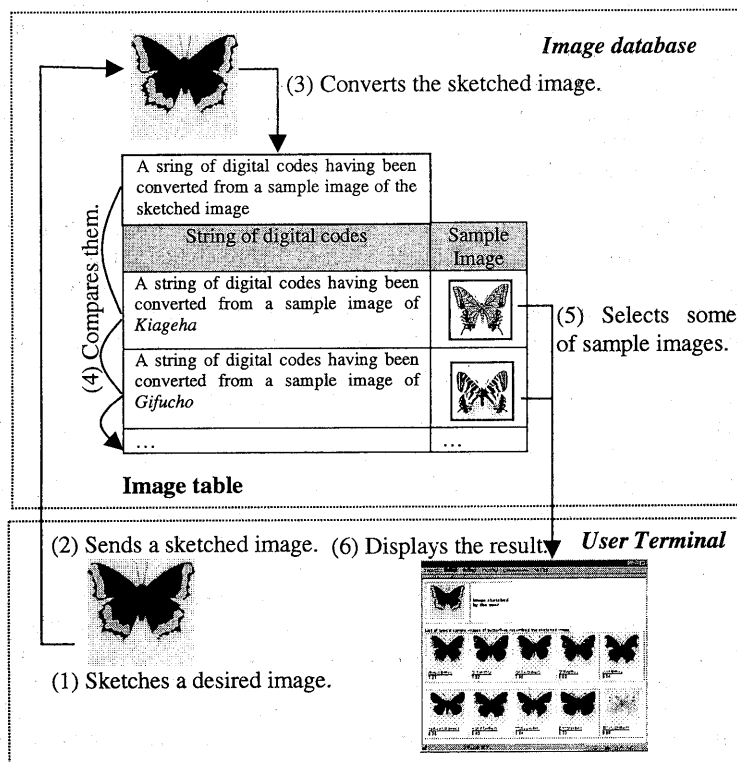


Figure 2. Outline of the image retrieval using a sketched image.

2.2 Image Retrieval Using a Sketched Image on Database System of Butterflies

The outline of the image retrieval using the image sketched by the user on the image database system is as shown in Figure 2. An image table consisting of a code table and an image table is stored in an image database. The image table lists strings of digital codes and sample images. Strings of digital codes are derived from the sample images of 78 species of butterfly.

When the user retrieves the desired images of butterflies in the image database, the user sketches the image of a desired butterfly (See (1) in Figure 2.) and sends it from the user terminal to the image database (See (2) in Figure 2.). The sketched image which is sent from the user terminal is converted to another string of digital codes (See (3) in Figure 2.). This string of digital codes which is converted from the sketched image is compared with other strings of digital codes derived from the sample images on the image table in the image database (See (4) in Figure 2.). The sample images which resemble the sketched image are selected to check whether the codes of the sample images are related to the string of digital codes of the sample image (See (5) in Figure 2.). Then, the selected sample images are displayed on the user terminal as the result of the retrieval (See (6) in Figure 2.).

3. Measurement of Performance

3.1 Processing Time

The processing time is defined as the interval from the time that the specified keywords (in Figure 1) or the sketched image (in Figure 2) is sent from the user terminal to the image databases to the time that the result of the retrieval is verified after the completion of the retrieval. Hence, the times required for operations (2) through (4) in Figure 1 and (2) through (5) in Figure 2 are included in the processing time. However, operations (1) and (5) in Figure 1, and operations (1) and (6) in Figure 2 are excluded from the considerations, since the times required for these operations depend on the operations' skill and operating programs. Then the operations (1) and (5) in Figure 1, and the operations (1) and (6) in Figure 2 are defined as those carried out during the extended processing times for the image retrieval, respectively.

Measurement was carried out using the hardware and software as listed on Table 1. The program execution flow is shown in Figure 3.

Table 1. Hardware and software used in the image retrieval.

	Hardware	Software
Image database	SUN sparc (sun4d) Memory:128Mbytes	OS: Solaris2.4
		DBMS: Illustra ORDBMS
		Image Retrieval: Illustra VIR Datablade
User terminal	IBM PC/AT	OS:Win95
		Applications: Built on Visual Basic

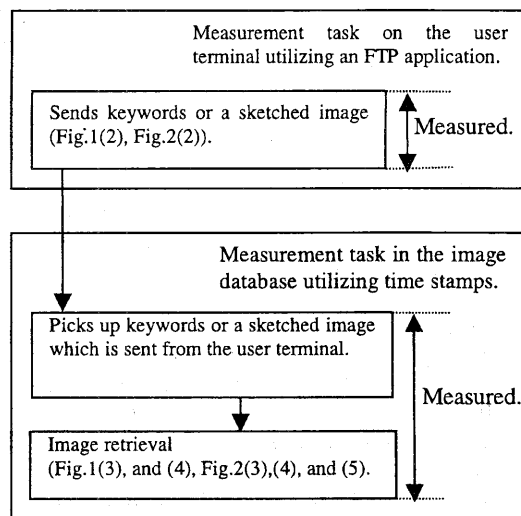


Figure 3. Measurement tasks.

3.2 Success Rate

The success rates of the image retrieval using keywords and the image retrieval using a sketched image were defined as follows. When the user took an action to retrieve each of 78 different butterflies using up to 21 keywords chosen from a set of 21 keywords, the desired images of butterflies with the first ten priorities are picked up to estimate the success rate of the image retrieval using keywords. The success rate is thus defined as the average of the success rates taken from the images with the first ten priorities. When the user took an action to retrieve each of 78 different butterflies using the image which is sketched by the user, the desired images of butterflies with the first ten priorities are picked up to estimate the success rate of the image retrieval using a sketched image. The success rate is obtained as the average of these success rates.

Table 2. Processing time and success rate for the retrieval.

Retrieval using keywords		Retrieval using an image	
(1) Specifies keywords.	*1	(1) Sketches a desired image.	*1
(2) Sends keywords.	1.0 sec.	(2) Sends a sketched image.	18.30 sec.
(3) Compares the set keywords with the keywords listed on the image table.		(3) Converts the sketched image.	
(4) Selects some of sample images.		(4) Compares the converted string of digital codes with the strings of digital codes listed on the image table.	
(5) Displays the result.	*1	(5) Selects some of sample images.	*1
(6) Displays the results			
Success rate	18%	Success rate	77%
References	Fig.1	References	Fig. 2

*1: Excluded from the processing time, and however included in the extended processing time for the image retrieval.

4. Performance

The processing time and success rate of the image retrieval are listed on Table 2.

The processing time was 1.0 second when one keyword was used for retrieving the desired images of butterflies from the image database storing 78 species of butterfly. The processing time for retrieving the desired images of butterflies from the same image database as above using a sketched image was 18.30 seconds. The success rate of the image retrieval using keywords was 18% although that of the image retrieval using a sketched image was 77 %.

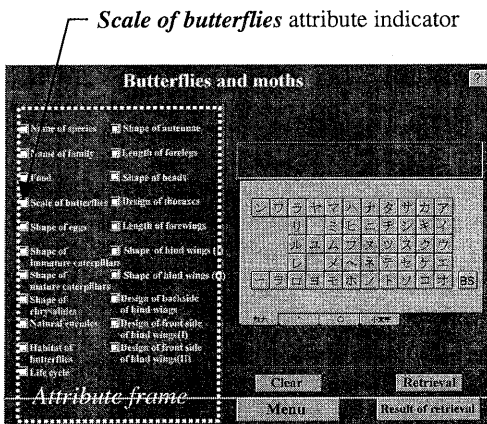
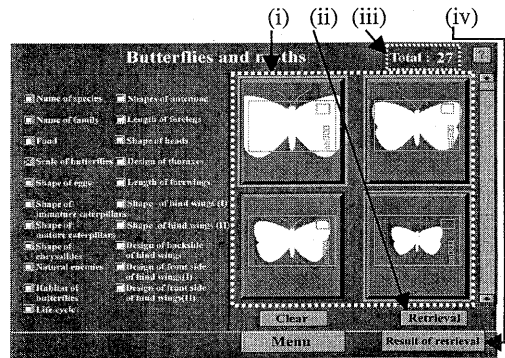


Figure 4. Attribute frame for retrieving butterflies.



- (i) Images indicating the scales of butterflies.
- (ii) Access key used for the image retrieval of corresponding butterflies.
- (iii) Number of corresponding butterflies.
- (iv) Access key used for displaying the images of corresponding butterflies on the CRT.

Figure 5. Image retrieval in accordance with the scale of butterflies.

5. Parameters

5.1 Keywords Used for Image Retrieval

Twenty one attributes of butterflies stored in the image database are regarded as keywords for the image retrieval. Figure 4 shows the attribute frame for retrieving butterflies. Twenty one indicators for these keywords are appearing on the attribute frame in Figure 4.

For example, when the user selects one of the attributes "Scale of butterflies", the images demonstrating the scale of butterflies appear as shown in Figure 5. The user selects the image indicating the scale of the desired butterfly to determine the keyword specifying the scale of the desired butterfly.

5.2 Sketched Image Used for Image Retrieval

Figure 6 shows an example of the sketched images of a butterfly used in the measurement. The images sketched by the user on the user terminal are generated using such parameters as listed on Table 3. The sketched image of Figure 6 was used to measure the processing time for the image retrieval of this mode since the sketched image of Figure 6 was of the average size of 55 Kbytes on the experimental study. An application used as a supporting tool is supplied so as to make the user easily create the sketched image with the parameters listed on Table 3.

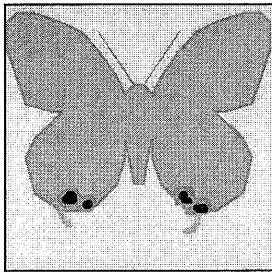


Figure 6. Example of the sketched image of butterfly.

Table 3. Parameters used to build the sketched images.

Items	Specifications
Formatting of an image of a butterfly	JPEG (high quality)
Number of pixels in a frame	512 × 512 pixels
Color of background	(R,G,B)=(230,230,230)
Shape of a butterfly	Expanding its wing as widely as it can
Size of a butterfly in a frame	Allocating a butterfly in the top of frame and expanding it in accordance with the size of frame

6. Theoretical Expressions

On the experiment of this study, the success rate of the image retrieval using keywords was 18%. When an action of the image retrieval was taken once, the success rate of the image retrieval using a sketched image was 77%. In accordance with the probability study, the success rate increases with the number of actions as discussed below.

Assume that P_k is the success rate of the image retrieval using keywords when an action of the image retrieval was taken once, and that P_s is the success rate of the image retrieval using a sketched image when an action of the image retrieval was taken once. When events of the image retrieval are independent each other, the success rate of the image retrieval using keywords P_a is given by

$$P_a = \sum_{n=1}^{\infty} (1 - P_k)^{n-1} P_k \dots (1)$$

where,

P_a : The success rate of the image retrieval using keywords when actions were successively taken n -times.

P_k : The success rate of the image retrieval using keywords when an action was taken once.

n : The number of actions of the image retrieval using keywords.

If the success rate P_a of the image retrieval using keywords when actions were successively taken n -times n is greater than the success rate P_s of the image retrieval using a sketched image when an action was taken once,

$$P_a \geq P_s \dots (2)$$

The following expression is obtained from expressions (1) and (2).

$$\sum_{n=1}^{\infty} (1 - P_k)^{n-1} P_k \geq P_s \dots (3)$$

Where, the success rate P_k was 0.18 and that P_s was 0.77 on the experiment of this study. Then, n is obtained from the expression (3) as

$$n \geq 7.4 \dots (4)$$

From the expression (4), the number of actions was 8 when the success rate of the image retrieval using keywords was greater than the success rate of the image retrieval using a sketched image.

7. Practical Use on Internet Line

The image retrieval system of butterfly was constructed on the multimedia network for internet on the basis of the proposed image retrieval using a sketched image. The user can retrieve the desired image of butterfly using a sketched image on internet.

On the experimental study, when actions were taken at least eight times during the image retrieval using keywords, the success rate became the same as that of the image retrieval using a sketched image although it was very low compared to that using the sketched image. Figure 7 shows an example of the volume of data needed for the image retrieval. Although a large amount of data, i.e., 80 Kbytes or more, is required for the images resulting from the image retrieval, whether images are retrieved using keywords or a sketched image, an action taken once for the image retrieval using a sketched image eliminates excessive data, i.e., 640 Kbytes (for the actions taken eight times), even though excessive data is needed for the image retrieval using keywords when multiple actions were taken on the internet line. The image retrieval using a sketched image reduces a large volume of data to be transferred through the internet line between the image database and user terminal. The proposed method of the image retrieval using a sketched image is thus effective for operations on the internet line since the volume of data needed for transactions is quite few compared with that needed in case of keywords, even though the success rate of the image retrieval using a sketched image is high.

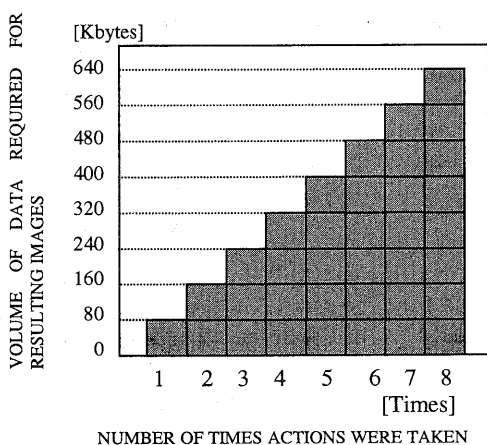
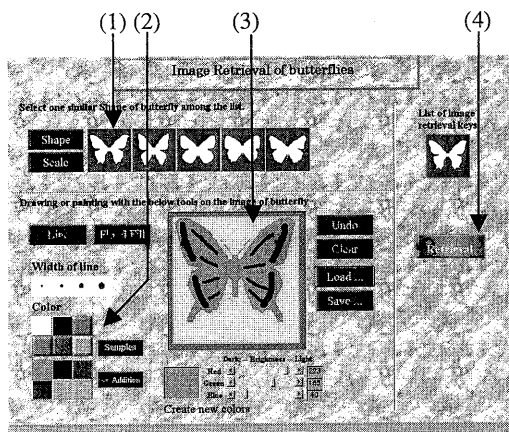


Figure 7. Volume of data required for the images resulting from the image retrieval.

As an example, Figures 8 through 10 show how to retrieve the desired image of butterfly using the image sketched on the user terminal. Figure 8 shows an application used as a supporting tool which is supplied to make the user easily create the sketched image in retrieving the desired image of butterfly. The user sketches an image of the desired butterfly on the user terminal using the application. In frame (1), control accesses to select the shape of the wings of the desired butterfly. In frame (2), control supplies the painting tool to draw the texture to the wings of the desired butterfly. Frame (3) indicate the sketching area for the image of the desired butterfly. First, the user selects the shape of the wings of the desired butterfly, and then the selected shape appears in the sketching area. Second, the user draws the texture using the painting tool to complete the sketched image of the desired butterfly. When the sketched image is complete, the user retrieves images similar to the sketched image on the access key (4) of Figure 8.



- (1) Access frame to select the shape of the wings of the desired butterfly.
- (2) Painting tool to draw the texture to the wings of the desired butterfly.
- (3) Sketching area.
- (4) Access key for the image retrieval.

Figure 8. Application for creating a sketched image.

Figure 9 shows the displayed images obtained as a result of the image retrieval. When the user specifies a sample image, then the desired image of butterfly appears as shown in Figure 10.

The images obtained as a result of the retrieval using a sketched image are demonstrated on the CRT utilizing the internet browser of Figure 9. When selecting the image of the desired butterfly among the others as shown in Figure 9, the user can access to an internet homepage built for displaying the desired images of butterfly, as shown in Figure 10.

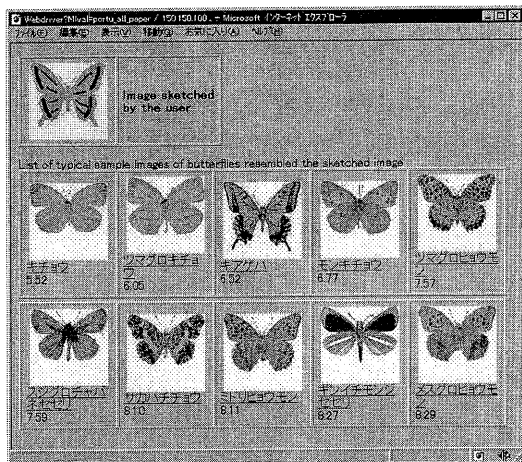


Figure 9. Result of the image retrieval using a sketched image.

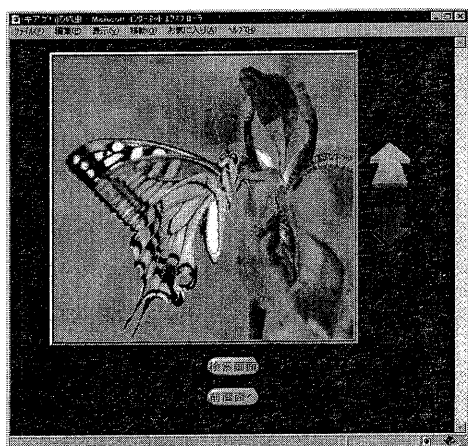


Figure 10. Homepage built for displaying the desired images of butterfly.

8. Conclusion

(1) A new measurement method for the processing time and success rate was proposed.

(2) The processing time and those success rate of the image retrieval using keywords and of the image retrieval using a sketched image are found to be evaluated.

(3) The processing time for the image retrieval using keywords was 1.0 second when one keyword was used for retrieving the desired images although the processing time was 18.3 seconds for the image retrieval using a sketched image.

(4) The success rate in the image retrieval systems was measured using the proposed measurement method on this study. The success rate for retrieving images of butterflies from the image databases storing 78 species of butterfly using keywords were 18% although that of the image retrieval using a sketched image was 77%.

(5) The success rate of the image retrieval using keywords when an action of the image retrieval was taken 8 times was greater than that of the image retrieval using a sketched image in accordance with the theoretical expressions in Chapter 6.

(6) The image retrieval using a sketched image was confirmed to be effective for operations on the internet line because only one action was taken for the image retrieval using a sketched image, which resulted in reduction of the processing time on the internet.

References

- [1] Xia Wan, et al, "IMAGE RETRIEVAL WITH AN OCTREE-BASED COLOR INDEXING SCHEME," Proceedings of the ISCAS'97, pp.1357-1360, 1997.
- [2] C.Oriji, et al, "Multimedia Object Storage and Retrieval," Proceedings of the International Symposium on Multimedia System, Information Processing Society of Japan, pp. 368-375, 1996.
- [3] Y.Sakai, et al, "Searching Multimedia Information in Distributed Environment," IEICE Transactions on Communications, Vol.E79-B, No.8, pp.989-998, August 1996.
- [4] Donald A.Adjeroh, et al, "Multimedia Database Management-Requirements and Issues," IEEE Multimedia, pp.24-33, July-September 1997.
- [5] K.Suzuki, et al, "New Type of Electronic Encyclopedia for Boys and Girls", Proc. IEEE ISCE'97, pp.300-303, (December 1997).
- [6] K.Suzuki, et al, "Retrieval of Butterfly from Its Sketched Image Utilizing Image Navigation Directory on Multimedia Network," IEEE ICECS'98. To be presented in September 1998.