Query-by-Sketch Image Retrieval Using Similarity in Stroke Order*

Takashi HISAMORI^{†,††}, Nonmember, Toru ARIKAWA^{†a)}, Student Member, and Gosuke OHASHI[†], Member

SUMMARY In previous studies, the retrieval accuracy of large image databases has been improved as a result of reducing the semantic gap by combining the input sketch with relevance feedback. A further improvement of retrieval accuracy is expected by combining each stroke, and its order, of the input sketch with the relevance feedback. However, this leaves as a problem the fact that the effect of the relevance feedback substantially depends on the stroke order in the input sketch. Although it is theoretically possible to consider all the possible stroke orders, that would cause a realistic problem of creating an enormous amount of data. Consequently, the technique introduced in this paper intends to improve retrieval efficiency by effectively using the relevance feedback by means of conducting data mining of the sketch considering the similarity in the order of strokes. To ascertain the effectiveness of this technique, a retrieval experiment was conducted using 20,000 images of a collection, the Corel Photo Gallery, and the experiment was able to confirm an improvement in the retrieval efficiency.

key words: data mining, Expected Search Length (ESL), image retrieval, relevance feedback

1. Introduction

The use of large image databases is becoming increasingly popular with the popularization of digital consumer electronic products and high-speed networks in recent years, and it is becoming desirable to construct a system that enables efficient retrieval of images that users want. Against such a backdrop, various studies on content-based image retrieval have been conducted with great success [1]–[15]. However, bridging the semantic gap between low-level visual features and high-level users' subjective views remains an unresolved task of image-retrieval research. Recently, a retrieval technique using relevance feedback to reflect a user's intention on the system is being introduced, achieving positive effects [16]–[24]. The query-by-sketch image retrieval method using the relevance feedback proposed by the authors has made it possible to bridge the image and the sketch by correlating the input sketch with the relevant image through a user's own relevance evaluation, thus reducing the semantic gap [23] and [24]. The authors also made it possible to display the retrieval result automatically by each

stroke of a sketch, noting the fact that the sketch used in a query is drawn using on/off signals from input devices such as a mouse or stylus, by reading the sketch and retrieving simultaneously the switching of signals from on to off without burdening the user. Moreover, it is now possible to retrieve relevant images with a fewer numbers of strokes by correlating all the sketches entered by each stroke with images. However, the fact that the effects of relevance feedback according to the correlated sketch are grossly dependant on the order of strokes is expected to remain a problem. Although it is theoretically possible to consider all the possible orders of strokes, that would cause a realistic problem of creating an enormous amount of data. Consequently, the technique introduced in this paper intends to improve retrieval efficiency by effectively using relevance feedback by means of conducting data mining of the sketch considering the similarity in the order of strokes. This paper consists of six sections including this section, in which the backdrop of the study and the purpose of the study are described. The overview of the query-by-sketch image retrieval is described in Sect. 2 and the technique using data mining is described in Sect. 3. The image retrieval experiment and the obtained results are described in Sect. 4, and the discussion of the result of data mining is described in Sect. 5. The conclusion of the paper is described in Sect. 6.

2. Query-by-Sketch Image Retrieval

2.1 Overview

We first present an overview of the previously proposed query-by-sketch image retrieval method [23]. Figure 1 shows an overview of the query-by-sketch image retrieval method. The process, corresponding to the numbers in Fig. 1, is as follows: 1) The user inputs a sketch image into the system as a query; 2) The system extracts the feature from the sketch; 3) The system calculates the Euclidean distance between the features of the sketch and the features in the feature database; 4) The system outputs images according to the value for the distance in ascending order. We use an image feature extraction method focusing on relative positional relations among binary line images. Figure 2 shows the proposed Relative Direction Frequency (RDF) feature extraction process. On the pixel of interest, the number of edge pixels in each of eight regions, determined as shown in Fig. 2 (a), is counted. The number of pixels in each region is denoted as c_x , as shown in Figs. 2 (b) and 2 (c). Then,

Manuscript received September 9, 2009.

Manuscript revised January 4, 2010.

[†]The authors are with the Department of Electrical and Electronic Engineering, Shizuoka University, Hamamatsu-shi, 432–8561 Japan.

^{††}The author is with Increment P Corp, Tokyo, 153–8665 Japan.

*This work was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan under grant number 21700155.

a) E-mail: f0930075@ipc.shizuoka.ac.jp

DOI: 10.1587/transinf.E93.D.1459

Fig. 1 Overview of the previous proposed query-by-sketch image retrieval using relevance feedback. (1) The user inputs a sketch into this system. (2) The system extracts the feature. (3) The system calculates the similarity measure. (4) The system outputs retrieval results. (5) The user identifies positive images. (6) The system performs feedback. (7) The user inputs a sketch. (8) The system extracts the feature. (9) The system calculates the similarity measure. (10) The system outputs retrieval results. When the distance between the feature of the sketch and the relevant features is similar, the system outputs relevant images.

Fig. 2 Relative Direction Frequency feature extraction process. (a) Example of edge images, (b) and (c) sum of edge pixels in each region, (d) and (e) normalization of edge pixels, (f) and (g) value of binary pattern b_x is calculated by the thresholding technique, (h) binary to decimal conversion, and (i) histogram.

s_x is the normalized value obtained by Eq. (1), as shown in Figs. 2 (c) and 2 (f). The value of the binary pattern b_x is calculated by the thresholding technique using Eq. (2), as shown in Figs. 2 (d) and 2 (g). The values of C and Th denote the total number of line pixels and the threshold value, respectively. The threshold value is determined empirically to be 0.15.

$$s_x = \begin{cases} \frac{c_x}{C-1} & (C \neq 0) \\ 0 & (C = 0) \end{cases} \quad (1)$$

$$f_x = \begin{cases} 0 & (s_x \leq Th) \\ 1 & (s_x > Th) \end{cases} \quad (2)$$

Fig. 3 Overview of the semantic gap and the relevant label. To reduce the semantic gap between low-level features and high-level semantics, user sketches play an important role in relevance feedback. When users mark positive images among output images with “relevant” label, the input sketch are relevant to these images as relevant label in feedback.

By expanding the concept of the Local Binary Pattern (LBP) [15] in the field of texture analysis, the values “0” or “1” in each direction are arranged counterclockwise from f_0 to express an eight-bit binary number. The binary number ($b_7, b_6, b_5, b_4, b_3, b_2, b_1, b_0$) is transformed into a decimal number d ($0 < d < 255$), as shown in Fig. 2 (h). This process enables the relative position of edge pixels to be preserved. We then vote for the corresponding bin d in the histogram shown in Fig. 2 (i). A 256-dimensional image feature is obtained by applying the above-mentioned process to all edge pixels and normalizing the processed pixels. This process enables shift-invariance and scale-invariance to be realized. Rotation-invariance and symmetry-invariance are easily represented by shifting the binary number.

2.2 Relevance Feedback

In query-by-sketch image retrieval using relevance feedback, it is possible to reduce the semantic gap by marking images with a relevant label. Figure 3 shows an overview of the semantic gap and the relevant label for “human” image. The input sketch in Fig. 3 is drawn using seven hiragana (four characters from the Japanese hiragana syllabary: “へ,” “の,” “も,” and “じ” are used), and the most popular “human face” sketch. The two upper “へ” express the eye-brows, the two “の” express the eyes, “も” expresses nose, the lower “へ” expresses the mouth, and “じ” expresses the line of the face. Because “へ の へ の も へ じ” is a very simple sketch, it is often used as a sketch to express “human.” Therefore, in this paper, if “へ の へ の も へ じ” is used as an input sketch, not only “face” images, but also “human” images are considered to be the retrieval target.

In Fig. 3, the high-level semantics of the input sketch are similar to that of the “human” image. However, the low-level features of the input sketch are not necessarily similar to that of the “human” image. The low-level feature of the input sketch must be more similar to that of a previously in-

put sketch than that of the “human” image. If the previously input sketch is relevant to the “human” image, the use of the previously input sketch in relevance feedback enables the retrieval performance to be improved, because the previously input sketch plays an important role in bridging the semantic gap between the input sketch and the “human” image. The process of query-by-sketch image retrieval using relevance feedback corresponding to the numbers in Fig. 1 is as follows: 5) The user indicates positive images among the output images; 6) The system makes the feature of the sketch relevant to the features of the “relevant” images as feedback; 7) The user then inputs a sketch into the system as a query; 8) The system extracts the feature from the sketch; 9) The system calculates the Euclidean distance between the features of the sketch and the features in the feature database, including the relevant features; 10) The system outputs images according to the value of the distance in ascending order. When the feature of the sketch is similar to the relevant features, the system outputs relevant images.

3. Query-by-Sketch Image Retrieval Using Data Mining

3.1 Utilization of Each Stroke and Stroke Order of the Sketch

The achievement of query-by-keyword retrievals with query prediction make possible to reduction of the burden of the user, because the user can retrieve it without inputting the keyword completely. As with query-by-keyword retrievals, the request for the application of the query prediction to query-by-sketch image retrieval is increased. The input sketch in query-by-sketch image retrieval consists of each stroke and stroke order. Therefore, authors propose the running retrieval method paying attention to each stroke and stroke order of the sketch, in order to achieve the prediction of the images user want. Figure 4 shows the overview of

Fig. 4 Overview of the running retrieval method, paying attention to each stroke and stroke order of the sketch. By correlating all the sketches input in the past with the image by each stroke, it becomes possible to obtain sketches similar to the input sketch with fewer numbers of strokes from the correlated sketches.

the running retrieval method. Figure 4 shows the overview of the running retrieval method, paying attention to each stroke and stroke order of the sketch. It becomes possible to retrieve the image using the sketch as the bridge by correlating all the sketches input in the past with the image by each stroke, obtaining sketches similar to the input sketch with fewer numbers of strokes from the correlated sketches. However, this has a problem in that the relevance feedback effect cannot be achieved easily if there is a slight deviation in the stroke order in the input sketch. A simple method here to obtain the relevance feedback effect without dependence on the stroke order would be to have all the sketches that have all conceivable stroke orders correlated with the image. Table 1 shows the relation between the number of strokes of the sketch and the number of patterns of all conceivable stroke orders. The correlation procedure is easy as the number of stroke order patterns is small if the number of strokes of the sketch is small, but the number of stroke order patterns increases enormously as the number of strokes of the sketch increases, making the process of correlating all the patterns difficult.

Thus, the authors decided to pay attention to the partial similarity of the stroke order of the sketch. Figure 5 shows an example of a sketch where the subject matter is a “car.” While the wheels are drawn first and the body next in the sketch An, the body is drawn first and the wheels next in the sketches Bn and Cn. Additionally, the total number of strokes and the final shapes are different among the sketches A9, B5 and C4. As can be seen here, the sketches provide

Table 1 The relation between the number of strokes of the sketch and the number of patterns of all conceivable stroke orders.

Strokes of sketch	1	2	3	...	10	...	n
Patterns of stroke orders	1	2	6	...	3,628,800	...	$n!$

Fig. 5 An example of a sketch where the subject matter is a “car.” The total number of strokes and the final shapes are different among the sketches A9, B5 and C4. However, if we look at the sketches paying attention to the differences in these sketches every two strokes, i.e., the sketches of every other stroke, such as an, bn, and cn, the similarity that the wheels are drawn continuously is found from a2, b5 and c4.

Fig. 6 The retrieval objects when the difference sketch a2 is used as an input. The retrieval objects here are difference sketches bn and cn, which are different from the input in terms of types, and the difference sketches with the most similar feature are sought for among them, respectively.

a low probability of similarity as there is a big difference in the stroke order and the total number of strokes, depending on the user, so it is difficult to use it efficiently in combination with the relevance feedback. However, if we look at the sketches paying attention to the differences in these sketches every two strokes, i.e., the sketches of every other stroke, such as an, bn, and cn, the similarity that the wheels are drawn continuously is found from a2, b5 and c4. If such difference sketches can be extracted from the accumulated difference sketches, an efficient utilization of them in combination with the relevance feedback can be expected.

3.2 Frequent Pattern Mining of the Sketch

In order to extract similar difference sketches from the accumulated difference sketches in this method, we obtained an indicator that indicates how similar each difference sketch is to other difference sketches by focusing on the similarity relations of the difference sketch features. Similarity retrieval using each difference sketch as an input was conducted as a round robin in order to retrieve similar difference sketches. Figure 6 shows the retrieval objects when the difference sketch a2 is used as an input. The retrieval objects here are difference sketches bn and cn, which are different from the input in terms of types, and the difference sketches with the most similar feature are sought for among them, respectively. Similarity retrieval is further conducted using the difference sketches thus obtained by the search as the input. Figure 7 shows the overview of the similarity difference sketch retrieval method. In Fig. 7, the arrows indicate that the sketch at the start point is the sketch used as the input in the similarity retrieval, and the sketch at the end point is the sketch having the shortest Euclidean distance from the input sketch feature quantity, i.e., the sketch with a similar feature quantity.

Figure 7 (a) is a case where the focus is on the difference sketch a2 of Fig. 6, showing that the feature quantity of a2 is the most similar to the feature quantity of b5 among all bn, and it is the most similar to the feature quantity of

Fig. 7 Overview of the similarity difference sketch retrieval method. (a) is a case where the focus is on the difference sketch a2 of Fig. 6, (b) is a case where the focus is on the difference sketch a3 of Fig. 6. This method pays attention to the number of sets of arrows indicating said similarities, and uses it as the indicator of the relative similarity relations of the difference sketch in focus.

c4 among all cn. It also shows that the feature quantity of b5 is the most similar to that of a2 among all an, and to that of c4 among all cn, while the feature quantity of c4 is the most similar to that of a2 among all an, and to that of b5 among all bn. It can also be seen here that the difference sketches a2, b5, and c4 are in relation similar to each other. This method pays attention to the number of sets of arrows indicating said similarities, and uses it as the indicator of the relative similarity relations of the difference sketch a2 in focus. Figure 7 (b) is a case where the focus is on the difference sketch a3 of Fig. 6, showing that the feature quantity of a3 is the most similar to the feature quantity of b4 among all bn, and it is the most similar to the feature quantity of c4 among all cn. It also shows that the feature quantity of b4 is the most similar to that of a7 among all an, and to that of c4 among all cn, while the feature quantity of c4 is the most similar to that of a2 among all an, and to that of b5 among

Table 2 The values T of the difference sketches.

an
T	3	0	1	2	1	0	0	0
bn
T	2	0	0	3				
cn
T	0	1	3					

all bn. It can also be seen here that the difference sketches a3, b4, and c4 are not in relations similar to each other and the number of arrow pairs that are opposite to each other is also zero.

Thus, by applying similar process to each one of the difference sketches and counting the number of arrow pairs, the value T indicating the relative similarity relation can be obtained for each of them. The values T of the difference sketches in Fig. 6 are shown in Table 2. If there are N kinds sketches correlated to a certain image, the maximum value T_{max} of T can be expressed by Eq. (3).

$$T_{max} = \sum_{k=1}^N (k-1) = \frac{N(N-1)}{2} \quad (3)$$

In this method, it is decided to extract the difference sketch if T assumes a value in excess of a certain ratio relative to T_{max} , in other words, if it satisfies the Eq. (4).

$$T \geq T_{th} = \alpha T_{max} \quad (0 \leq \alpha \leq 1) \quad (4)$$

If the sketches An, Bn and Cn are correlated with the “car” image, the three sketches an, bn and cn become the objects of frequent pattern mining, a kind of data mining, so that the maximum value of T is 3 from the Eq. (1). As the value T of the difference sketch a2 in Fig. 7(a) is 3, a2 is extracted as one of the similar difference sketches when α of the Eq. (4) is assumed 1. On the other hand, as the value T of the difference sketch a3 in Fig. 7(b) is zero, a3 is no extracted as a similar difference sketch. As such, it is possible to extract only the sketches that are similar to each other. An optimum coefficient α that provides the relevance feedback to many images with a small number of difference sketches was searched for in the preliminary experiment, and the value α of 0.8 was adopted based on said experiment for the retrieval experiment to be described in Sect. 4.

3.3 Application to Image Retrieval

Figure 8 shows the overview of the query-by-sketch image retrieval method using data mining of difference sketches. In

Fig. 8 Overview of the query-by-sketch image retrieval method using data mining of difference sketches. The retrieval process is conducted by matching and comparing the feature quantity of the difference input sketch with the feature quantity of each difference sketch extracted by means of the data mining, in parallel with the sketch image retrieval method without using data mining.

this method, the retrieval process is conducted by matching and comparing the feature quantity of the difference input sketch with the feature quantity of each difference sketch extracted by means of the data mining, in parallel with the sketch image retrieval method without using data mining. Even if the overall stroke order of the input sketch is different from that of the input sketch in the past as a result of conducting the two retrieval processes in parallel, the image correlated with the particular sketch is retrieved so long as they are similar partially. In other words, it is possible to indicate the retrieval result efficiently even if the sketch such as a “car,” in which there can be substantial differences in the stroke order or the total number of strokes depending on the users. The authors thought here that it is possible to alleviate the psychological burden on the user by not only retrieving the images efficiently by indicating the retrieval results by means of difference input sketches, but also visually devising a better indication method for the result. Therefore, the retrieval result by the difference input sketches is indicated using a popup window in this method. It is believed that the psychological burden of retrieval on the user can be alleviated by not just displaying the retrieval result, but also by displaying it with emphasis through a popup window. Moreover, as a non-relevant image can be retrieved in a higher position if the Euclidean distance d between the input difference sketch and the difference sketch correlated with the image is relatively shorter even though it is absolutely longer, it is decided in this method to display only those images that satisfy a condition d less than the threshold value. In the initial experiment, an optimum threshold value d_{th} is retrieved that does not degrade the retrieval efficiency, and adopted a value of 0.2 for d_{th} in the retrieval experiment to be described in Sect. 4.

4. Experimental Results

In order to confirm the effectiveness of this method, retrieval experiments were conducted on 20,000 images of a collection, the Corel Photo Gallery, using “glass,” “car,” and “human” as the retrieval targets. The images judged by all three subjects as relevant are defined as relevant images, and consisted of 93, 748 and 842 “glass,” “car,” and “human” relevant images, respectively. Meanwhile, the “glass,” “car,” and “human” sketches correlated by the relevance evaluation were 1,715, 2,276 and 4,851 sketches, respectively. The relevance feedback and the data mining are the method to improve the performance of query-by-sketch image retrieval. To confirm the effectiveness of this method, it is essentially necessary to compare the retrieval efficiency of this method with the state-of-the-art methods. However, as stated in Ref. [26], it is difficult to simply compare the retrieval efficiency of this method with other methods, because the effect of the relevance feedback and/or the data mining is derived from combination with performance of the image retrieval. Therefore, in this retrieval experiment, three retrieval results of method A “without relevance feedback without data mining,” method B “with relevance feedback without data mining,” and method C “with relevance feedback with data mining” are compared in order to confirm the pure effectiveness of relevance feedback and data mining in query-by-sketch image retrieval. As the experiment method, five correct images were randomly set up from the retrieval target images to which difference sketches were correlated, and one correct image was set up for each retrieval. In other words, the retrieval experiment was conducted in such a way as to find the preset correct image with a minimum number of strokes. It was ruled here that when a correct image was found in the retrieval result initial image, i.e., the top 20 main windows, or top 8 popup windows, it was considered that the correct image was found; and the retrieval was considered as completed when either the correct im-

age was found, or when the sketch was completed. Also, five sketches collected from subjects were used as the input. Figures 9, 10 and 11 are the retrieval result examples for the retrieval targets of “glass,” “car,” and “human” respectively, showing the input sketch and the difference of the input sketch at the top left and the bottom left, respectively. For a comparison purpose, Figs. 9 (a), 10 (a) and 11 (a) represent the retrieval results of method A “Query-by-sketch image retrieval without relevance feedback without data mining,” Figs. 9 (b), 10 (b) and 11 (b) represent the retrieval results of method B “Query-by-sketch image retrieval with relevance feedback without data mining,” and Figs. 9 (c), 10 (c) and 11 (c) represent the retrieval results of method C “Query-by-sketch image retrieval with relevance feedback with data mining,” respectively. Also, Tables 3, 4 and 5 show the rate of finding correct images and the average number of strokes when the retrieval was completed when “glass,” “car,” and “human” were used as the retrieval targets respectively.

It can be seen here that while no relevant images concerning “glass” were found as shown in Fig. 9 (a), 20 relevant images concerning “glass” were found thanks to the relevance feedback effect, as shown in Fig. 9 (b). In Fig. 10 (c), two images are indicated in the popup window in addition to the result of Fig. 9 (b) thanks to the data mining effect. Moreover, it can be seen from Table 3 that the correct image finding rate increased when the relevance feedback is used, but no increase is found in the retrieval result due to the use of the data mining. It is believed that this is due to the fact that there was a sufficient effect of relevance feedback without having to use data mining, judging from the fact that the average number of strokes in method B “Query-by-sketch image retrieval with relevance feedback without data mining” was as small as 2.7. In Fig. 10 (a), there was only one relevant image concerning “car” found at the 1st rank, and no change was found in the upper rank images even when the relevance feedback was used. However, in Fig. 10 (c), five images are indicated in the popup window in addition to the result of Fig. 10 (b) thanks to the data mining effect.

Fig. 9 Examples of retrieval results for “glass” images. (a) Query-by-sketch image retrieval without relevance feedback without data mining. (b) Query-by-sketch image retrieval with relevance feedback without data mining. (c) Query-by-sketch image retrieval with relevance feedback with data mining.

Table 3 Detection rate of target images and average strokes for “glass” image.

	A	B	C
Detection rate of target images (%)	8.0	32.0	32.0
Average strokes	3.4	2.7	2.7

Table 4 Detection rate of target images and average strokes for “car” image.

	A	B	C
Detection rate of target images (%)	0	0	32.0
Average strokes	4.6	4.6	3.6

Table 5 Detection rate of target images and average strokes for “human” image.

	A	B	C
Detection rate of target images (%)	0	40.0	48.0
Average strokes	11.0	7.7	7.0

(a)

(b)

(c)

Fig. 10 Examples of retrieval results for “car” images. (a) Query-by-sketch image retrieval without relevance feedback without data mining. (b) Query-by-sketch image retrieval with relevance feedback without data mining. (c) Query-by-sketch image retrieval with relevance feedback with data mining.

(a)

(b)

(c)

Fig. 11 Examples of retrieval results for “human” images. (a) Query-by-sketch image retrieval without relevance feedback without data mining. (b) Query-by-sketch image retrieval with relevance feedback without data mining. (c) Query-by-sketch image retrieval with relevance feedback with data mining.

It can be confirmed that all five images indicated are relevant images concerning “car.” Also, although no increase due to the use of the relevance feedback can be found in the retrieval result in Table 4, it is possible to confirm that the retrieval is completed with a fewer number of strokes as it was possible to find correct images by using the data mining. While three relevant images concerning “human” were found at the 11th, 15th and 19th positions in Fig. 11 (a), four relevant images were found at the 11th, 15th, 19th and 20th ranks thanks to the effect of the relevance feedback in Fig. 11 (b). Moreover, in Fig. 11 (c), one image is indicated

in the popup window in addition to the result of Fig. 11 (b) thanks to the data mining effect. The particular indicated image can be confirmed as a relevant image concerning “human.” Also, it can be confirmed from Table 5 that the correct image finding rate increases as a result of using the relevance feedback and that the retrieval is completed with a fewer number of strokes, thus proving that the retrieval rate improved by using the data mining.

The Recall-Precision graph is normally used as a scale of quantitative evaluation of the image retrieval method. Most of the evaluation scales including the Recall-Precision

Fig. 12 Average ESL for "glass" images. "A" is Query-by-sketch image retrieval without relevance feedback without data mining. "B" is Query-by-sketch image retrieval with relevance feedback without data mining. "C" is Query-by-sketch image retrieval with relevance feedback with data mining. (a) ESL graph (ESL 1600). (b) The enlarged partial graph in graph (a) (ESL 100).

Fig. 13 Average ESL for "car" images. "A" is Query-by-sketch image retrieval without relevance feedback without data mining. "B" is Query-by-sketch image retrieval with relevance feedback without data mining. "C" is Query-by-sketch image retrieval with relevance feedback with data mining. (a) ESL graph (ESL 1600). (b) The enlarged partial graph in graph (a) (ESL 100).

Fig. 14 Average ESL for "human" images. "A" is Query-by-sketch image retrieval without relevance feedback without data mining. "B" is Query-by-sketch image retrieval with relevance feedback without data mining. "C" is Query-by-sketch image retrieval with relevance feedback with data mining. (a) ESL graph (ESL 1600). (b) The enlarged partial graph in graph (a) (ESL 100).

Fig. 15 Example of data mining result for “human” sketches. (a) Difference sketches (b) Extracted sketches by data mining.

Fig. 16 Example of data mining result for “glass” sketches. (a) Difference sketches (b) Extracted sketches by data mining.

Fig. 17 Example of data mining result for “car” sketches. (a) Difference sketches (b) Extracted sketches by data mining.

graph can be said to evaluate the retrieval-system side judging from the fact that they all measure the convenience of retrieval. On the other hand, the Expected Search Length (ESL) is proposed as a scale for evaluating user-side cost in the retrieval [27], [28]. As described in Sect. 3, this method is considered to serve as a means of alleviating the psychological burden on the user as it emphatically indicates the retrieval result in the popup window. Therefore, this method provides the quantitative evaluation using the ESL evaluation scale that handles the user’s retrieval cost, not the Recall-Precision graph that handles the output side of the retrieval system. ESL represents the non-relevant image number that needs to be examined until the user becomes satisfied. Considering a graph with the number of images that satisfy the user on the horizontal axis and ESL on the vertical axis, a detailed analysis becomes possible in coordination with the change of the number of images that satisfy the user [28]. Figures 12 (a), 13 (a) and 14 (a) show the av-

erage ESL when “glass,” “car,” and “human” are used as the retrieval targets, respectively, while Figs. 12 (b), 13 (b) and 14 (b) show the enlarged partial graphs. The graph legends are method A “Query-by-sketch image retrieval without relevance feedback without data mining,” the method B “Query-by-sketch image retrieval with relevance feedback without data mining,” and the method C “Query-by-sketch image retrieval with relevance feedback with data mining,” respectively. From Fig. 12 (a), we can see that the graphs of the method B where the relevancy feedback is used and the graph of the method C lie below the graph A where the relevancy feedback is not used, when “glass” is used as the retrieval target. Since ESL represents the user’s retrieval cost, it can be confirmed that the user’s retrieval cost is reduced by using the relevance feedback. Also, from Fig. 12 (b), it is obvious that there was no data mining effect to speak of as the graph of the method B and the graph of this method C are approximately equal, and it is confirmed that they match

with the tendency of the result of Table 3. From Figs. 13 (a) and 14 (a), where the retrieval targets were “car” and “human,” one can see that the use of the relevancy feedback reduces the user’s retrieval cost judging from the fact that the graphs of the method B and this method C, both of which using the relevancy feedback, are located below the graph of the method A. Moreover, one can also see from Figs. 13 (b) and 14 (b) that the user’s retrieval cost was reduced as a result of the data mining judging from the fact that the graph of this method C is located below the graph of the method B. Thus, we see that this method is a retrieval method that reduces the burden on the user.

With regard to the retrieval speed, the average retrieval speed is 416 msec (on an Intel Core 2 Duo 1.83 GHz, 2.00 GB RAM), so that it is possible to retrieve images in real-time.

5. Discussion

Difference sketches extracted by using the data mining are discussed. Figures 15, 16 and 17 are examples of the result of the data mining of difference sketches of “human”, “glass” and “car” respectively. Figure 15 shows that difference sketches having a set of a mouth and a nose are extracted from sketches of “human”. Figure 15 (b) shows that relationships between these two objects are vertical in the difference sketches. Figure 16 shows that difference sketches having a stem of glass are extracted from sketches of “glass”. Figure 16 (b) shows that relationships between these two objects are vertical in the difference sketches, like sketches of “human”. Figure 17 shows that difference sketches having a pair of wheels and/or a set of one wheel and body are extracted from sketches of “car”. Figure 17 (b) shows that relationships between these two objects are horizontal in the difference sketches. The direction of “human” and “glass” sketches is vertical; on the other hand, the direction of “car” sketches is horizontal. Thus, it is shown that the directions of the sketches are identical with those of the difference sketches extracted using the data mining. It is inferred from these results that the users tend to draw the sketch by the stroke order corresponding to the directionality of the object. Therefore, the results in Fig. 15, 16 and 17 suggest correlation between shape and the stroke order of sketches.

6. Conclusion

In this paper, we proposed to improve retrieval efficiency by effectively using the relevance feedback by means of conducting data mining of the sketch considering the similarity in the order of strokes. Query-by-sketch image retrieval experiment using similarity in stroke order was conducted using 20,000 images of a collection, the Corel Photo Gallery, and this method provided the quantitative evaluation using the ESL evaluation scale that handles the user’s retrieval cost. As a result, it is possible to find relevant images with a fewer number of strokes by data mining using similarity

in stroke order, and it could be confirmed from ESL that the user’s retrieval cost is reduced.

Acknowledgements

The authors wish to thank Dr. Y. Shimodaira of Shizuoka University for his advices.

References

- [1] R. Datta, D. Joshi, J. Li, and J.Z. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” *ACM Comput. Surv.*, vol.40, no.2, article 5, pp.1–60, 2008.
- [2] D. Forsyth, J. Malik, and R. Wilensky, “Searching for digital pictures,” *Scientific American*, vol.276, no.6, pp.88–93, 1997.
- [3] M.S. Kankanhalli, B.M. Mehtre, and H.Y. Huang, “Color and spatial feature for content-based image retrieval,” *Pattern Recognit. Letters*, vol.20, pp.109–118, 1999.
- [4] M. Abdel-Mottaleb and N. Dimitrova, “CONIVAS: CONTENT-based image and video access system,” *ACM Multimedia 96*, pp.427–428, 1996.
- [5] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, “Query by image and video content: The QBIC system,” *Computer*, vol.28, no.9, pp.23–32, 1996.
- [6] A. Pentland, R.W. Picard, and S. Sclaroff, “Photobook: Content-based manipulation of image databases,” *International Journal of Computer Vision*, vol.18, no.3, pp.233–254, 1996.
- [7] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, and T.S. Huang, “Supporting similarity queries in MARS,” *Proc. 5th ACM International Multimedia Conference*, pp.403–413, 1997.
- [8] W.-Y. Ma and B.S. Manjunath, “Netra: A toolbox for navigating large image databases,” *Multimedia Systems*, vol.7, no.3, pp.184–198, 1999.
- [9] Z. Yang and C.-C.J. Kuo, “A semantic classification and composite indexing approach to robust image retrieval,” *Proc. IEEE International Conference on Image Processing*, 1999.
- [10] K. Hirata and T. Kato, “Query by visual example,” *Extending Database Technology ’92*, pp.56–71, 1992.
- [11] A.K. Jain and A. Vailaya, “Shape-based retrieval: A case study with trademark image databases,” *Pattern Recognit.*, vol.31, no.9, pp.1369–1390, 1998.
- [12] M. Abdel-Mottaleb, “Image retrieval based on edge representation,” *Proc. IEEE International Conference on Image Processing*, 2000.
- [13] B.M. Mehtre, M.S. Kankanhalli, and W.F. Lee, “Shape measure for content based image retrieval: A comparison,” *Pattern Recognit.*, vol.33, no.3, pp.319–337, 1997.
- [14] F. Mokhtarian and A. Mackworth, “Scal-based description and recognition of planar curves and two-dimensional shapes,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.8, no.1, pp.34–43, 1986.
- [15] T. Ojala, M. Pietikainen, and D. Harwood, “A comparative study of texture measures with classification based on feature distributions,” *Pattern Recognit.*, vol.29, no.1, pp.51–59, 1996.
- [16] F. Jing, M. Li, H.-J. Zhang, and B. Zhang, “An efficient and effective region-based image retrieval framework,” *IEEE Trans. Image Process.*, vol.13, no.5, pp.699–709, 2004.
- [17] W. Niblack, X. Zhu, J. Hafner, T. Breuel, D. Ponceleon, D. Petkovic, M. Flickner, E. Upfal, S. Nin, S. Sull, B. Dom, B.-L. Yeo, S. Srinivasan, D. Zivkovic, and M. Penner, “Updates to the QBIC system,” *Proc. SPIE: Storage and Retrieval for Image and Video Database*, VI3312, pp.150–161, 1997.
- [18] P. Wu and B.S. Manjunath, “Adaptive nearest neighbor search for relevance feedback in large image datasets,” *Proc. IEEE International Conference on Image Processing (ICIP 2001)*, pp.87–98, 2001.

- [19] J. Tešić and B.S. Manjunath, "Nearest neighbor search for relevance feedback," Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol.2, pp.643–648, 2003.
- [20] Y. Rui, T.S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," Proc. IEEE International Conference on Image Processing, pp.815–818, 1997.
- [21] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: A power tool in interactive content-based image retrieval," IEEE Trans. Circuits Syst. Video Technol., vol.8, no.5, pp.644–655, 1998.
- [22] K. Tieu and P. Viola, "Boosting image retrieval," Int. J. Comput. Vis., vol.56, no.1-2, pp.17–36, 2004.
- [23] G. Ohashi and Y. Shimodaira, "Query-by-sketch image retrieval using relevance feedback," Journal of Systemics, Cybernetics and Informatics, vol.1, no.5, pp.25–28, 2003.
- [24] T. Hisamori and G. Ohashi, "Query-by-sketch interactive image retrieval using rough sets," Proc. ISIC.IEEE International Conference on Systems, Man and Cybernetics, pp.1223–1229, 2007.
- [25] W.S. Cooper, "Expected search length: A single measure of retrieval effectiveness based on the weak ordering action of retrieval systems," American Documentation, vol.19, no.1, pp.30–41, 1968.
- [26] S. Nakajima, S. Kinoshita, and K. Tanaka, "Image database retrieval based on relevance feedback by difference amplification," IEICE Trans. Inf. & Syst. (Japanese Edition), vol.J87-D-I, no.2, pp.164–174, Feb. 2004.
- [27] M.D. Dunlop, "Time, relevance and interaction modeling for information retrieval," SIGIR 1997 Proc., pp.206–213, 1997.
- [28] T. Agata, M. Nozue, N. Hattori, and S. Ueda, "A measure for evaluating search engines on the world wide web," Library and Information Science, no.37, pp.1–11, 1997.

Gosuke Ohashi received his B.E., M.E. and D.E. degrees from Keio University in Yokohama, Japan in 1992, 1994 and 1997, respectively. He had been an Assistant Professor since 1997 and he is currently an Associate Professor in Electrical and Electronic Engineering at Shizuoka University. He had been a visiting researcher at University of California, Santa Barbara from 2003 to 2004. His research interests include image processing and image retrieval. He is a member of IEEE, IEEJ and ITE.

Takashi Hisamori received his B.E. and M.E. degrees from Shizuoka University in Hamamatsu, Japan in 2007 and 2009, respectively. He is now with Increment P Corp. He is a member of IEEE.

Toru Arikawa received his B.E. degree from Shizuoka University in Hamamatsu, Japan in 2007. He is currently with the Department of Engineering of the Graduate School of Engineering at Shizuoka University. His research interests include image processing and image retrieval.