RETRIEVAL OF HAND-DRAWN SKETCHES WITH PARTIAL MATCHING

Wing Ho Leung and Tsuhan Chen

Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Email: {wingho, tsuhan}@andrew.cmu.edu

ABSTRACT

Sketching is a natural way of input that provides an effective means of illustration. A sketch consists of multiple strokes that can be captured by pen-based devices to be stored in a database for future retrieval. In our prior work, we proposed a sketch retrieval method that is suitable for whole matching. However, the problem gets complicated if the retrieval is performed for partially matched sketches. In this paper, we propose a method to solve the partial matching problem for retrieving sketches. In the matching stage, the score is first computed by considering the stroke feature similarity alone. For a sketch consisted of multiple strokes, spatial relations between every pair of strokes (bistroke) can be used together with the bistroke features for finding the correspondence to further improve the retrieval performance. A potential application of sketch retrieval for partial matching is to retrieve materials from the lecture notes for reviewing without the trouble of describing it using keywords.

1. INTRODUCTION

Sketching is a natural way of expressing ideas that may sometimes be difficult to be described in words. Children are able to draw sketches before they learn how to read and write. Pen-based devices allow people to capture a sketch in electronic format. A sketch consists of multiple strokes that are sequences of the 2-D x and y coordinates of the points sampled by the pen-based device. It will be useful if we are able to store the sketches into a database and allow the user to retrieve them at a later time by drawing a simple sketch query. Query by sketch falls into the category of content-based image retrieval (CBIR). QBIC [1] was the first CBIR system and it also supports query by sketch. Global features such as area, circularity, eccentricity, etc., are used in shape matching. Matusiak et al. [2] proposed another approach to sketch-based images database retrieval by using Curvature Scale Space (CSS) to match contours. In Sciascio and Mongiello's system [3], the Fourier descriptors are used for shape comparison and they use relevance feedback to improve the retrieval performance for content-based image retrieval over the web. All the above systems assume that the query consists of a single shape.

Lopresti et al. [4][5] reported their work on matching hand-drawn pictures which they call "pictograms". This approach has a drawback that it treats the same hand-drawings with different stroke orders as a poor match. In order to make the system less sensitive to the stroke order, Lopresti and Tomkins [6][7] proposed to match the strings block by block. However, poor match may still result if a stroke is drawn in reverse direction (i.e., when the start point and the end point of a stroke interchange). Under these approaches, string matching is performed for the alignment based on the time sequence. They may work well for hand-writings or pen gestures [8] when the strokes have certain sequence pattern but may not be suitable for unstructured free-form hand-drawings.

The Query by Visual Example (QVE) reported by Kato et al. [9] used correlation of the corresponding blocks between the edge maps for evaluating similarity. Due to the variations in drawing style, this correlation approach will hardly match two rough sketches. Del Bimbo and Pala [10] proposed to use elastic matching to retrieve images from the database based on the user sketch. This energy minimization technique may be too time consuming when it requires many iterations for the solution to converge.

In our prior work [11], we proposed a retrieval method for hand-drawn sketches. It is based on string matching by the alignment of the spatial order among the boundaries of the minimum bounding rectangles of the strokes in each of the x and y projections. In [12], we include the similarity in spatial relations between strokes in the computation of the overall similarity score. These approaches are good for *whole* matching such as trademark retrieval with query by sketch [13], i.e., when the query sketch and the relevant sketch in the database have roughly the same number of strokes as shown in Figure 1. The algorithm is also robust to distortion when there are missing strokes. However, in the case of *partial* matching as shown in Figure 2, the query sketch may be only a portion of the sketch in the database. In this case, our previous proposed approach will not perform well because when there are much more strokes in a sketch from the database than the query sketch so it is more likely to find a set of spatially aligned strokes. This may lead to a high similarity score even for irrelevant sketches, thus reducing the retrieval precision. Besides, because the stroke correspondence is based on the spatial alignment according to the stroke order, there would be a discontinuous step in the similarity score if the spatial order between two strokes is exchanged due to drawing variation.

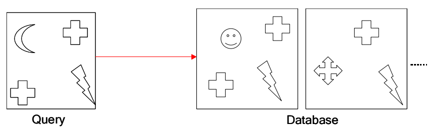

Figure 1 Retrieval with Whole Matching

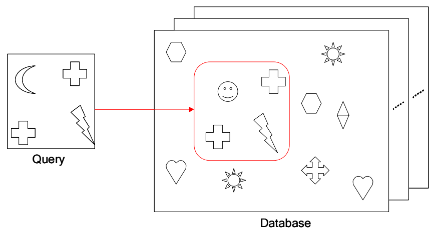

Figure 2 Retrieval with Partial Matching

In this paper, we propose a novel retrieval method for partial matching of hand-drawn sketches. The matching is performed component-wise where a component can be either a stroke or can consist of two strokes (bistroke). The correspondence between stroke feature points depend solely on stroke feature similarity. On the other hand, the correspondence between bistroke feature points depend on both the stroke similarity and the spatial relation similarity simultaneously. The spatial relation similarity is a continuous function over the difference of the relative spatial distance of a pair of strokes between two sketches to avoid the discontinuity in the similarity score when the spatial order between two strokes is flipped.

Sketch retrieval with partial matching is useful to find relevant information after jotting and storing notes with a pen-based device. For example, in a classroom, the teacher may write and draw the lecture notes on the whiteboard that can be captured and stored page by page. Later students can retrieve relevant pages from the lecture sketch database by drawing a simple query. For example, a student can draw the chemical structure of benzene as the query and then the system will retrieve the page that contains a similar chemical structure with the associated description about its name, chemical formula and properties. With the partial retrieval capability, it is not necessary to perform segmentation of a page into sketches before the matching, i.e., we do not need to know which regions of the page form a unit and then decide which regions to be matched with the query.

This paper is organized as follows. In Section 2 we provide the system description of our approach. In Section 3 we describe our experiment and presents the results. The conclusions and future work are in Section 4.

2. SYSTEM DESCRIPTION

Figure 3 shows the system diagram of our approach:

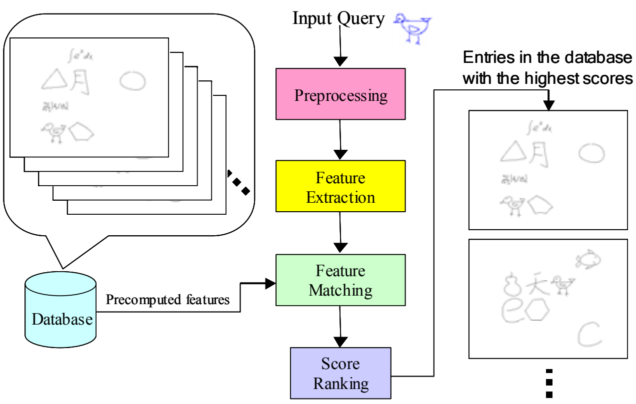

Figure 3 System diagram of our approach

When the user sketches a query, the 2-D x and y coordinates of each stroke of the sketch are captured. Firstly preprocessing is applied to reduce the effect of the variation in drawing speed and drawing style. Then geometric features are extracted for each stroke and the shape likelihood of each stroke is computed. Each entry in the database is described by the term *page* because it is composed by smaller sketches and a portion of a page may match with the query sketch. In the matching stage, the features of the query are compared with the precomputed features from each entry in the database to determine the similarity scores. These scores are sorted and the system will retrieve the entries in the database that have the highest similarity scores.

2.1 Preprocessing

The original 2-D x and y coordinates of each sketch are first resampled to 256 points. If the input query consists of multiple strokes, the number of resample points are distributed according to the proportion of the arc-length of each stroke. Neighboring points of the same stroke will have equal distance after the resampling. This resampling process reduces inconsistencies due to different writing speed.

A single stroke may sometimes be drawn as smaller broken strokes. In order to account for these different styles, stroke merging is performed to merge broken strokes based on the proximity of the end points. A stroke is first checked whether its two end points are close. If they are not close, then that stroke is considered as open stroke and it may be merged with another open stroke if the end point of one stroke is close to the end point of the other.

2.2 Feature Extraction

Geometric features are extracted from each stroke. Different features are used to determine the likelihood that each stroke falls in each basic shape: line, circle and polygon. Some examples of the features are shown in Figure 4. A likelihood value between 0 and 1 is assigned for each stroke with respect to each shape.

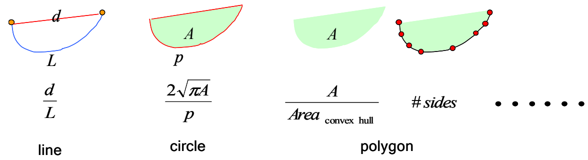

Figure 4 Example features

2.3 Feature Matching

The feature matching module is used for computing the similarity between a query sketch with multiple strokes and a page of sketches also consisting of multiple strokes. Even though the query sketch may appear in only part of a page, our algorithm does not require segmentation of the page into sketches because the matching is performed component-wise. A component may be a stroke when stroke feature matching is performed or a component may consist of two strokes within a sketch when bistroke feature matching is performed. These two matching strategies are described in the following sections.

2.3.1 Stroke Feature Matching

Every stroke in a query sketch or in a page from the database can be represented as a point in the stroke feature space. The query sketch and a page from the database are thus two sets of points in the stroke feature space. The goal of the stroke feature matching is to find a one-to-one correspondence between the set of points of the query and the set of points of the page in the stroke feature space. To accomplish this goal, we first construct the similarity matrix between the strokes of the query and strokes of the page. For each element in the matrix, we compute the similarity score between the associated strokes by considering the shape likelihood of both strokes and the similarity in the extracted geometric features between them. Assume that there are N strokes in the query sketch and M strokes in the page, the similarity matrix will contain $N{\times}M$ elements. We search for the element in the similarity matrix with the highest score and then its row index and its column index will indicate the corresponding stroke feature points between the query and the page. The similarity matrix is updated by removing the row and the column containing that element. This process is repeated until the similarity matrix is empty. The complexity of the stroke feature matching is thus $O(NM(\log N + \log M))$. After finding the correspondence between the stroke features of the query and the stroke features of the page, the overall similarity score between the query and the page is computed by summing up the similarity scores of the corresponding stroke features.

2.3.2 Bistroke Feature Matching

A bistroke feature consists of stroke features of a pair of strokes together with the spatial relation between them. The spatial relation is represented by the distance between the centroids of the minimum bounding rectangle of the pair of strokes, normalized by the sum of stroke length to make it scale invariant. As a result, the query with *N* strokes will correspond to a set of $\binom{N}{2}$ points in the bistroke feature space. Analogous to the stroke feature matching, the goal of the bistroke feature matching is to find a correspondence between the set of points of the query and the set of points of the page in the bistroke feature space. However, if we use the same approach as the stroke feature matching for finding the correspondence, then the complexity of the bistroke feature matching will become

$$O\left(\binom{N}{2}\binom{M}{2}\left(\log\binom{N}{2} + \log\binom{M}{2}\right)\right) = O\left(N^2M^2(\log N + \log M)\right)$$

which is too computation intensive. As a result, instead of finding a one-to-one correspondence between the sets of bistroke feature points, we will simply find the nearest bistroke feature point in the page for each of the bistroke feature point in the query. The complexity of the bistroke feature matching is thus reduced to $O\left(\binom{N}{2}\binom{M}{2}\right) = O(N^2M^2)$. Although it is now possible for multiple bistroke feature points from the query to correspond to a single bistroke feature point from the page, it is unlikely for this to happen too often since there are relatively large number of bistroke feature points located in a high dimensional space. As a result, the requirement of the one-to-one correspondence for bistroke feature matching may not be as strict as the case for stroke feature matching.

The similarity score for the bistroke feature matching is computed as $SIM(B_Q, B_P) = SIM(s_{Q1}, s_{P1})SIM(s_{Q2}, s_{P2})SIM(R_{Q12}, R_{P12})$

where $B_Q = \{s_{Q1}, s_{Q2}, R_{Q12}\}$ and $B_P = \{s_{P1}, s_{P2}, R_{P12}\}$. B_Q is a bistroke feature point in the query and is defined by two stroke features s_{Q1} and s_{Q2}, and the spatial relation between these two strokes R_{Q12}. B_P, s_{P1}, s_{P2} and R_{P12} are the corresponding terms for the page. It should be noted that $SIM(s_{Q1}, s_{P1})$ and $SIM(s_{Q2}, s_{P2})$ can be looked up from the similarity matrix constructed in the stroke feature matching module. The spatial relation similarity $SIM(R_{Q12}, R_{P12})$ is an exponential function of the difference between the relative spatial relations, i.e., the difference of the relative distances between the centroids of the minimum bounding rectangles of each of the two strokes in the query and each of the two strokes in the page. After finding the correspondence between the bistroke features of the query and the bistroke features of the page, the overall matching similarity score between the query and the page is computed by summing up the similarity scores of the corresponding bistroke features.

3. EXPERIMENTS AND RESULTS

For the data collection, a collaboration tool called "mimio" [14] that consists of a capture bar mounted on a physical whiteboard and a wireless pen is used to capture the 2-D coordinates of the strokes electronically. The sketches are drawn by four people and are collected at two different time sessions that are 8 months apart. During each session, we ask each of the four people to draw 5 repetitions for each of the 35 classes of sketches to account for the different drawing styles. Figure 5 shows a few examples of different classes of sketches. Figure 6 shows example sketches drawn by the same person at the two sessions. There is significant variation in the drawings even though the same user draws them.

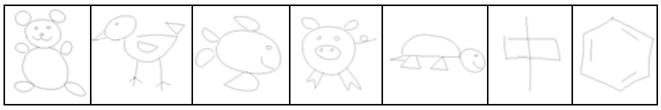

Figure 5 Some examples of different classes of sketches

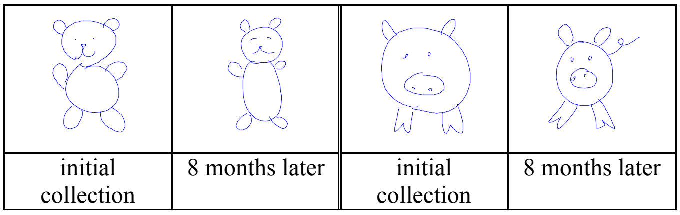

Figure 6 Example sketches drawn by the same user at two different time instants (8 months apart)

For the experiment, we use those sketches from the second collection that consists of at least two strokes as the queries. The objective of this experiment is to evaluate the retrieval performance of our proposed algorithm for partial matching. Each element in the database is constructed by combining sketches from different classes to form a page. In each page, we randomly select seven sketches belonging to different classes from the initial collection and put them together by translation and scaling to form a page. It should be noted that the query sketches and the pages of sketches in the database are collected at a different time in order to simulate an actual retrieval scenario in which the database is collected first and then retrieval is performed at a later time. In this database, there are 100 pages in total. For each of the 35 classes of sketch, there are 20 pages in the database that contain a sketch from that class. Figure 7 shows two example pages in the database that contain a sketch from the class "fish". For each query, we retrieve the elements from the database in the descending order of similarity scores.

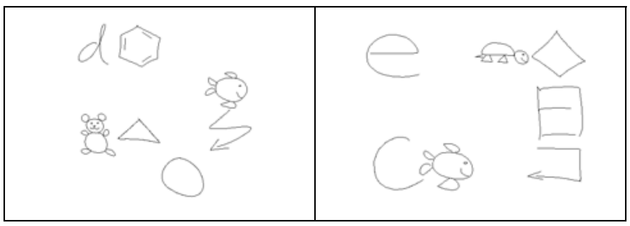

Figure 7 Two example pages in the database

To evaluate the retrieval performance, the precision and recall graph [15] is plotted based on the ranks of those relevant pages in the database that contain a sketch from the same class as the query sketch. The resulting graph shown in Figure 8 is obtained by averaging over all the queries. In a recall and precision graph, the higher the curve, the better the retrieval performance since for the same recall value, a higher curve signifies a higher precision value. We compare our retrieval result with several other approaches. From Figure 8, it can be seen that the retrieval performance using the dynamic programming approach is very low. Using our previous approach for whole matching by using spatial relations between strokes, there is not much improvement. With the seven Hu moment invariants [16] as features, the retrieval performance is also very low because these features are more suitable for global matching. By matching the histogram of edge directions, the result is better than the previous approaches. It can be seen that by using stroke feature matching, there is a significant improvement in the retrieval performance. By using the bistroke feature matching, the retrieval performance is further improved compared with the stroke feature matching. Quantitatively, the stroke feature matching shows an improvement of 62% increase in terms of the average precision compared with the edge histogram matching while the bistroke feature matching shows an additional gain of 56% over the stroke feature matching.

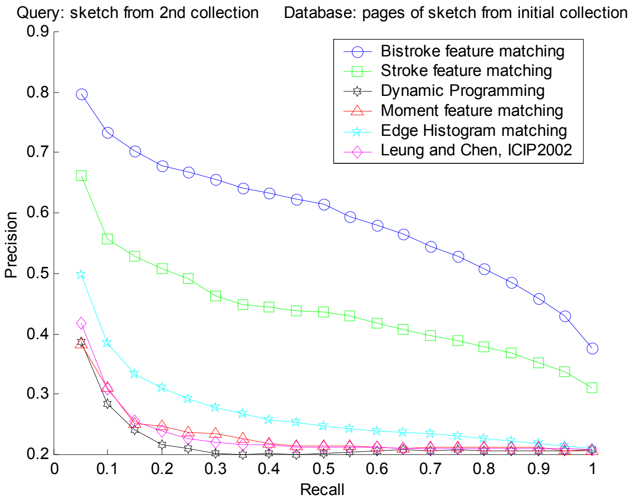

Figure 8 Retrieval Performance for Partial Matching

4. CONCLUSIONS AND FUTURE WORK

We proposed a method for retrieving hand-drawn sketches for partial matching. Experiments show that our approach outperforms some existing methods in terms of the retrieval performance. There is a huge potential for sketch retrieval application with partial matching capability. For instance, a student may draw a simple sketch query and the retrieval system will archive the relevant pages of lecture materials that were captured in electronic format during class. Our approach can match rough sketches drawn by different people. With the partial sketch capability, it is no longer necessary to perform segmentation of a page into sketches before the matching. For future work, we are working on matching the sketches in a hierarchical way and using relevance feedback to refine the retrieval result.

5. REFERENCES

[1] Faloutsos C., Barber R., Flickner M., Niblack W., Petkovic D. and Equitz W. "Efficient and Effective Querying by Image Content". In *Journal of Intelligent Information Systems*, 3:231-262, 1994.

[2] Matusiak S., Daoudi M. and Blu T. "Sketch-Based Images Database Retrieval". In *Proc. of the Fourth Intl. Workshop on Multimedia Information Systems (MIS'98)*, pp. 185-191, Istanbul, Turkey, September 1998.

[3] Sciascio E. D. and Mongiello M. "Query by Sketch and Relevance Feedback for Content-Based Image Retrieval over the Web". In *Journal of Visual Languages and Computing*, 10(6), pp. 565-584, 1999.

[4] Lopresti D. and Tomkins A. "On the searchability of electronic ink". In *Proc. of the Fourth Intl. Workshop on Frontiers in Handwriting Recognition*, pp. 156-165, December 1994.

[5] Aref W., Barbara D., Lopresti D. and Tomkins A. "Ink as a first-class datatype in multimedia databases". In S. Jajodia and V. S. Subrahmanian, editors, *Multimedia Databases*. Springer-Verlag, New York, 1995.

[6] Lopresti D. and Tomkins A. "Temporal domain matching of hand-drawn pictorial queries". In *Proc. of the Seventh Conf. of The Intl. Graphonomics Society*, pp. 98-99. August 1995.

[7] Lopresti D., Tomkins A. and Zhou J. "Algorithms for matching hand-drawn sketches". In *Proc. of the Fifth Intl. Workshop on Frontiers in Handwriting Recognition*, pp. 233-238, September 1996.

[8] Long A. C., Landay, J. A., Rowe, L. A. and Michiels J. "Visual similarity of pen gestures". In *CHI 2000, ACM Conf. on Human factors in Computing Systems*, 2(1), pp. 360-367, April 2000.

[9] Kato T., Kurita T., Otsu N. and Hirata K. "A Sketch Retrieval Method for Full Color Image Database". In *Proc. of the 11th Intl. Conf. On Pattern Recognition*, pp. 530-533. The Hgues, The Netherlands, August 1992.

[10] Del Bimbo, A and Pala, P. "Visual image retrieval by elastic matching of user sketches". In *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 2, pp. 121-132, February 1997.

[11] Leung W. H. and Chen T. "User-Independent Retrieval of Free-Form Hand-Drawn Sketches". *IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, Vol. 2, pp. 2029-2032, Orlando, FL, May 2002.

[12] Leung W. H. and Chen T. "Retrieval of Sketches Based on Spatial Relations Between Strokes". *IEEE Intl. Conf. on Image Processing*, Vol. 1, pp. 908-911, Rochester, NY, September 2002.

[13] Leung W. H. and Chen T. "Trademark Retrieval with Contour vs Skeleton Classification". *IEEE Intl. Conf. on Multimedia and Expo.*, Lausanne, Switzerland, August 2002.

[14] Virtual Ink Corporation, mimio. [Online]. Available: http://www.mimio.com/

[15] van Rijsbergen C. J. *Information Retrieval,* Butterworths, London, 1979.

[16] Hu M. "Visual Pattern Recognition by Moment Invariants". *IRE Trans. on Information Theory*, vol. IT-8, 1962.