RETRIEVAL OF SKETCHES BASED ON SPATIAL RELATION BETWEEN STROKES

Wing Ho Leung and Tsuhan Chen

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

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

ABSTRACT

In this paper we propose a method to retrieve sketches stored in the form of multiple strokes, by extracting the shape information for each stroke and by considering the spatial relation between the strokes. The spatial relation between the strokes can be used to determine the stroke correspondence between the query sketch and the database sketch. Moreover, the similarity between the relationship among query strokes and database strokes can be used as part of the matching score. Sketch retrieval is useful in various applications, such as retrieval of free-form hand-drawings, hand-written or printed text, and trademarks. For free-form hand-drawings, such as lecture notes captured digitally from a whiteboard, a user can easily search through a database of hand-drawings by inputting a sketch about what he/she is looking for, without the trouble of describing it using keywords. For hand-written or printed text, such as a protest sign in a video scene, sketch retrieval can be used to recognize the text without the need to first identify each character. For trademarks, such as company logos, sketch retrieval is useful in adding automation to the feature extraction and the retrieval processes.

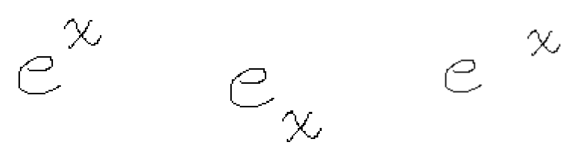
1. INTRODUCTION

Recently pen-based devices such as Personal Digital Assistants (PDA) and electronic whiteboards have become more and more common to the general public. The captured database is stored in the form of sketch which consists of strokes that are sequences of coordinates of the points sampled by the pen-based device. It is necessary to have an efficient sketch retrieval scheme in order to allow users to search for relevant information from the database. Retrieval in the sketch database is equivalent to finding a stroke or multiple strokes from the database that are a good match to the query stroke(s). Sketch retrieval can be applied to various applications such as free-form hand-drawings, hand-written or printed text and trademarks.

Searching through a collection of free-form hand-drawings provides a good motivation for sketch retrieval. For example, in the classroom, the lecture notes written by the teacher on the whiteboard can be captured electronically by a collaboration device such as "mimio" [1]. Later students can retrieve relevant lecture materials from the hand-drawn sketch database by sketching a query drawing. Lopresti et al [2][3] reported their work on matching hand-drawn pictures which they call "pictograms". This approach has the 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 [4][5] 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). In our prior work [6], we proposed a sketch retrieval method for general unstructured free-form hand-drawings.

In addition to free-form hand-drawings, sketch retrieval can also be used in identifying hand-written or printed text. This is different from traditional text recognition, since text recognition deals with a limited set of characters while the sketch domain is a more general set that can be language independent. Besides, the goal of traditional text recognition is to recognize the context of the text by mapping it to a set of pre-defined characters. On the other hand, the goal of sketch retrieval is to find an entry from the database that is similar to the query. Potential applications of sketch retrieval in this area include identifying text written or printed on a placard during a protest, or automatic recognition of road signs on a highway.

Another application of sketch retrieval is trademark matching. In some trademark retrieval system, the trademarks are first annotated with keywords that are then used for retrieval. However, this process requires a lot of manual labor in order to assign keywords and the annotation for the same trademark may not be consistent with different users. By using our sketch retrieval system, the features of the trademarks are extracted automatically therefore manual annotation is not required. As a potential application, a user can simply provide the trademark image or a sketch of the trademark as the query, and then similar trademarks from the database can be retrieved together with the related product or company information.



Sketch 1 Sketch 2 Sketch 3

Figure 1 Example Sketches

If each stroke of a sketch is treated independently, then the spatial relation between the strokes is not considered. It can be seen in Figure 1 that with respect to sketch 1, although sketch 2 and sketch 3 contain the strokes displaced by the same distance (but in different direction), sketch 3 is perceived as more similar to sketch 1 than sketch 2. As a result, it is important to consider spatial relation in determining similarity. In image indexing and retrieval, 2D string and its variants [7][8] have been proposed to represent the spatial relation between labeled objects or regions in an image. In addition, the spatial relation can be quantified by using topological neighborhood graph [9][10][11]. We would like to modify these ideas of using spatial relation and apply them in our sketch retrieval system. While previous work in [2][3][4][5] assume that the stroke matching is based on the drawing order, we will use the spatial relation between the strokes to determine the stroke correspondence between the query sketch and the database sketch and include this factor in calculating the matching score.

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 for free-form hand-drawings. In Section 4 we reports the application of our sketch retrieval system for hand-written text and trademarks. The conclusions and future work are in Section 5.

2. SYSTEM DESCRIPTION

Figure 2 shows the system diagram of our approach:

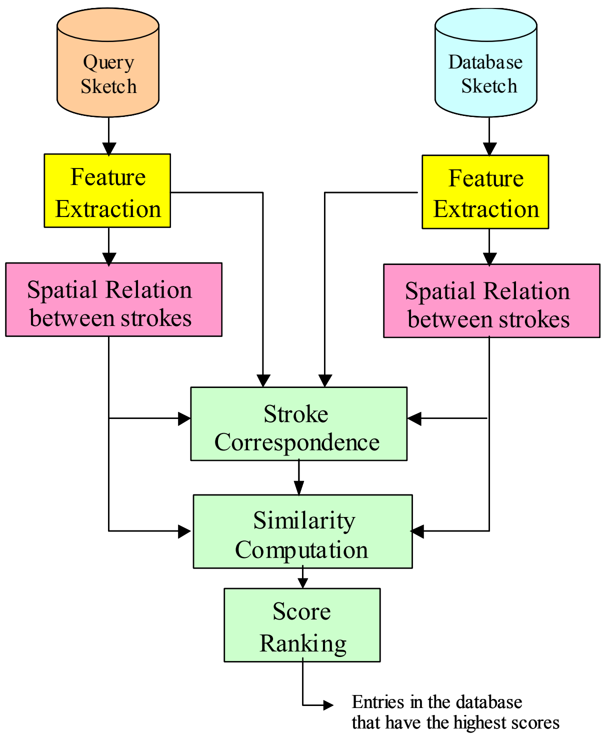


Figure 2 System diagram of our approach

In our previous work [6], the spatial relation between strokes is not used. In this paper, we study the effect when the spatial relation is used for the stroke correspondence alone and when the spatial relation is used for both the stroke correspondence and the similarity score computation.

2.1 Feature Extraction

Features are extracted from each stroke. These features are used for shape estimation, as shown in Figure 3, to determine the likelihood that each stroke falls in each basic shape type: line, circle and polygon. A confidence measure that takes the value between 0 and 1 is assigned for each stroke with respect to each shape type. A more detailed description of the feature extraction module can be found in [6].

Shape Type	Features
Line	- height of triangle formed by the two end points and each sample point - ratio between the sum of distances of each pair of neighboring points and the distance between end points
Circle	- perimeter efficiency - area between the number of points of the convex hull and the number of point of the original stroke samples - angle that the stroke samples go around the center of the stroke
Polygon	- ratio between area formed by stroke samples and area of its convex hull - area between the number of points of the convex hull and the number of point of the original stroke samples - angle that the stroke samples go around the center of the stroke

Figure 3 Shape types and the corresponding features

2.2 Spatial Relation between strokes

When considering the spatial relation between strokes, the *x* and *y* coordinates of the minimum bounding rectangles of the strokes are used. The spatial relation between strokes is determined by the order of these coordinates and this information may be used for both the stroke correspondence and the similarity score computation.

2.3 Stroke Correspondence

In order to determine the similarity between the query sketch and the database sketch, we need to find out the correspondence between the query strokes and the database strokes. The stroke correspondence can be assigned according to the shape information between strokes. In this case, in Figure 4, the triangles in sketch 2 and sketch 3 are corresponding strokes to the triangle in sketch 1 and the rectangles in these two sketches are corresponding strokes to the rectangle in sketch 1. If the similarity is entirely based on shape, then sketch 2 and sketch 3 are both similar to sketch 1. On the other hand, if the spatial relation between the strokes is also considered in determining the stroke correspondence, then we can make a further distinction that sketch 2 is more similar than sketch 3 with respect to sketch 1. The orders of the strokes need to be preserved in determining the correspondence. In Figure 4, in the horizontal direction, the stroke correspondence of the sketches is the same as the previous case when only shape information is considered. However, in the vertical direction, only one of the object pairs (the triangle pair and the rectangle pair) can be corresponding strokes for sketch 2 and sketch 3 with respect to sketch 1 but not both because the two object pairs have conflicting orders in the vertical direction. The object pair to be selected as the corresponding strokes depends on the shape similarity of the object pairs. For example, if the triangle pair has a larger similarity score than the rectangle pair, then the triangles will be the only corresponding strokes in the vertical direction. The horizontal and vertical directions are considered separately for the stroke correspondence and then the similarity score is the combined result in both directions. In [10][11], the spatial relation between multiple objects is used for similarity computation assuming that the object correspondence is given. In our approach, object correspondence is not known in advance, and the spatial relation is used to determine the object correspondence automatically.

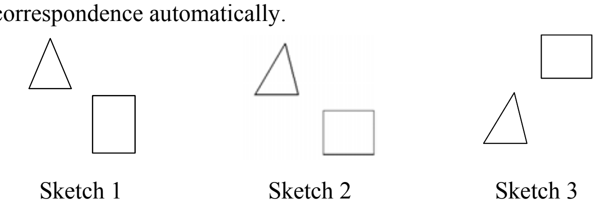


Figure 4 Example Sketches

2.4 Similarity Computation

The matching (similarity) score between the strokes is calculated after the correspondence between the query strokes and the database strokes are known. If the spatial relation is not considered, the matching score between the query sketch and the database sketch can be computed by summing up the matching scores between the corresponding strokes minus an unmatched cost. There are two cases of unmatched strokes: 1) a stroke in the query sketch that does not match any of the strokes in the database sketch; 2) a stroke in the database sketch that does not match any of the strokes in the query sketch. The unmatched cost is proportional to the number of unmatched strokes. On the other hand, if the spatial relation between the strokes is considered, then the matching score computation is modified by including the similarity of spatial relations calculated based on the topological neighborhood graph as described in [10].

3. EXPERIMENTS AND RESULTS

We perform an experiment to analyze the gain in the retrieval performance when the spatial relation between strokes is used. For our database we collect 12 classes of sketches that consist of at least 3 strokes from different people. Initially, for each sketch we have 20 repetitions made by 4 different people to account for the different drawing styles. Eight months later, the same people are asked to redraw the sketches 5 more times per person. Figure 5 shows a few examples of these sketches in the database. Figure 6 shows example sketches drawn by the same person at two different time instants (8 months apart). There is significant variation in the drawings even if they are drawn by the same user. We then compute the matching scores between a query and database sketches using 3 different approaches. Based on the rank of those sketches that fall in the same category, we plot the precision and recall graph [12] in order to analyze the retrieval performance. The result is shown by the two curves in Figure 7 and in Figure 8. In Figure 7, each of the sketches from the initial collection is used as the query while in Figure 8, each of the sketches drawn 8 months later is used as the query, matching with the initial collection of the sketch database. For each query sketch, all the database sketches that belong to the same class, whether they are drawn by the same user or different users, are considered as relevant retrieved items, thus making our system user-independent. In both cases, it can be seen that the retrieval performance is the worst if spatial relation between strokes is not used while the retrieval performance is improved when the spatial relation is used for stroke correspondence and it is further improved when the spatial relation is used for both stroke correspondence and similarity score computation.

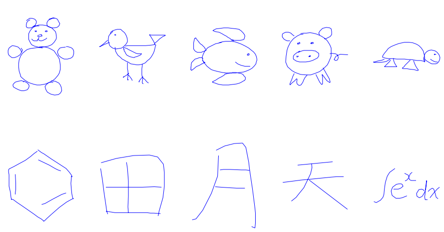


Figure 5 Some examples of the sketches in the database

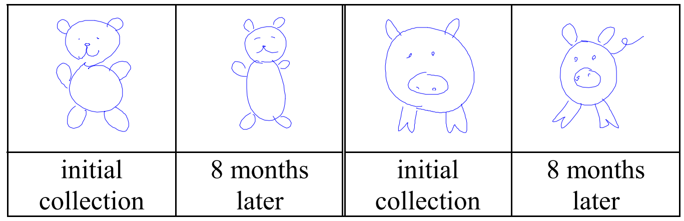


Figure 6 Example sketches drawn by the same user at two different time instants (8 months apart)

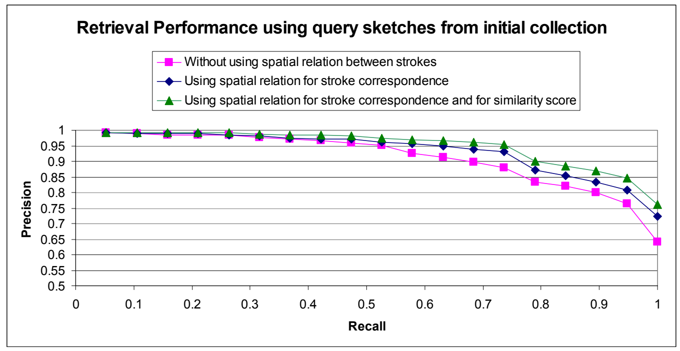


Figure 7 Retrieval Performance

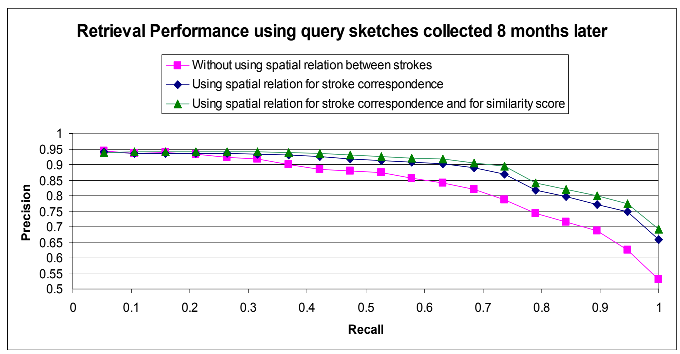
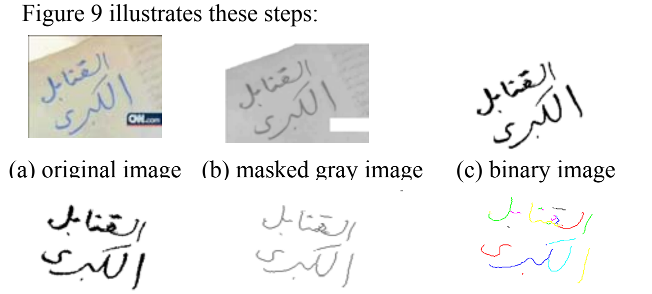


Figure 8 Retrieval Performance

4. APPLICATIONS IN HAND-WRITTEN TEXT AND TRADEMARKS

Our sketch retrieval can also be used in other applications such as hand-written text and trademarks. Preprocessing can be done to extract the sketch from the region of interest of an image. We first mask out the region of interest and then perform thresholding to convert it into a binary image. A morphological filter is applied to the binary image for thinning. The image can be rotated to align the principle axis with the horizontal direction. Afterwards a contour following algorithm is applied to extract the strokes from the resulting image. After the strokes are extracted from the query image, then we use our sketch retrieval system to find similar sketch or its associated image. Figure 9 illustrates these steps:

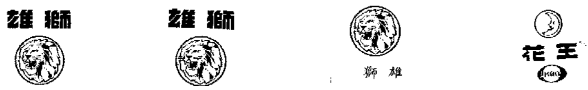


(d) rotated image (e) thinned image (f) extracted strokes

Figure 9 Stroke Extraction from an image

Query Score = 1.0 Score = 0.94 Score = 0.67

Figure 10 Example retrieval results for text



Query Score = 0.98 Score = 0.95 Score = 0.77

Figure 11 Example retrieval results for trademark

Figure 10 and Figure 11 show an example query and some retrieved results for the text database and the trademark database respectively. For the text database, after the stroke extraction from the input image, the database sketches are retrieved according to the similarity scores. The trademark database consists of over 600 trademark images and we retrieve database trademarks that are similar to the query trademark according to the similarity score. The first column in both Figure 10 and Figure 11 contains the retrieved result that has high similarity score while the retrieved result in the third column has low similarity score. The second column in Figure 10 and Figure 11 contains the retrieved result that is similar to the query but with some difference in the spatial relation, and it has intermediate similarity score after using spatial relation in the stroke correspondence and the similarity score computation.

5. CONCLUSIONS AND FUTURE WORK

This paper presents our system for retrieving sketches by using spatial relation between multiple strokes. Experiments show that the retrieval performance is increased when spatial relation is used to find the stroke correspondence and to be considered as a factor in the matching score. Sketch retrieval can be used in various applications such as free-form hand-drawings, hand-written text and trademarks, as shown in this paper. For future work, we would like to explore other ways of combining the stroke matching score and the spatial relation matching score. Moreover, we will perform more experiment for hand-written text and trademarks and explore sketch retrieval in other potential applications.

6. ACKNOWLEDGEMENT

The authors would like to thank Telecom Lab (TL), Taiwan for providing the trademark database.

7. REFERENCES

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

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] Leung W. H. and Chen T. "User-Independent Retrieval of Free-Form Hand-Drawn Sketches". To appear in *ICASSP2002*, May 2002. Available: http://amp.ece.cmu.edu/.

[7] Chang S. K., Shi Q. Y. and Yan C.W. "Iconic indexing by 2-D Strings". In *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. PAMI-9, No. 3, pp. 413-428, May 1987.

[8] Wang Y.-H. "Image indexing and similarity retrieval based on A new Spatial Relation Model". In *Intl. Conf. on Distributed Computing Systems Workshop*, pp. 16-19, April 2001.

[9] Egenhofer M. J. and Al-Taha K. K. "Reasoning About Gradual Changes of Topological Relationships". In *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, Frank A. U., Campari I. And Formentini U., eds., Lecture Notes in Computer Science no. 693, pp. 196-219, Springer-Verlag, 1992.

[10] Nabil M., Ngu A. H. H. and Shepherd J. "Picture Similarity Retrieval Using the 2D Projection Interval Representation". In *IEEE Trans. on Knowledge and Data Engineering*, Vol. 8, No. 4, pp. 533-539, August 1996.

[11] Nabil M., Shepherd J. and Ngu A. H. H. "2D Projection Interval Relationships: A Symbolic Representation of Spatial Relationships". In *Advances in Spatial Databases: Fourth Int'l Symp., SSD '95*, Lecture Notes in Computer Science no. 951, pp. 292-309, Springer-Verlag, 1995.

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