

Sketch Off: A Sketch Recognition Competition

Paul Corey

Sketch Recognition Laboratory,
Department of Computer
Science, Texas A&M University
Mail Stop 3112, College Station,
TX 77839

pfc9416@cse.tamu.edu

Brian Eoff

Sketch Recognition Laboratory,
Department of Computer
Science, Texas A&M University
Mail Stop 3112, College Station,
TX 77839

bde@cse.tamu.edu

Tracy Hammond

Sketch Recognition Laboratory,
Department of Computer
Science, Texas A&M University
Mail Stop 3112, College Station,
TX 77839

hammond@cse.tamu.edu

ABSTRACT

In this paper we describe the competition to be conducted at the Sketch Recognition Workshop of IUI 2009. The Sketch Recognition Competition promotes discussion, innovation, and competition within the sketch recognition community.

Author Keywords

Competition, Sketch Recognition.

ACM Classification Keywords

H.5.2 User Interfaces: *Input devices and strategies*, I.5.0 Pattern Recognition: General

INTRODUCTION

Sketch recognition is a rapidly developing field focused on the automatic recognition and interpretation of freely sketched drawings, generally input using a digitizing pen or tablet. To foment discussion and analysis of various existing sketch recognition techniques, promote the development of new, innovative methods, and instill the sketch recognition community with both camaraderie and a good-natured spirit of competition, we will be hosting the first Sketch Recognition Competition during the Sketch Recognition Workshop at IUI 2009.

Other computer disciplines have held similar competitions with great success. Perhaps one of the best known competitions is TREC, held at the Text REtrieval Conference. Annually since 1992, researchers in the area of text retrieval compete to create the best algorithms for one of several tracks which focus on some subproblem in text retrieval, such as blogs or legal documents. TREC provides data sets and scoring criteria for each track. Since its inception, TREC submissions have led the field of text retrieval research[6]. Like TREC, we offer two tracks for

competition; however, as this will be the first competition for sketch recognition, the competition scope is considerably more limited.

Similarly, AAAI holds the General Game Playing Competition. This competition seeks to move from the traditional, specialized game playing machines, like Deep Blue, to agents that can compete at any well-specified game. This competition promotes the full use of machine learning, knowledge representation, and decision making, and has produced novel techniques applicable to a variety of domains[1].

It is our hope that, like TREC and the General Game Playing Competition, this competition will promote progress in sketch recognition by providing uniform data and testing for a variety of sketch recognition methods and by inspiring competition between sketch researchers. Ideally, by examining interesting, and difficult, sketch domains, innovative sketch recognition will be designed promoting the field as a whole.

OVERVIEW

The competition has been divided into two tracks. Track 1, designed to be a less difficult task, focuses on recognition of single, isolated symbols. It is intended for comparison of existing techniques and to introduce beginners to the field. Track 2 requires recognition and interpretation of full diagrams, consisting of multiple symbols within a single sketch. It is intended to promote experimentation and innovation by including difficult, open research tasks.

TRACK 1 – SINGLE SHAPES

The focus of the first track is the recognition of isolated sketched symbols. Each sketch input consists of a single symbol, possibly drawn using multiple pen strokes. Symbol recognition is a fundamental task of sketch recognition, and numerous techniques have been proposed to solve this problem. This track is targeted towards both experienced sketch recognition researcher, who will be able to apply their own techniques to the competition domain, and novices (to sketch recognition) whose experience in related fields such as machine learning or pattern recognition may allow them to adapt methods from those

fields to a sketch recognition domain.

Each sketch used in Track 1 contains a single symbol, possibly composed of multiple strokes. However, sketchers are not constrained to drawing in a particular style, with a specified number or ordering of strokes, or in a specified orientation. The symbol set for Track 1 consists of the symbols: Sinusoid, Spiral, Helix, Arrow, Rectangle, Ellipse, 0, 1, 2, 3, 4, Bowtie, =, and X. Figure 1 depicts an example of each of the symbol types.

For this track, a simple accuracy metric is used to determine the best technique. The team which correctly classifies the greatest number of symbols wins.

TRACK 2 – FULL DIAGRAMS

In Track 2, competitors must recognize and interpret multiple symbols within a single sketch. This track seeks to transition from simple recognition, or classification of strokes into symbols, to a more complete sketch based application. In addition to recognition of symbols, the diagram must be meaningfully interpreted, requiring recognition of the association and links between symbol, as well as what these links mean.

The diagrams used in Track 2 are sketched finite state machines. Each sketch contains a fully specified state machine, and competitors must recognize the components of the state machine and interpret those components in order to determine whether or not the machine will accept or reject input strings.

This domain is more difficult than Track 1, and incorporates open research problems to some degree. First is the stroke grouping problem. To correctly classify a symbol, each algorithm must also determine which strokes make that symbol. Second, the recognition algorithms must be able to distinguish text ('A', 'B', and 'C') from shapes (states and transitions). Much of the research done in this area has focused on identifying shapes from text in handwritten documents[2, 7]. However, in this domain shapes and text are heavily intermingled, and the number of text and shape stroke are closer to equal. Both of these differences pose serious difficulties for these previous approaches. Third is the annotation problem. The competitors must be able to correctly determine which of the text symbols is associated with which shape. To successfully interpret the finite state machine, the correct transition label must be associated with each transition.

Examples of the finite state machines drawn for this track are shown in Figures 2, 3, and 4. Unlike standard FSMs, these are drawn using triangles as the start state to reduce some of the complexity of the diagram.

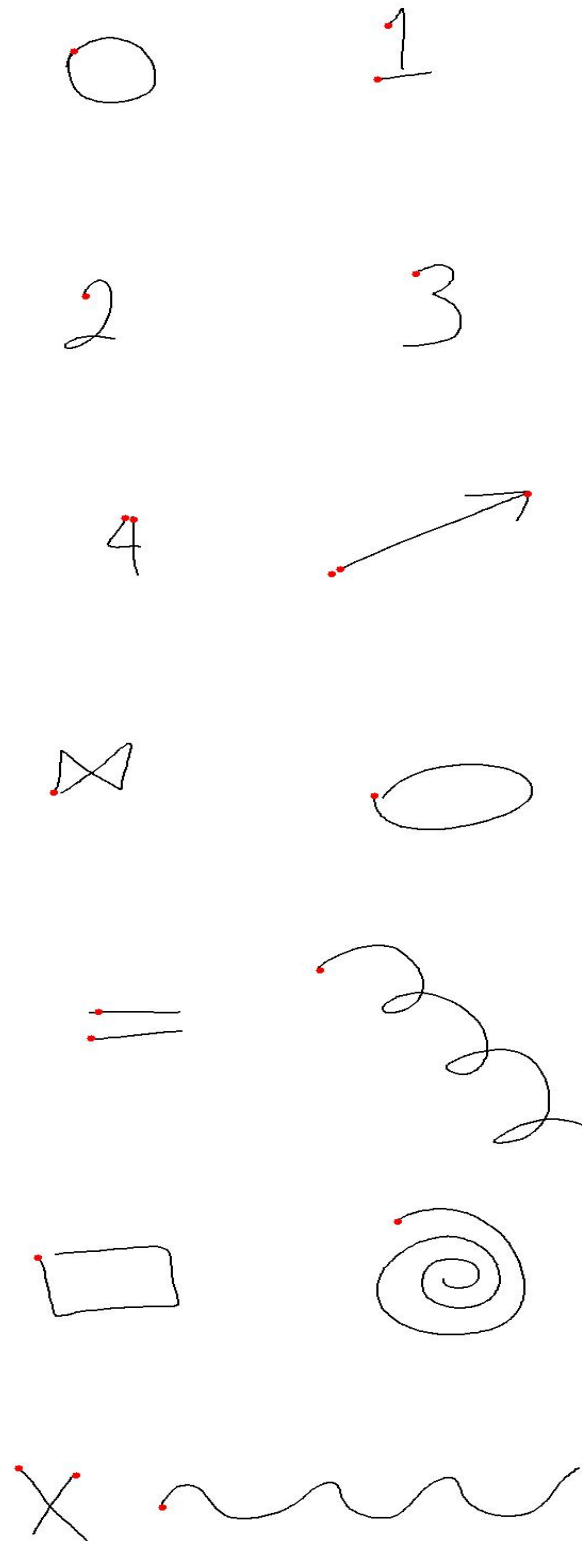


Figure 1: Examples of sketched symbols for Track 1.

The scoring system for Track 2 takes into account four areas of the recognition and interpretation process. Each of the four areas can receive up to 25 points, for a total of 100 possible points for each diagram. The first item is number of symbols recognized. Competitors lose points for recognizing too many or too few symbols. Table 1 details the points awarded for number of symbols recognized. Next, grouping is scored. The strokes of each recognized symbol must match a ground truth grouping. Competitors receive a percent of the possible points for correct groups, equal to the number of correct groups divided by the total number of groups. Third, recognition of symbols is examined. For each correctly labeled symbol, a percentage of the points will be awarded equal to the number of correctly labeled symbols divided by the number of correctly grouped symbols. Lastly, interpretation of the diagram is scored. With each diagram, the competitors are given a set of input strings for the FSM. A percentage of the points equal to the percent of the input strings correctly determined to accept or reject will be awarded.

| Number of Symbols | Points Awarded |
|------------------------|----------------|
| Correct | 25 |
| +/- 1 Symbol | 20 |
| +/- 2 Symbols | 15 |
| +/- 3 Symbols | 10 |
| +/- 4 Symbols | 5 |
| +/- 5 Symbols, or more | 0 |

Table 1: Point awarded for number of symbols recognized.

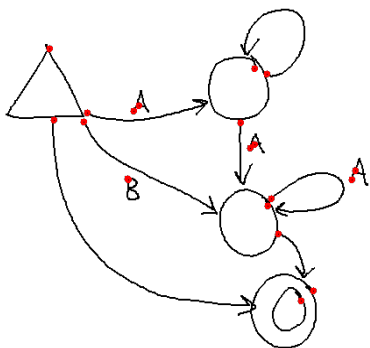


Figure 2: A sketched finite state machine diagram for Track 2.

Example Recognizers

For comparison purposes, we will provide several recognizers implemented from algorithms that have become benchmarks within the sketch community.

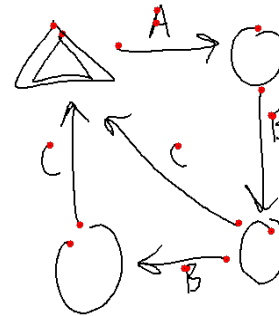


Figure 3: Another sketched finite state machine diagram for Track 2.

For Track 1, four recognition systems will be implemented. The first is LADDER, a geometric recognition system. It breaks strokes into geometric primitives, such as lines and arc, and builds symbols from these primitives based on geometric and spatial constraints[4]. The second system is based on Rubine's GRANDMA, which recognizes single stroke gestures, using linear classification based on several features[5]. A commonly used method of appending strokes will be used when presented with multiple stroke symbols. Third, a vision based approach using Hausdorff distances will be implemented, first converting strokes into an image[3]. Lastly, Wobbrok's \$1 Recognizer will be implemented[8]. This is a single stroke template matching

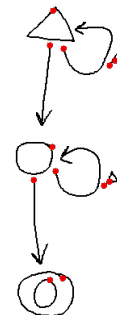


Figure 4: A third sketched finite state machine from Track 2.

algorithm, and multiple strokes will be handled by appending strokes similar to the procedure for the Rubine recognizer.

For Track 2, the same recognizers will be used. However, additional grouping algorithms will be implemented to segment strokes into individual symbols. The first grouping algorithm will be a naive grouper based on the overlap of stroke bounding boxes. The second grouping algorithm will group strokes based on time between ending one stroke and beginning the next. A simple threshold will be used to determine when to segment strokes. Due to time constraints, these algorithms will be available on the competition website shortly after the conference.

These methods provide a baseline for comparison during the competition. Competitors will be able to judge their techniques not only against each other, but also against these standards.

CONCLUSION

Hopefully, this competition will inspire the sketch recognition community to create new, innovative techniques to solve the problems posed by the competition tracks, as well as foster a sense of community and open a forum for discussion of sketch recognition techniques. In time, this competition may provide direction, motivation, and research achievements similar to TREC and AAAI's General Game Playing Competition. To learn more about the competition or find out the winners of the competition, please visit the Sketch Recognition Workshop webpage, <http://srl.csd.tamu.edu/workshops/2009/iui/>. Additionally, all test and training data will be posted on the website after the competition, as well as the scripts used to score the entries. Good luck to all the competitors!

ACKNOWLEDGMENTS

This research is supported by the NSF IIS Creative IT Grant #0757557 Pilot: Let Your Notes Come Alive: The SkRUI Classroom Sketchbook.

REFERENCES

1. AAAI-08 General Game Playing Competition. <http://www.aaai.org/Conferences/AAAI/2008/aaai08generalgame.php>.
2. Bishop, C.M.; Svensen, M.; Hinton, G.E., "Distinguishing text from graphics in on-line handwritten ink," *Frontiers in Handwriting Recognition, 2004. IWFHR-9 2004. Ninth International Workshop on*, vol., no., pp. 142-147, 26-29 Oct. 2004.
3. Kara LB, Stahovich TF. An image-based, trainable symbol recognizer for hand-drawn sketches. *Computers & Graphics* 2005; 29(4):501-17.
4. Hammond, H. Randall Davis, "LADDER, a sketching language for user interface developers," *Computers & Graphics, Volume 29, Issue 4, August 2005*, Pages 518-532
5. Rubine, D. 1991. Specifying gestures by example. In *Proceedings of the 18th Annual Conference on Computer Graphics and interactive Techniques SIGGRAPH '91*. ACM, New York, NY, 329-337.
6. TREC. Text REtrieval Conference. <http://trec.nist.gov>.
7. Shilman, M.; Viola, P. "Spatial Recognition and Grouping of Text and Graphics," *Sketch-Based Interfaces and Modeling, 2004*, Eurographics Association, 91-95.
8. Wobbrock, J. O., Wilson, A. D., and Li, Y. 2007. Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes. In *Proceedings of the 20th Annual ACM Symposium on User interface Software and Technology* (Newport, Rhode Island, USA, October 07 - 10, 2007). UIST '07. ACM, New York, NY, 159-168.