Building a Training and Test Corpus of Hand Drawn Sketches

Michael Oltmans and Randall Davis

What: We have constructed a database of hand drawn sketches and the tools necessary to analyze this database. The data set represents both complete sketches and the individual strokes annotated with hand-assigned labels (line, arc, ellipse, polyline, polygon, or spiral).

The analysis tools help us browse and replay the sketches, optionally synchronized with a QuickTime video of the user, and analyze a number of features of the strokes themselves, for example, stroke speed. Using these tools we have begun to explore the differences in sketching style between different sketching tasks and between different users.

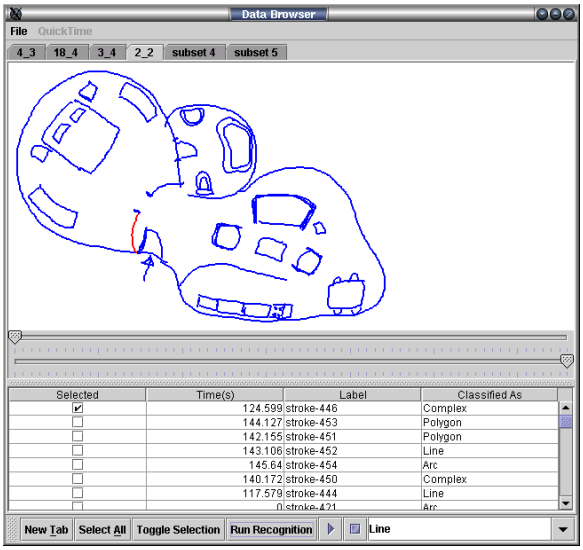

Figure 1: The sketch media player

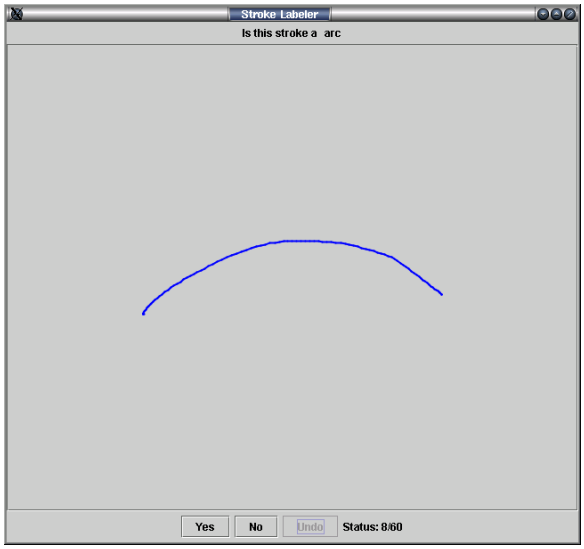

Figure 2: Stroke labeling interface

Why: The recognition and understanding of hand drawn sketches is a relatively new and emerging field. As a consequence of its age our field lacks standard data sets for testing and training sketch recognizers. Our goal is to create such a data set to be shared among the growing sketch-understanding community.

By studying the different sketching styles used by different people, or by the same person performing different tasks, we hope to gain a better understanding of which properties of hand-drawn sketches are general, and which are highly tuned to specific people and tasks, making it possible to improve our recognizers.

How: We collected the data by asking 23 subjects (to date) to draw between 4 and 7 rough sketches of floor plans for apartments. Subjects were asked to sketch their current apartments, brainstorm several apartments they would like to live in, draw a more formal version of their favorite sketch, and to draw redraw their favorite apartment one final time while verbally describing the design to the moderator. All of the sketches were collected on a TabletPC running Windows XP Tablet PC Edition. Subjects were allowed to erase using the eraser end of the stylus. The interface erased whole strokes and did not allow the subject to erase a portion of a stroke. The final task of describing the design was video taped.

We constructed a Postgres database to store the strokes, including timestamps for each point, which enables us to "replay" the sketch.. The next task was to hand-label the strokes with their class: line, arc, polyline, polygon, ellipse, and spiral. As some strokes are ambiguous (see Figure 3) and do not fit cleanly into just one of these categories, each stroke is labeled with a binary vector, with one element for each class. . A simple interface (shown in Figure 2) presents each stroke to a human labeler (members of our research group) and asks if the stroke was in one of the above categories. To accommodate possible

ambiguity, the human labeler is presented with each stroke six times, once for each class. Each stroke is labeled by at least two different people so we can identify and resolve conflicts between labelers.

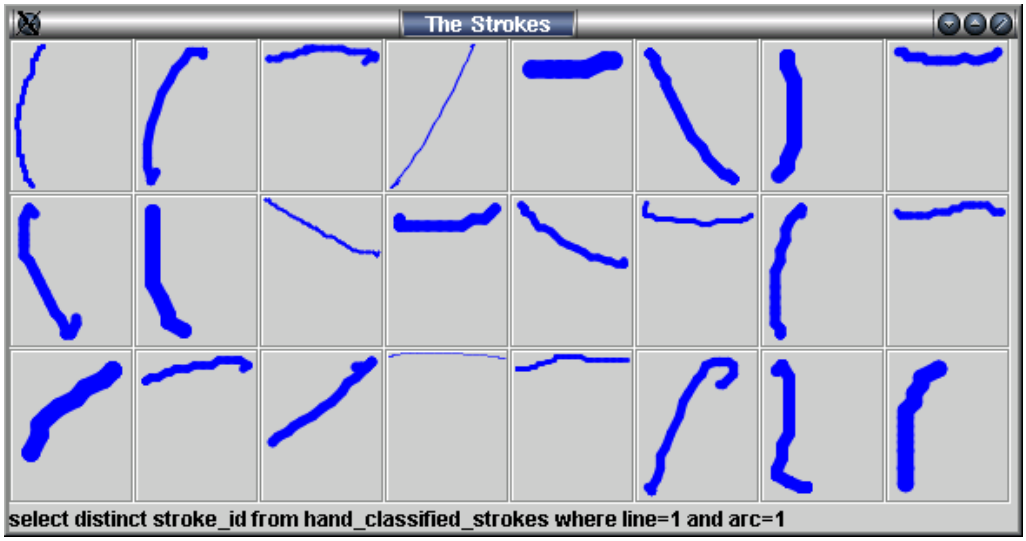

Figure 3: It is unclear without more context whether these strokes are lines or arcs

We have written a number of tools for browsing the collection of sketches, which allow us analyze the sketching process rather than just the final sketch. The most powerful tool is a sketch media player (shown in Figure 1. The player shows a grayed-out version of a sketch and uses the timestamps on the strokes to draw them one at a time in the proper sequence and at the proper time. The player also allows for the selection of groups of strokes by clicking on them, selecting them with a bounding box, or by specifying simple criteria such as timestamps or the label assigned by our current recognition algorithm (e.g. select all lines).

Aaron Adler extended the player so that a QuickTime video, recorded while the user was drawing, can be synchronized with the sketching sequence. This allows us to watch both the subject and the drawing surface without the difficulties of videotaping a laptop screen that may not be visible due to glare or obstruction by the subjects as they are drawing.

Progress: We have to date collected a total of 127 sketches from 23 subjects with a combined total of over 11,000 strokes. We have thus far labeled 279 strokes and will continue the labeling process until we have an evenly distributed subset of entire dataset.

We have also begun preliminary explorations into constructing user profiles and the differences in styles between tasks. Our preliminary analysis suggests that there are some differences in features such as average stroke speed between these various classes. It remains to be determined how strong these effects are and how than can be used to improve our low level stroke recognizers.

Future: Moving forward we will be formalizing the analysis of the variations in features across the users and tasks and attempting to incorporate that knowledge into our recognition architecture. We also plan to get the database into a publishable format for distribution to researchers at other institutions.

We also need to collect more data. We would like to collect sketches from other domains and perhaps using different media, for example, Mimio boards. We are also considering running more subjects in the experiment using different erasing semantics (i.e. deleting partial strokes instead of whole strokes) to analyze the effects this has on sketching style.

We believe the dataset will provide an important basis for tackling the recognition of complicated collections of strokes, such as over-tracing, line continuations, and structures like dashed lines.

Research Support: This work is supported in part by the MIT/Microsoft iCampus initiative and in part by MIT's Project Oxygen.