# Early Processing in Sketch Understanding

Tevfik Metin Sezgin & Randall Davis

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

**The Problem:** Freehand sketching is both a natural and crucial part of design, yet is almost totally unsupported by current design automation software. We are combining the flexibility and ease of use of paper and pencil with the processing power of a computer to produce a design environment that feels as natural as paper, yet is far smarter. One of the most basic steps in accomplishing this goal is converting the original pen strokes into the intended geometric objects, the early processing part of the task.

**Motivation:** Current CAD systems provide extensive support for producing precise and detailed shop drawings of devices and are of considerable utility. But in the early stages of a design process, when the intent is to brainstorm and test out different ideas, a less precise model, or perhaps a rough sketch, is what most designers need, hence the extensive use of pencil and paper. But while this feels natural and easy to use, it has the obvious drawback that the result is at best an image, rather than a good model of the device. Hence our desire to provide both the flexibility and naturalness of paper and the advantages of computational models.

**Previous Work:** Mark Gross has worked on recognition and interpretation of diagrams in design [1]. His Electronic Coctail Napkin architecture provides a computational environment for working with design diagrams. In their joint work, Mark Gross and Ellen Do stressed the importance of employing a natural input scheme during design [2].

Earlier work in our group [3, 4] produced a basic sketch recognition system for the mechanical devices domain. This system recognizes some basic geometric objects and domain specific shapes, but focused primarily on recognizing mechanical components rather than careful processing of the basic input data. The present system is far more robust, in part because it views the sketch as a drawing being created while the computer watches, and hence can make use of both traditional geometric information (e.g., direction change) and information about pen velocity.

**Approach:** We use a digitizing tablet to capture pen's position in time, which permits examining pen speed, acceleration and direction of movement. Using these properties as distinguishing features we extract potential vertices of each stroke using two different recognition schemes that assign a certainty to the vertices. Each recognition method suggests a candidate fit approximating the input stroke. These fits are fed into a blackboard system that generates more candidate fits using the two input fits and returns the best fit.

[origin=c]0[scale=.5]figures/4segs.eps

Figure 1: A stroke that contains both cuves and straight lines (shown in black). The output generated by the recognizer system consists of Bézier curves and straight lines (shown in red, green, and blue).

[origin=c]0[scale=.6]figures/engine3$_{b}efore.eps[origin=c]0[scale=.6]figures/engine3_{a}fter.eps$

Figure 2: A sample snapshot of a rough sketch for an engine (left), and the output generated by the recognizer (right).

**Difficulty:** One of the major problems is the noise in the input. Even strokes that appear straight contain significant amount of noise, a consequence of freehand drawing. A variety of well-known filtering algorithms fail to produce any significant improvement because the number of data points available is generally small compared to what is needed. A second, related difficulty is that although commercially available digitizing tablets can support sampling rates above 100 Hz, most existing applications do not require such high sampling rates, so the actual report rate remains around 40-50 Hz.

**Impact:** Our system gives the user the freedom and convenience of using a digitizing tablet like a paper, and puts the processing power of a computer on the user's service. It frees the user from the trouble of selecting among different tools for drawing different shapes. In addition it saves time and effort by eliminating the need to export sketches on paper to electronic format.

**Future Work:** Our current system is to some extent tuned to the input capture device in use, and while changes to the constants can be done manually, a self-configuring system would be far better. In addition, alternate methods of filtering the noise in the data may improve the performance.

Beyond these short-term improvements, we aim to integrate this low level recognition system with higher level sketch understanding and interpretation modules. Our sytem serves as the link between the raw input data and their low level geometric representations. The next step involves mapping these geometric objects to actual domain specific entities. This will certainly enable more useful analysis and aid the early design process.

**Research Support:** This work is supported in part by the Ford/MIT Collaboration under the Virtual Engineering project, in part by the Lockheed Martin Space Operations Company under CSOC Contract NAS9-98100, and in party by the MIT Oxygen Collaboration.

**References:**

[1] M.D. Gross. Recognizing and Interpreting Diagrams in Design. *2nd Annual International Conference on Image Processing*, pp. 308-311, 1995

[2] M.D. Gross and E. Do. Ambiguous Intentions: a Paper-like Interface for Creative Design. *Ninth Annual Symposium for User Interface Software and Technology*, pp. 183-192. Marc Brown and Ramana Rao (eds.), Seattle, 1996.

[3] M. D. Muzumdar ICEMENDR: Intelligent Capture Environment for Mechanical Engineering Design. Master's Thesis, MIT, 1999

[4] A Foundation for Intelligent Multimodal Drawing and Sketching Programs. Master's Thesis, MIT, 1999