Vasu Agrawal

14 November, 2014

15-112 Term Project Proposal

Problem Statement:

There is an increasing focus on being able to control technology without physically interacting with it. To this end, my project will be program which enables the user to interact with his or her computer using gestures drawn in the air. To make this technology as accessible as possible, I will rely solely on the user's webcam, instead of expensive additional hardware.

Steps to solution:

1. Obtain a camera feed.
2. In controlled lighting, find a contour of the user's hand.
3. Detect the position of the center of the hand (palm).
4. Find the positions of the user's fingers.
5. Track these hand positions over multiple frames.
6. Determine which of a few basic gestures the user's hand is making, if any.
7. Initiate basic actions with recognized gestures.
8. Allow user to train new gestures.
9. Optimize and improve recognition with more complicated algorithms.
10. Provide a hookable API.
11. Interface program with an open source game to demonstrate capabilities.

Modules:

I intend to use OpenCV 2 and its Python wrapper. OpenCV provides both computer vision and machine learning capabilities. Depending on the complexity necessary, I may alternatively implement my own basic machine learning, or use another machine learning library.

Gestures:

I plan on being able to differentiate between gestures using their orientation and shape. This means that a counterclockwise circle is different from a clockwise circle, which are both different from a line. Further classification can be done based on the overall size of the gesture as well as the speed at which it was drawn. Basic template gestures will be lines, circles, arcs, and infinity curves.

Proposed Algorithm:

In order to be able to classify gestures, I will first preload certain basic templates, such as circles, lines, and other various arcs. These gestures will be stored as a list of points along the curve which approximate the gesture well. Since the tracked gestures will also be a list of points, a distance based mapping can be done on the tracked gesture points to see which of the template points they correspond to. From this correspondence, an error can be derived. The gesture with the lowest error will then be considered valid, provided that the error is under some threshold (to ensure that a gesture was actually drawn). Further smoothing can be done on the tracked points to account for inaccuracies and imperfections in the tracking algorithm.