Camera Based 3D Input Method

Peter Pong Yen-Wen Liu
Supervisor: Professor Anind
Carnegie Mellon University, Pittsburgh, Pennsylvania

Introduction

Motivated by a frustration with the limitations of conventional electronic input devices (such as the ubiquitous mouse and keyboard), we propose a input method, that allow users to input and control data by simply performing certain predefined gestures in front of their computer screens. A webcam captures the hand motion and image data are interpreted to enable users to manipulate data directly with their hands. This input method will extend the workspace environment to three dimensions, and will allow for intuitive and immediate data manipulation.

In particular we would like to investigate the feasibility of a input method that would deliver the standard multi-touch capability but at the same time allow user to interact in three dimension without the need for extra hardware.

The Setup

Logitech Quickcam Pro 9000
o Premium autofocus
o 2.0 megapixel performance
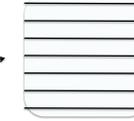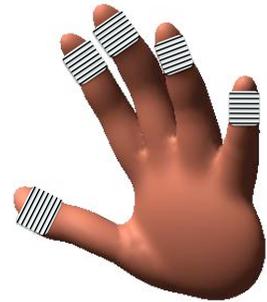
Finger Wraps
o These labels wraps around the finger so they can be captured by the camera regardless of the hand's orientation.
o The horizontal lines are evenly spaced so that the we can estimate the distance between the fingers and the camera.
o We can possibly also encode different information on each label to differentiate between the fingers

References:

[1] Dietz P., Yerazunis W., Leigh D. Very Low-Cost Sensing and Communication Using Bidirectional LEDs.

[2] Nock R., Nielsen F. Statistical Region Merging.

[3] Han J. Y. Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection.

[4] Hudson S. E. Using Light Emitting Diode Arrays as Touch-Sensitive Input and Output Devices

Theory of Operation

The camera captures the image and it performs statistical segmentation on the objects in the image and picks out where the labels are located in the image. It then tracks the labels from the background and processes the labels to distinguish which finger the label is on and to extract 3D positional data
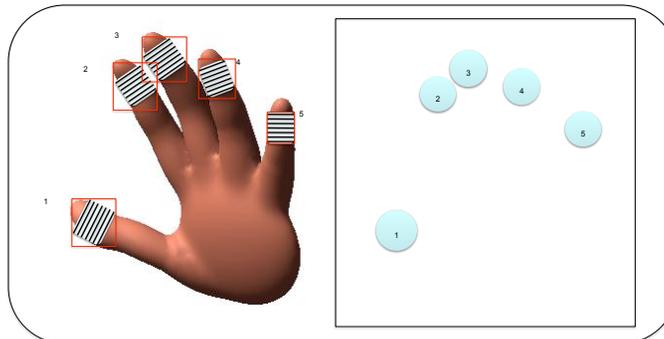

Current Progress

We are currently able to capture at about 10-15 fps using the Logitech webcam and have implemented segmentation and color tracking algorithms in our program to tack the user's finger. Our program is capable of tracking multiple fingers at the same time. However we are still working on other aspects of the project such as inferring the distance of the user's finger and recognizing user's gesture. The screen shots on the right are some of the demo apps we have created.

Demo Programs
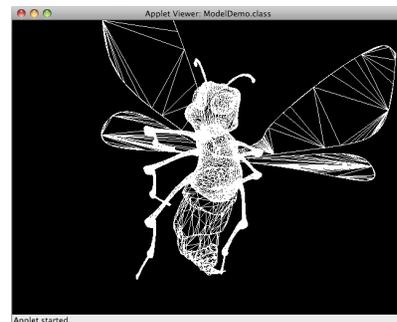
Demo 1. This is a simple demo program, in which users are able to manipulate the 3D model by waving their hands in front of the camera. Currently users are able to translate and rotate the model. However in the future we would like to include features that would allow users to modify the 3D meshes, such as stretching the mesh or adding/removing edges from it.
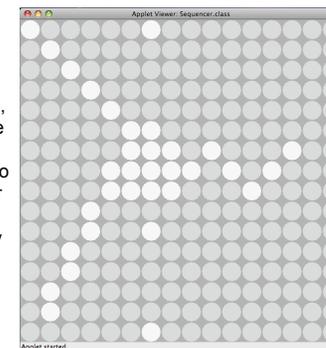
Demo 2. This is an interactive sequencer. It works like the Yamaha Tenori-on. Each dot on the 16x16 grid is a button that can be turn on/ off when user "presses" on it by gesturing in front of the camera. There is an imaginary line that continuously scans across the grid from left to right and will play a note for each of the buttons that are turned on in the column the line is currently on.