Specific goal of the project (Give a problem definition. Why is this an interesting and difficult problem?)

The video version of this “Theremin Problem” is an interesting problem because there is a loss of information when using just .

A real Theremin can use oscillators to track the user’s hands in all three dimensions. However, a simple camera does not have any information on depth. This was one of the many complications we had to find a solution for. Additionally, a virtual Theremin has to account for possible inaccuracies in hand position detection, hand movement detection, changes in lighting, and changes in the background. Furthermore, after gaining all visual data, a virtual Theremin needs to use that using that to predictably(I don’t get this adverb…) and reliably reflect a change in audio output.

* Methods (based on material from class, the literature, or on your own ideas)

Before we could output any sound, we first needed to be able to segment the hands. To do this, we decided to use hand segmentation using skin detection which is a method similar to the one used in the hand gesture homework earlier this semester. To make processing of the image more manageable, the image was split in half down the center into two separate images.

Volume:

The first half of the image was used to control the volume of the audio. The segment was then iterated through pixel by pixel. For each pixel, the RGB values were taken and thresholding was performed on their values based for skin detection (based for?). This resulted in a grayscale image, and the contours of the grayscale image were obtained using the opencv findContours function.

A function was then written to iterate through all contours in order to find the one with the largest contour. The largest contour was then taken to be the left hand. We also decided that the centroid of the contour would be the best data to be used in controlling the volume. In order to find the centroid we…

Different amplitude equations

-finally chose to use decibel

Centroid Moments

Optical Flow

-gray scale image

-skin detect

-finally skni detect and contours (best result yet)

Matrix Norm

-l2 norm of flow vectors

* Experiments

A bunch of optical flow experiments

Audio experiments

* Results
  + Volume can pan but the transition could be worked on
  + Can sense hand motion for vibrato, the audio could be worked on
* Discussion of results (Is the method successful? Are your results satisfactory? What are the limitations of the method used? Did you improve on the state-of-the-art? Give a critical evaluation.)
* Conclusions.

github: