

## CV Studio

With recent advances in computer vision, there is an opportunity to enhance one's musical experience by creating a system that translates gestures and movements into various instruments, allowing a user to essentially act or dance to create music.

### Description:

CV Studio will be a way for users to create music loops and lay down ideas using gesture commands. For example, users will be able to stomp their feet for a kick drum, clap their hands to play a clap sound, or hit an "air drum" to play drum sounds. These gestures will be supplemented with a glove that contains pressure pads so that a user could combine a button with a gesture to create more combinations of commands. These sounds will be able to be played in a loop so that the user can build on what they have already created.

### Scope:

This will be a software and hardware project. On the software side, we will be using computer vision libraries to process video data so that we can process gestures. On the hardware side, we will attach and wire pressure sensors to a glove so that it can be used as a source of input to change sounds and add more commands. Our project is not to build a new digital audio workstation, but rather to route gesture commands to a pre-existing one.

### Requirements:

- We want to have the latency of a sound playing when a user triggers it to not be noticeable. (100 ms)
- We want to have accurate detection of gestures and reduce false positives so that the application plays what the user intends it to. (Accuracy of 90%)
- We want the system to be able to play fast songs. (8 sounds per second)
- We want the system to be able to play complex songs. (4 simultaneous gestures)
- We want the system to be persistent. (save and load settings)

The preliminary implementation of this project is as follows. The Xbox Kinect camera module will transmit video data to a laptop, which will be running our computer vision software designed to recognize gestures. This software will output MIDI information corresponding to the recognized movements, which will be routed to a digital audio workstation which will then perform the proper action on existing music software (Maschine 2). These actions will include both MIDI for playing notes as well as basic keystrokes such as redo, undo, save and load.

For the glove system, we will route pressure pads on each fingertip to a Raspberry Pi, to process the inputs and send the results to the same laptop, to combine into more actions.

Our MVP will only consist of the Kinect camera wired through a laptop running our CV software. This product will be able to recognize a minimum of five gestures with an accuracy rate of 90% across multiple user tests. A user will be able to stand in front of the Kinect camera and

Mark Prettyman, Chris Ng, Tony Lu

generate a multitude of sounds using the gestures of their choice. The time delay between when a gesture is done and the sound is played will be less than 100ms.