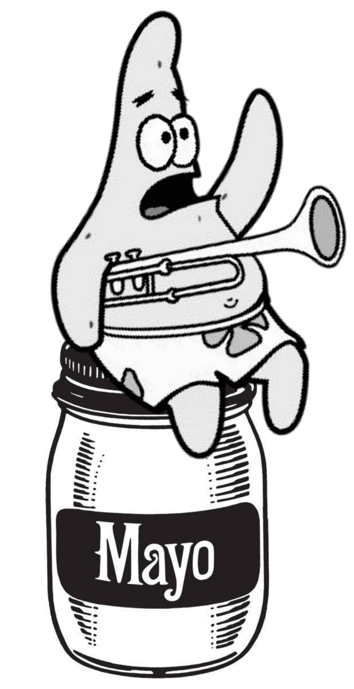
Is Mayonnaise an Instrument?

Team D1: Harry Fernandez, Min Gun Kim, Tomas Vancura
18-500 Capstone Design, Spring 2022
Electrical and Computer Engineering Department
Carnegie Mellon University



Product Pitch

Our product gives beginner musicians a tactile way to learn the concepts behind sound synthesis and serves as a helpful addition to a seasoned producer's workflow.

Synthesizing sounds electronically is often slow and unintuitive. Our project allows electronic musicians to use objects in their environment as MIDI (Musical Instrument Digital Interface) instruments through the help of a computer vision headset and motion-sensing glove. This provides users with a more intuitive way to interact with electronic synthesizers and makes learning the underlying concepts of sound synthesis easier.

System Architecture

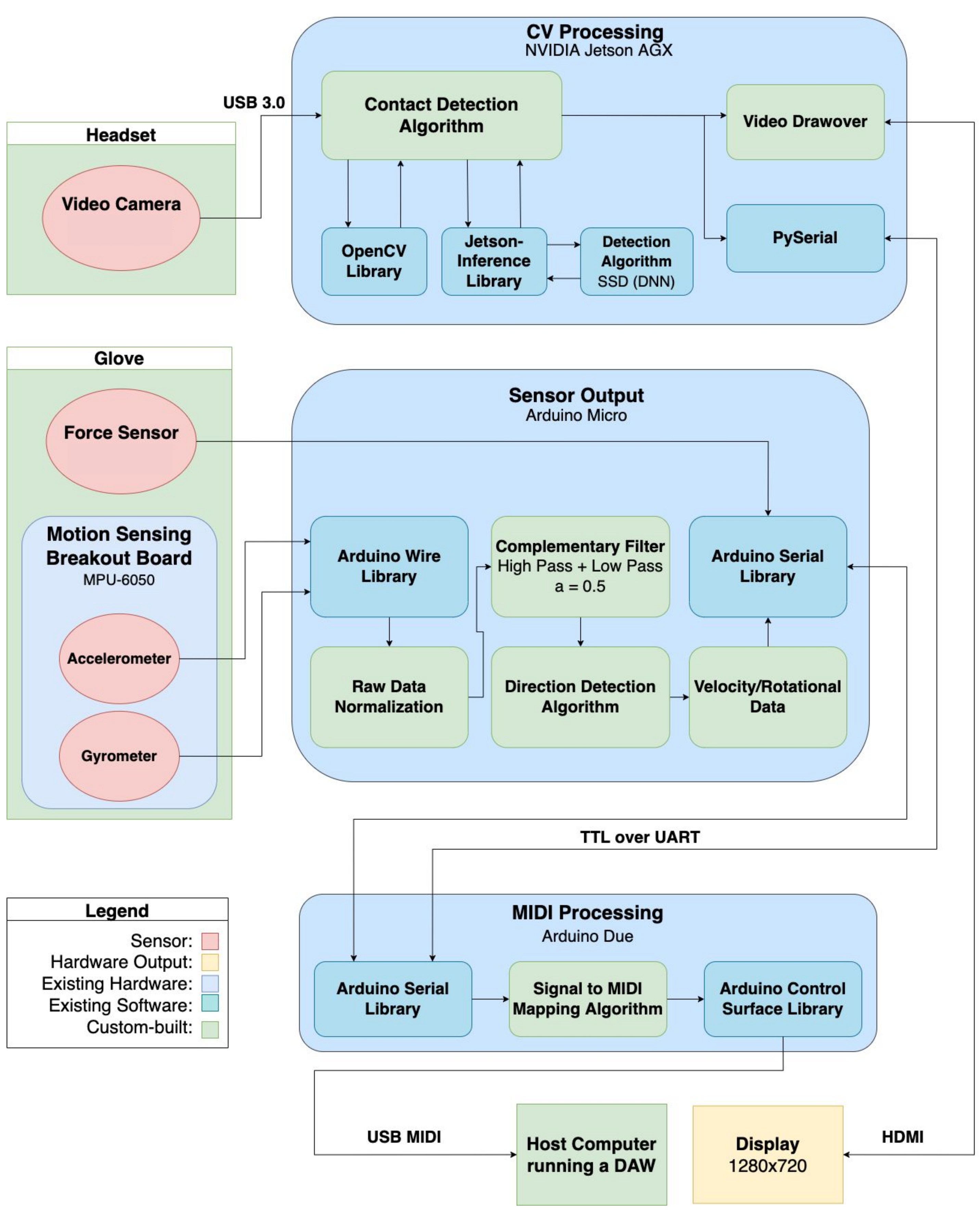


Figure 1. Block diagram detailing the system

Three main subsystems

- *Computer Vision:* Responsible for identifying objects and predicting the object touched
- *Sensor Output:* Responsible for measuring rotation and observing object contacts
- *MIDI Processing:* Responsible for aggregating/mapping data and outputting it to MIDI

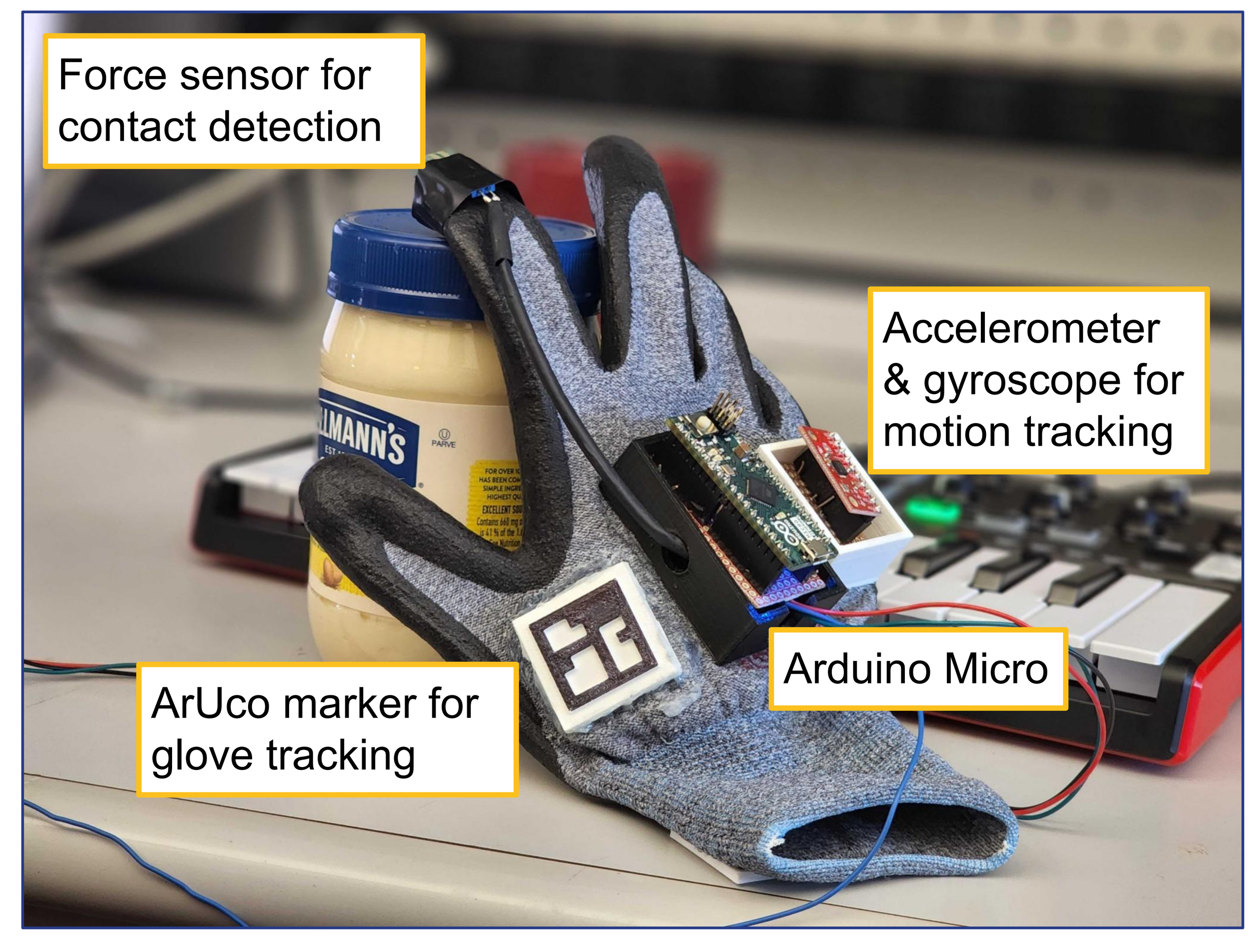
Conclusions & Additional Information

View our status reports and project documents here!

http://course.ece.cmu.edu/~ece500/projects/s22-teamd1/

Overall, our system exceeded our expectations, and user tests revealed that the responsiveness of our project makes it a helpful tool for musicians to create and experiment with electronic music. While we are very satisfied with the state of the project, given more time we would love to improve the robustness of the system, increase the number of supported objects, and make the system even easier to use with ergonomic improvements.

System Description



The *Computer Vision Subsystem* detects what types of objects are in the scene via Deep Learning. The user can choose to track an object by overlaying the glove and a detected object very briefly, which is then remembered by the system and sent to the *MIDI Processing Unit*.

When the force sensor on the *Motion-Sensing Glove* detects contact, the accelerometer data is processed and sent to the *MIDI Processing Unit*.

Figure 2. The motion sensing glove and its components

All of this data is then processed and sent to a host computer as MIDI messages, allowing for different objects to generate different sounds. The generated sound can be further modified by rotating the held object in space.

Connecting the device to a computer running any standard software synthesizer allows the user to map objects and motions to parameters like filter cutoff and pitch bend. In this way, our device can be interfaced with just like *any other normal MIDI instrument.*

Complementing other, existing MIDI instruments is where our product shines. When paired with other MIDI instruments, a musician can augment their sounds in real time as they play music. This allows for vast experimental creation.

Figure 3. A Music Technology student playing our instrument during user validation testing

System Evaluation

Accuracy tests

- Computer vision was tested over 14 trials with 3 objects
- **Object detection** accuracy measures the consistency of detecting the correct object
- **Prediction accuracy** measures the robustness of the currently held object being identified
- We nearly met our **goal of 90%** contact prediction accuracy for each object (maximum error of 5%)

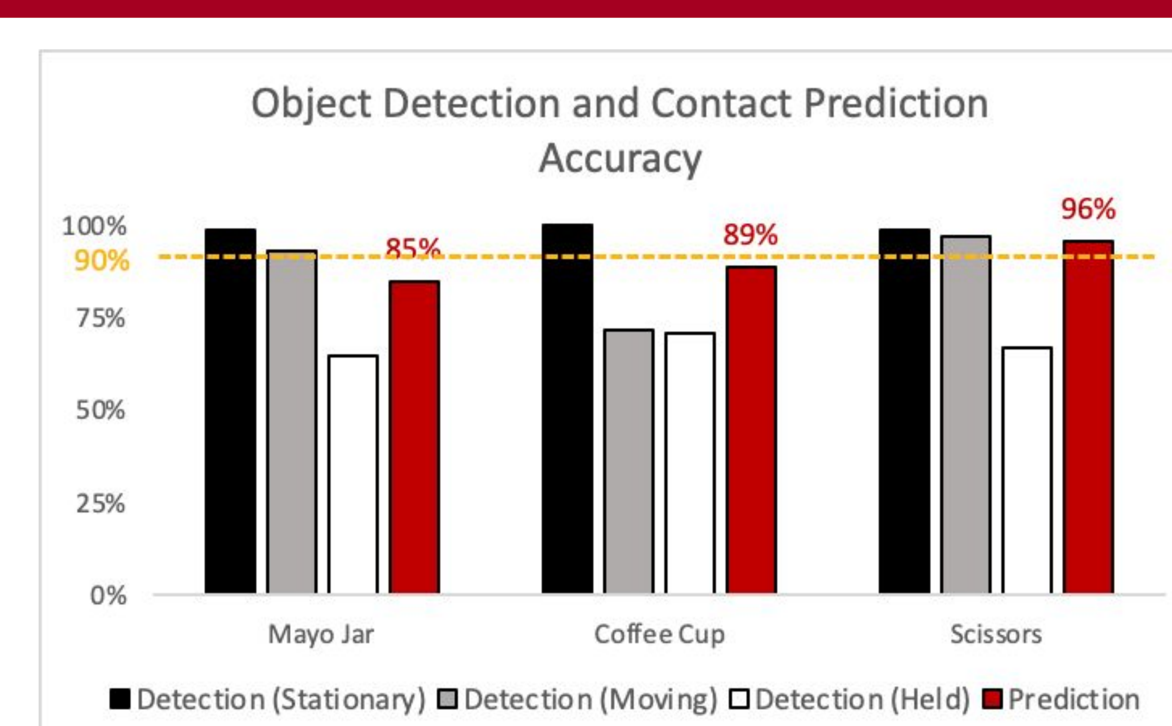


Figure 4. Computer vision accuracy tests

Table 1. Latency performance of the subsystems

Subsystem	Goal	Actual
Due <-> Micro	1.5ms	~0.8ms
Due <-> Jetson	1.5ms	~1ms
Due <-> Max	1.5ms	<1ms
End-to-End	30ms	0.8ms - 2.5ms

Latency tests

- Tested pairwise between components as well as end-to-end
- Each performance metric is the average transaction period over 10,000 simulations

User responses

"Super cool! I'm really looking forward to the next steps in the development of this project!"

"That was cool! That was fun! That was really fun!"

Electrical & Computer ENGINEERING

Carnegie Mellon