# Team D1: Is Mayonnaise an Instrument?

# Harry Fernandez, Min Gun Kim, Tomas Vancura



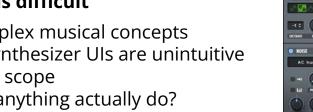
#### **Use Case:**

#### Getting started is difficult

- Lots of complex musical concepts
- DAW and synthesizer UIs are unintuitive and huge in scope
- What does anything actually do?

#### **Experimentation is clunky**

- Not a lot of room for play and creativity • unless you really know what you're doing
- Almost all synthesizers use knobs, buttons, and faders
- Difficult to set up a quick workflow



???



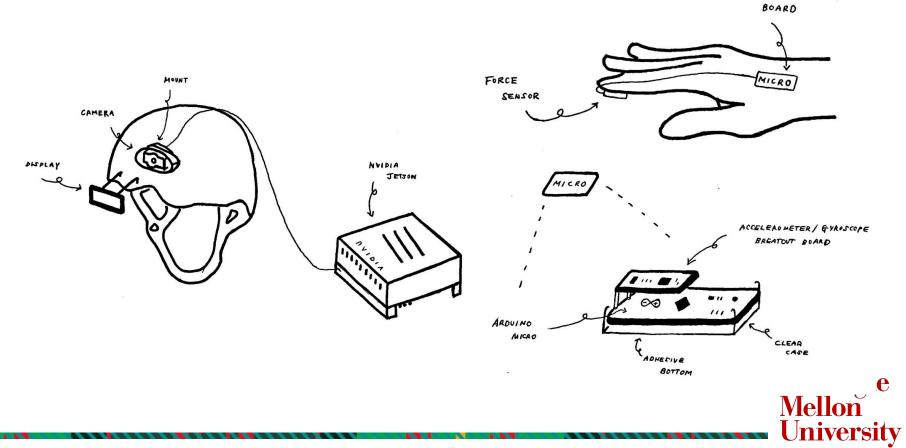
???

Carnegie Mellon University

???

### Solution

JENJOR PROCESSING



### **Use-Case Requirements**

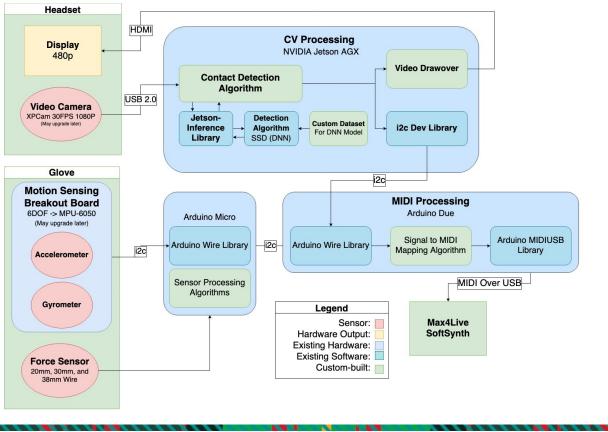
1.

Requirement	Metric
Capture video	60 FPS (min. 30 FPS)
Identify objects in video	≥ 3 distinct objects, ≤ 1m range at 90% accuracy
Determine when user is touching object	≤ 30 <sup>1</sup> ms end-to-end latency, ~ 99% accuracy
Determine position of held object	≥ 80% Movement detection accuracy
Translate/output positional, object data to MIDI	≤ 30 ms end-to-end latency
Visualize to a display	≤ 30 ms latency from picking up object to display

Xiaoyuan Gu, M. Dick, Z. Kurtisi, U. Noyer and L. Wolf, "Network-centric music performance: practice and experiments," in IEEE Communications Magazine, vol. 43, no. 6, pp. 86-93, June 2005, doi: 10.1109/MCOM.2005.1452835.

Carnegie Mellon

### System

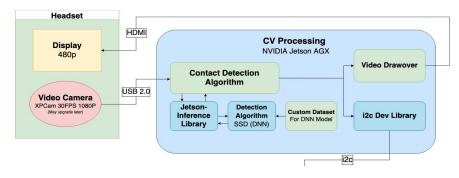


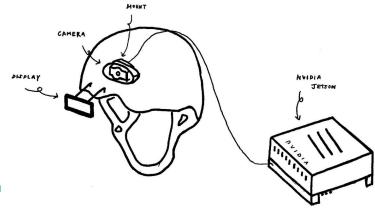
5

Carnegie Mellon

### **Headset & Computer Vision**

- Object detection and contact identification
  - What objects are in the camera's view?
  - Which object is the user (potentially) holding?
- Single Shot Detector DNN
  - Detect position, size, type of recognized objects in scene
  - Research suggests SSD can hit our performance metrics on Jetson boards
- Jetson-Inference
  - DNN AI image library from NVIDIA
- Unit Test: Three distinct benchmark objects
  - Measure success of accuracy across trials
  - Measure latency in detection





### **Object Detection Deep Learning Models**

- Multiple options for the SSD model
  - Pre-trained model (via Jetson-Inference Library, trained on COCO dataset)
    - Pros: Convenient

Time

Less

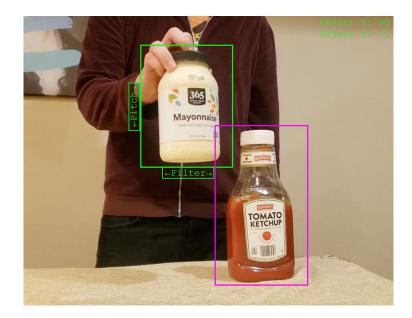
More

- Cons: 91 object classes, many are irrelevant
- Train new model (on Open Images data set)
  - Pros: Comprehensive training data (600 classes, 1.9M images, 16M bounding boxes)
  - Cons: Household object classes are very general
- Train new model (on custom dataset)
  - Pros: Control over specific types of classes
  - Cons: Collecting and labeling hundreds of images



## **Display and GUI**

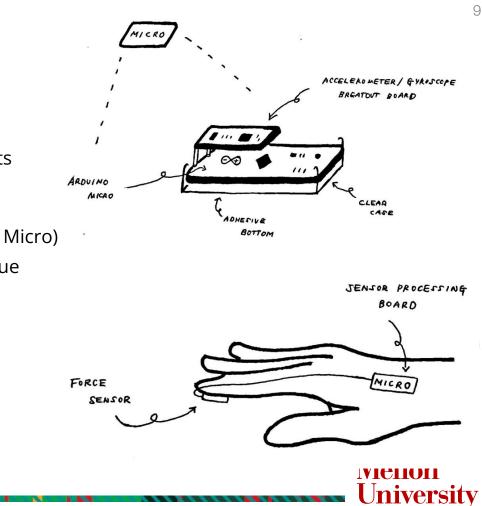
- Display will act as an AR HUD
- Visualizes the system's response to user actions
- Serves as a tutorial/entry point
  - Want to minimize learning curve
  - Will determine ease of use through user testing



Carnegie Mellon

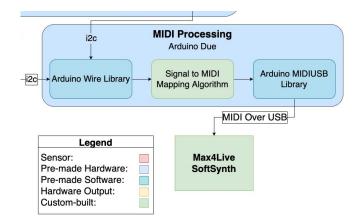
### Glove

- Force Sensor
  - Lies flat, detects contact with objects
- Accelerometer and Gyrometer
  - Track movements
- Sensor Data Aggregation Board (Arduino Micro)
  - Packages data and sends it off to Due
- Unit Test
  - Force sensor sensitivity
  - Motion simulation
  - Ergonomics/usability



### **Communication Protocol & MIDI Mapping**

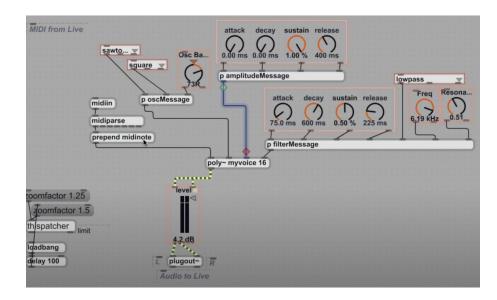
- Board-to-board communication over I2C (400kbps)
  - Jetson sends contact-detection packets
  - Arduino Micro sends sensor packets
- MIDI Mapping
  - MIDIUSB Library
  - Four main parameters for MVP
    - Object Type
    - Contact
    - X Coordinate
    - Y Coordinate
- Unit Test: I2C Latency with Unos



Carnegie Mellon

#### Integration

- Final product will be two wearables connected to a computer via USB
- Computer will be running a bespoke Max4Live Software Synthesizer



Carnegie Mellon

#### Schedule

#### Is Mayonnaise An Instrume...

#### **Project Logistics**

- Design Presentation
- Interim Demo
- **Final Presentation**
- 1 Slack Time
- 2 Finalize Hardware Requirements
- 3 Finalize CV Approaches
- 4 Place Hardware/Parts Orders
- 5 Create Design Presentation
- 6 Create Final Presentation and Demo

#### **CV** Detection Implementation

7 Research on CV Algorithm 8 Set up CV Board 9 Implement CV on to Board 10 T-Detection of an Object 11 T-Communication with MIDI Board 12 T-Communication with Video Boa... 13 T-Overall Integration 14 Add More Detectable Objects

#### Sensor Board Implementation

15 Acquire Data Refinement Method 16 Develop Data Refinement Method 17 T-Accuracy of Fingertip Touch Da... 18 T-Accuracy of Position Data 19 Create a Glove for the Board 20 T-Communication with MIDI Board 21 T-Communication with Video Boa... 22 T-Overall Integration 23 Add Gyrometer for Extra Features

#### **MIDI Board Implementation**

24 Research MIDI Packaging Method 25 Implement MIDI Generation on Bo... 26 T-MIDI Output on Max for Live 27 T-Communication with Other Boa... 28 T-Latency of MIDI Generation 29 Create a Max Patch for Synthesizer 30 T-Overall Integration 31 Improve Data Transfer Time

#### Video Board Implementation

32 Design UI for Visual Overlay 33 T-Communication with Arbitrary ... 34 Acquire Data from Other Boards 35 Create Headset for Display Mount 36 T-Overall Integration 37 Integrate CV and Sensor Features

