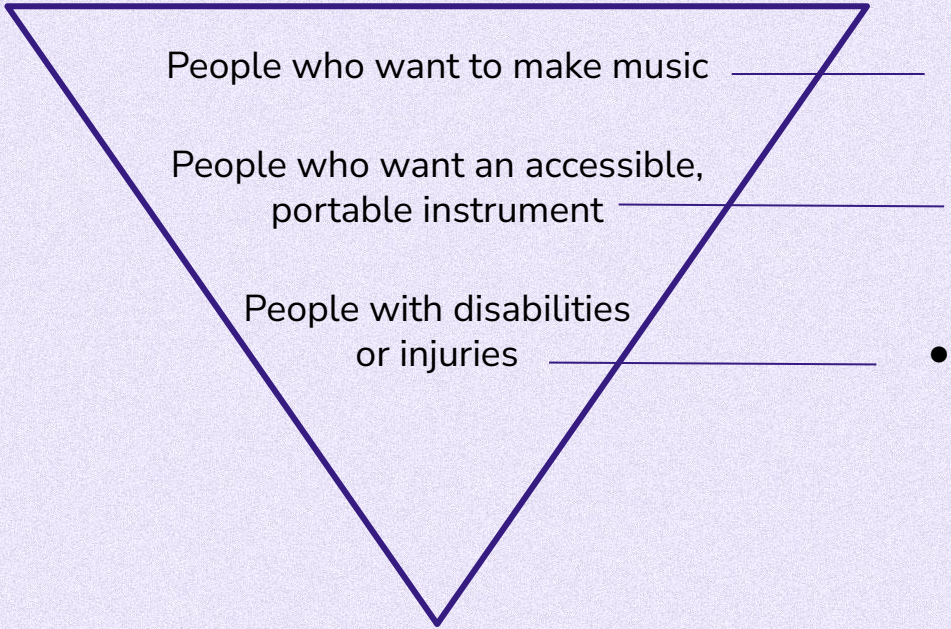
More Targeted Use Case

Who is our product for?



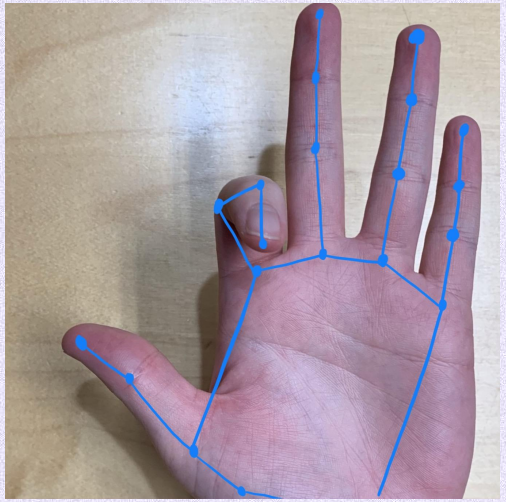
Use Case Requirements:

- Make an instrument with < 100 ms latency and customizable chord selection
- Non-bulky, lightweight, <100g total weight (slightly heavier than a smartwatch)
- Still have control over fingers, but no strength
  - E.g. can touch two fingers together, can bend fingers, can type
  - Ex. mild cerebral palsy, arthritis, stroke, Limb-Girdle Muscular Dystrophy, early stage ALS

Solution Approach & MVP Review

- **Chord Selection (Computer Vision)**
  - Detect up to 4 customizable chords based on finger bends
  - Target ≥ 95% accuracy within 0.3-1.5 m camera range
- **Strumming (IMU Wristband + Haptic Pads)**
  - IMU measures strumming motion speed and direction
  - Haptic pads register strum initiation events
  - Target strumming speed range: 60-300 bpm (1-5 strums / sec)

Eg. pointer finger bending switches chord to G



Quantitative Requirements

Left Hand Input

- Hand detection **0.3m to 1.5m** away from camera (laptop camera)
- **> 95%** accuracy on which finger/chord is selected
- **100 ms** latency (human reaction for music / drumming is 150-200ms)

Right Hand Input

- Hear difference of strumming **[60 bpm - 300 bpm]** (beats per minute) -> **[1-5 strums per second]**
  - Range set based on 100-140 typical song BPM
  - Using 0.2m strumming distance
  - [0.2 m/s -> 1m/s] speed detection
- Output different string sounds over continuous output range rather than discrete
  - Allows for “infinite” expression

Quantitative Requirements Cont'd

Mapping Input to Sound Output:

CHORD

- Select 1 chord between 4 chord options
- Set of 6 notes to play based on chord selection

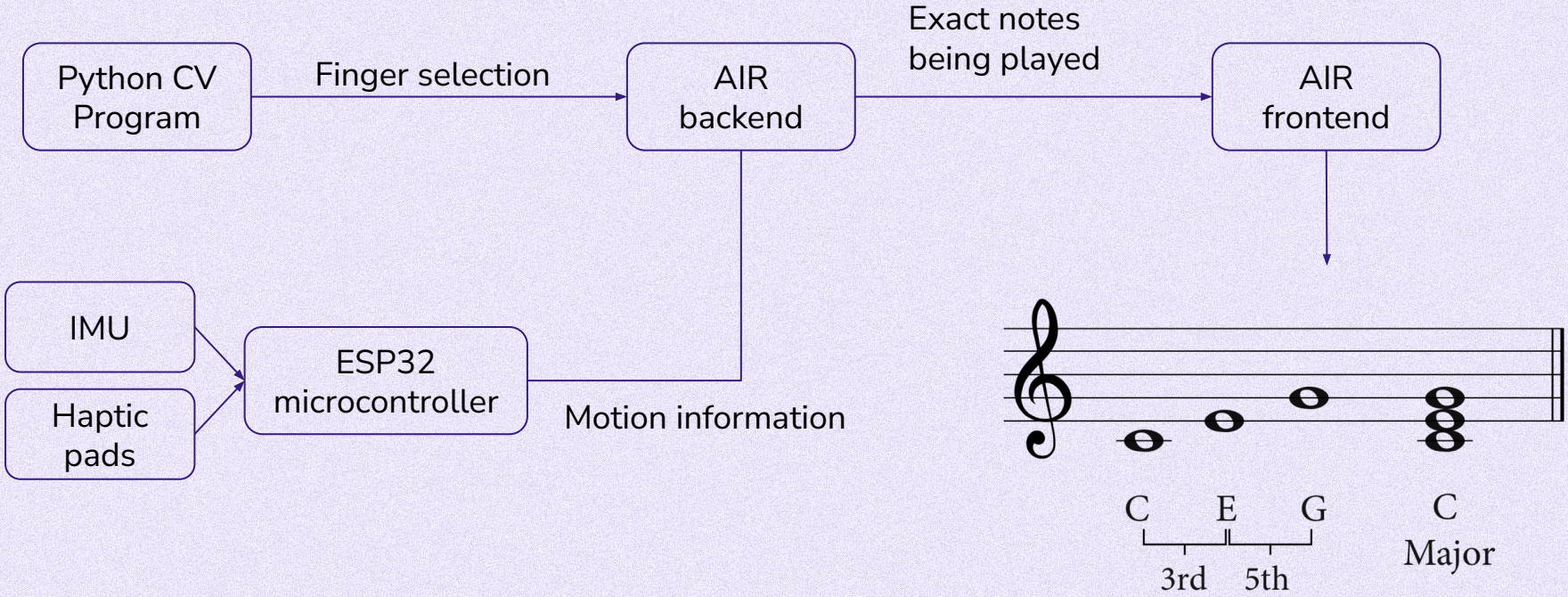
VOLUME & TIME BETWEEN NOTES

- Use transfer function to map speed (S) to volume **V**
- Also transfer function to map speed(S) to time **T** between notes of a chord

V (volume in dB) = min(S*slope, max_volume)

T (time in sec) = max(0, max_time - S*slope)

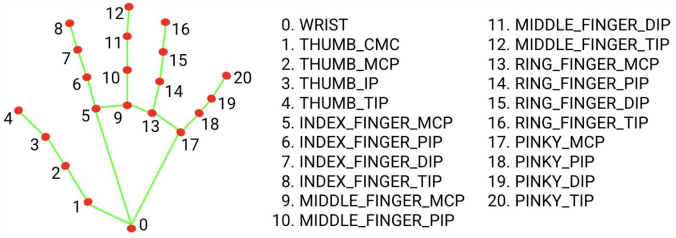
Overall System Specification



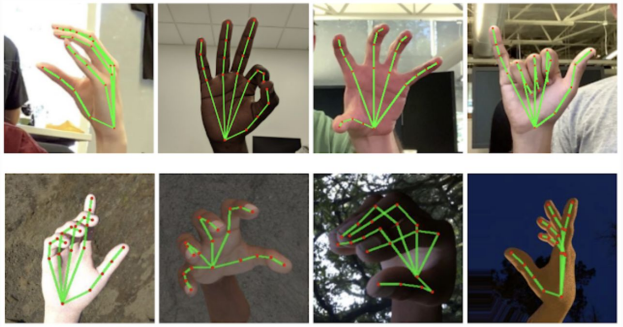
Implementation: Left Hand

Libraries

- MediaPipe hands library identifies points on a hand, and its x, y, z locations
- If fingertip point goes below join point, select that finger
- Can test different sensitivity models
- Output finger of highest confidence



hand_landmarks.png | :----------: | *Fig 2. 21 hand landmarks.* |



Source: https://mediapipe.readthedocs.io/en/latest/solutions/hands.html

Implementation: Right Hand

Hardware

- ESP32 S3 plus
  - Wifi communication for transmitting data to computer
  - Use I2C support for IMU integration
  - Link
- Makerfabs IMU:
  - I2C interface for communication
  - Link
- Haptic Pad:
  - Capacitance touch sensor
  - Will use built in I2C interface to integrate with ESP32 module
  - Link

Software

- Arduino program running on ESP32 module collects information from IMU and haptic pad peripheral
- Transmits information through wifi to computer

Implementation: Main Program

Flask

Python Backend

Strumming Program

- LAN for ESP32 Communication
- UDP socket to receive info from IMU
- Outputs volume and time between strums

Chord Selection Program

- Performs CV with mediapipe
- Outputs which finger is selected

HTML/CSS Frontend

- Display for selecting chords to choose
- Send chords, volume, speed to js program

JavaScript

- Outputs sounds given inputs

Test & Verification Metrics

Testing/Verification

- Overall, 3 iterations to test (based on number of classes we have)
- Strum detection accuracy ≥ 98% with ≤ 1% false positives/negatives (aiming for 100%)
- Chord detection accuracy ≥ 98% (based on similar studies)
- End-to-end success rate ≥ 95% (correct chord + strum output)
- Input-to-sound latency < 100 ms (within human tolerance of 150–200 ms)
- Setup time ≤ 1 minutes (average guitar tuning time)
- Wireless range ≥ 2 m (hand-gesture detection needs short distance)
- Wristband weight < 100 g; comfortable for ≥ 2 hours wear (avg. rating ≥ 4/5)

***All user specific tests will be conducted with School of Music students in 57-584**
***Currently have 6 set class days for testing**

source:

Test Inputs, Outputs, and Risks

Test Inputs

- Gestures: different finger bends for all chords
- Strumming at 60–300 BPM (1–5 strums/sec)
- Strum distances: 0.2 m at 0.2–1.0 m/s
- Camera distances: 0.3–1.5 m, varied lighting
- Wear trials: single 2 hr + repeated 30 min sessions

Passing Outputs

- ≥ 95% accuracy for chord + strum combined
- Latency < 100 ms
- Tempo tracking error ≤ ±5 BPM
- Continuous mapping of strum velocity to sound

Risks & Mitigation

- IMU drift / noise → calibration, filtering
- CV sensitivity to lighting → define desired conditions for user
- Latency budget overrun → optimize audio pipeline
- Comfort variability → multiple strap sizes, soft lining

Project Management