

# CS445 Project Proposal

## Motivation

We will build a webcam-based musical controller that maps a single hand's motion to sound controls in real time, taking very loose inspiration from [Imogen Heap's Mi.Mu gloves](#). This project lets us practice practical computer vision (hand signals, feature engineering, smoothing) with a creative, measurable output. We hope to learn how to turn messy landmark data into stable controls. We hope to adjust this project as it develops and use it as a learning opportunity.

## Milestones

- Part 1 (Week 1-2): Get the camera working; detect the hand and find the index of fingertips. Send two numbers (x and y) to a music app using a virtual MIDI cable on Windows (like loopMIDI). Make sure a free synth app makes sound.
- Part 2 (Week 2-4): Make the controls feel steady and smooth (may add simple averaging so they don't jump). Map left/right and up/down to two sound knobs. Use a pinch gesture to start/stop a note.
- Part 3 (Week 4-5): Test how fast it responds, how steady it feels when we hold our hand still, and if it works in different lighting. Record a short demo video, tests, and finish writing up our results.

## Evaluation

- Response time feels quick: under about 0.2 seconds from moving hand to hearing change.
- Steadiness: when we hold our hand still, the values don't wobble too much (small changes only).
- Reliability: it finds the hand most of the time in three scenes (bright room, desk lamp, busy background), and pinch works at least 9/10 times.

## Resources

- Open Source/Github Repos:
  - [LG-Gesture-And-Voice-Control](#) – good reference for the gesture-detection pipeline
  - [arpeggiator \(hand-controlled\)](#) – good reference for showing how to map from hand input to sound/music (good paradigm)
  - [Sign-Language-Gesture-Recognition](#) – another good reference for hand landmarks/gesture recognition
  - [Hand-tracking topic on GitHub](#) – just general resource for more info
- Research/Academic Papers:
  - G. D. Nagalapuram, R. S, V. D, D. D and D. J. Nazareth, "Controlling Media Player with Hand Gestures using Convolutional Neural Network," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), Hassan, India, 2021, pp. 79-86, doi: 10.1109/MysuruCon52639.2021.9641567 – A parallel idea to pull ideas from
  - [Mapping strategies for gestural and adaptive control of digital audio effects \(Verfaille et al., 2006\)](#) – really close to our idea of mapping a gesture to an audio effect, but somewhat old

- [Towards a real-time mapping of Finger Gesture to Sound \(Matsoukas et al., 2012\)](#) – even more direct gesture to sound mapping (but maybe simpler than full voice-effect modulation)
- [A Real-Time Gesture-Based Control Framework \(2025\)](#) – most recent paper! And pretty close to the scope of our actual project!

## Group Assignments

- Member A: Camera + hand detection; get fingertip position into our program.
- Member B: Connect to the music app (virtual MIDI), set up mappings, and the pinch gesture.
- Member C: Testing (speed, steadiness, different lighting), demo video, and figures for the report.

Please note that we want to finalize the project and receive feedback before specific group member assignments, may want more/less overlap.