

Computer Vision

Arnav Gupta

April 19, 2024

Contents

1	Creative Coding	1
2	Colour-Based	2
3	Optical Flow	2
4	AI models	2

1 Creative Coding

p5.js is a JS library for creative coding, ultimately a simplified and enhanced HTML Canvas API.

To use p5.js, install types, include CDN file in HTML, include 1+ p5 sketch JS files, and use with JS. Also, use BrowserSync to run code.

Runs in either global or instance mode:

- **global**: all properties and methods are attached to the window
- **instance**: all properties and methods are bound to a p5 object

Implied callback functions are:

- **setup()** which runs once when program starts
- **draw()** which is executed 60 times per second (by default)

Implied event functions like **keyPressed()**.

Global variables like **mouseX**, **mouseY**, and **key** (last key pressed).

Computer vision: broad class of algorithms that allow computers to make intelligent assertions about digital images and video

Computer vision for human computer interaction is when a computer is controlled using a video camera.

2 Colour-Based

A simple computer vision algorithm identifies a special pixel based on some criteria and uses that pixel's location to control that computer, for each video frame.

This special pixel can be found by looking through all pixels for the most special one.

Can also count special pixels based on some criteria

3 Optical Flow

A more complex computer vision algorithm subtracts this frame from the last one, and uses the differences to estimate how far each pixel moved. Then uses the total pixel movement to control the computer.

An optical flow vector is calculated in each flow zone, which calculates the overall flow vector.

Can also subtract out background pixels, find blobs of connected pixels, and use blob position and/or shape to control the computer.

4 AI models

An even more advanced computer vision algorithm would send the frame to a deep learning model to recognize features, and use features to control the computer.

Configure model options, initialize the model. Have a callback for when predictions are made. Draw keypoint landmarks (ex. mouth top-middle and bottom-middle).