Hardware Considerations

Gimbal will be powered via USB, but stepper motors will run off 12V instead of 5V, so how will we power them? Either have a separate power supply (inelegant) or use a boost converter and current-limit the motors so that they can't draw more than a total of 0.5A (USB 2) or 0.9A (USB 3).

Alex Anderson-McLeod

Control system for the gimbal will consist of an ESP32 microcontroller connected to A4988 stepper drivers for controlling the motors, as well as to optointerrupters for determining when each axis is in its home position. The ESP32 will receive commands via serial from the face-tracking program on the host computer.

Alex Anderson-McLeod

Minimize mechanical limitations of the gimbal hardware to allow it to pan and tilt as far as possible. This is a major limitation of current solutions.

Alex Anderson-McLeod

It's probably good to use stepper motors for actuating the gimbal because of their high torque, precision, reliability, and ease of control.

Alex Anderson-McLeod

Make gimbal as compact and attractive as possible. Most optimal way to accomplish this is by designing the gimbal in CAD and then 3D-printing it.

Alex Anderson-McLeod

Backend Software

Since Pytorch is the most popular, most well-documented and arguably easiest to use ML library in existence right now, we will be using it for all our face-tracking code.

Alex Anderson-McLeod

We will start with an incredibly simple face classifier that uses a window-based scanning technique to determine the location of a face in an image. Once that is working, we will add additional capabilities, such as tracking the sides and rear of the face, and potentially cascading different models together, for improved results.

Alex Anderson-McLeod

All backend code on the host computer will be written in Python, while the code running on the ESP32 in the gimbal itself will be written in C++.

Alex Anderson-McLeod

We will need some fairly powerful GPUs for training our models. Where can we get low-cost access to a machine like that? Dr. Bakos says that he has a couple multi-GPU systems, so hopefully he'll let us use those.

Alex Anderson-McLeod

Since everything will run locally on the user's machine, we shouldn't need any central databases or servers or anything like that. Unless we implement the idea of taking the on-the-fly trained models from users, uploading them to a server, and using them to improve our core model.

Alex Anderson-McLeod

User Interface

None of my Capstone team members, myself included, have ever made a UI in Python, so we'll need to do some research into the best frameworks for this.

Alex Anderson-McLeod

UI should allow users to toggle the facial tracking feature on and off, allow them to choose what happens when tracking is lost, show a preview of the video feed, and allow manual control over the gimbal's position.

Alex Anderson-McLeod

Should we add the ability to control the camera with gestures, so that a limited featureset can be used without the UI, like with the Insta360 Link?

Alex Anderson-McLeod

Are there any additional/more advanced features that we should add to the UI that I'm not thinking of right now?

Alex Anderson-McLeod

Safety and Security

We should do all image-processing on the host computer without uploading anything to the cloud for the sake of user privacy.

Alex Anderson-McLeod

We plan on using data from the camera to train our model on-the-fly to improve its effectiveness. Given that we wouldn't be able to see any sensitive user images, would it be an invasion of privacy for us to upload these optimized models from users' computers to our servers to improve our core model?

Alex Anderson-McLeod

Given that I'm a big proponent of open-source software, we should release significant portions of our codebase as open-source in order to build trust with users, allow them to modify and improve things, and show them that we aren't engaging in any unethical behavior that many other tech companies practice when it comes to user data.

Alex Anderson-McLeod

Ensure that there are no gaps or crevices on the gimbal that a user could get a body part caught in while the gimbal is moving. If this is impossible, then make sure that the motors are current-limited to the point that they would not have enough torque to cause harm to the person if they were to get caught in the device.

Alex Anderson-McLeod

Time Management

I've already designed a pretty decent CAD model of the gimbal, which just needs to be 3D-printed and have the electronics installed.

Alex Anderson-McLeod

We will begin prototyping a simple facial-recognition proof-of-concept later this week, and hope to have this done within two weeks.

Alex Anderson-McLeod

We want to have the gimbal hardware and the embedded software that controls it done by October 15th.

Alex Anderson-McLeod

Once the proof-of-concept is done and able to recognize faces, it will need to be linked to the gimbal, which should take no more than a week. At this point, its performance can be compared to the Insta360 Link and other competing options to determine where there is room for improvement.

Alex Anderson-McLeod

Over the course of the next several months, we will steadily improve the design based on our research into the most cutting-edge AI-based face-tracking techniques, comparing with the Insta360 Link's performance along the way until we are satisfied with the performance of our offering.

Alex Anderson-McLeod

At this point, if time permits, we can add additional features, such as the ability to maintain tracking while the subject is turned away from the camera, that current camera gimbal options lack. This phase will run until the end of the school year.

Alex Anderson-McLeod