Providing a Personal Drone with Emotional Awareness

(CPSC 490 Team Project Proposal)

Maansi Dasari and Aman Heyer

Advisor: Professor Lin Zhong

Background

This project is an offshoot of the larger project known as POD: A Smartphone That Flies, which is being worked on by our advisor, Professor Lin Zhong, along with Guojun Chen of the Yale Efficient Computing Lab. POD was created as an alternative method of hands-free computing to AR glasses or headsets, and the personal drone has the ability to maneuver independently of the human body, providing great flexibility and opportunities for interactive control that are not possible with any other hands-free computing solution. The personal drone has been programmed by the Yale Efficient Computing Lab already to stay in front of the user's face as they move and it can respond to hand gestures and air clicks as well.

Project Overview

We aim to further the personal drone's ability to function as an emotionally aware computing companion in this project. Despite its best intentions, a drone that can move independently can be seen as threatening to a user who is unfamiliar with the system, and the last thing we want is for someone to feel uncomfortable with a drone in their face. Intelligently understanding when someone wants to interact and when they might want some space is an easy

task for a human and one that is required often in social contexts, and this is just one situation of many in which recognizing human emotions would be helpful.

To solve this problem, it might be enough to provide the user with total control of the drone's position with hand gestures, but the ultimate goal of POD is to be a helpful robotic assistant that can adjust its behavior without tasking the user with controlling every minute movement. If the personal drone is to be fully trusted as a social computing assistant in this way, it must be able to recognize different facial expressions, which often provide a lot of insight into a person's emotional state. In this project, we will address these concerns by creating an Android application that will give the personal drone the ability to discern between different human emotions and adjust its movements accordingly. The minimum functionality that we hope the drone will have at the end of the project is for it to be able to discern between positive and negative emotions, to back up from the user if they react negatively, and to approach the user (maintaining a safe distance) if they react positively. Our hope is that there will be time to broaden the personal drone's emotional vocabulary and program more advanced behaviors based on this knowledge.

Recognizing facial expressions is an old problem, and a lot of progress has been made in the field. There exist several Github libraries like EmoPy and deepface that are based on neural networks and are accurate enough for this project. Libraries such as these would require a good deal of work to integrate into an Android app, but it does seem possible to do so. There also exist services that have public APIs such as Microsoft Face API. The benefit of these types of services would be to make the offloading of computations trivial. Should none of the existing models prove satisfactory, we will use the FER (Facial Expression Recognition) dataset, which contains

almost 30,000 faces with labels for emotions, to train our own neural network and/or random forest models.

Deliverables

- Research different options for facial expression recognition. Options for this include:
    - Github libraries like *EmoPy* or *deepface* that have trained emotion recognition models
    - Microsoft Face API
    - Use MediaPipe Face Mesh API to obtain facial geometry from images, and train a model to classify the geometries into facial expressions
    - Training our own models on the FER dataset. Potential model types include random forests and neural networks.
- Adjust the distance between the drone and the user based on the facial expressions it recognizes: Positive Expression ⇒ A smaller distance; Negative Expression⇒A larger distance.
    - This can be done by using existing APIs built by the Yale Efficient Computing Lab but will require new integration.
- Expand the vocabulary of facial expression beyond {Positive, Negative} to incorporate more of the 7 base emotions {Happy, Sad, Neutral, Angry, Disgusted, Surprised, Afraid} and link with another senior thesis project (Jinny Choi's) to determine how the drone moves according to recognized facial expressions.

- [Extension] Offload major computations from phone to a cloud server to leverage greater computing resources. (If using the Microsoft Face API, this step will not be necessary)
- [Extension] Maintain an idea of who in a group is most comfortable with the drone based on their facial expressions and tone of voice and act accordingly
- [Extension] Expand emotion recognition capabilities to be able to detect voice commands and/or emotion in voices and respond accordingly. Options for this include:
    - Using an existing voice emotion recognition API like Empath API
    - Training a new model on voice emotion datasets
    - Integrating with facial expression recognition model to get an overall description of a person's emotions