Project Proposal

Team member List:

- Jessica Soong
- Rena Liu
- Cindy Huang
- Ping Cheng Ku

Problem Definition:

Our project will be comparing three different methods for two different tasks: (1) separating the hand from a motion video image, and (2) classifying gestures from the results of hand segmentation. For task (1) we will be implementing and comparing the results of the optical flow and depth methods. The optical flow segmentation method should result in a vector flow field as well as a sequence (or video) of objects that are moving in the original input. We can then take either input and feed it into task (2), which will be used to classify the hand gesture using deep learning. With the depth segmentation method, we will compare different levels of thresholding to segment, as well as adaptive thresholds (otsu's method).

For task (2) we will be trying a deep learning method on the various results from the segmentation task, as well as just the original input. For the deep learning method, there are potentially four identified separate inputs to test: (1) the original sequence, (2) the thresholded depth sequence, (3) the vector flow field of the sequence, and (4), the results of optical flow segmentation. For a non-learning based method, we will be trying a non-learning based method which uses border and convex-hull detection (Rios-Soria). The convexity method will specifically be classifying the image by counting the number of fingers held up in the image.

Problem Breakdown:

After deciding on our input dataset, which we have currently specified to be video motion images cropped to have the hand be in the foreground, we will implement the various methods of segmentation and classification. We will create classification input by performing the optical flow segmentation method, which will return a vector flow field, and performing depth segmentation, adjusting various parameters as well. We will then label these inputs with gestures (pointing up, closed fist, five fingers, etc.) both with a RCNN network and with convex hull.

Team Member Assignments:

Completing task (1) using the depth method will be much simpler than using optical flow and tracking, so once the depth segmentation is completed, the classification task can be started earlier. The team member assignments are broken down as follows:

Task	Method	Team Members
Segmentation	Depth	Jessica, Rena
	Optical Flow	Jessica, Rena, Cindy
Classification	Deep Learning	Ping-Cheng, Jessica
	Convexity Method	Cindy

Datasets:

- Multimedia Technology and Telecommunications Laboratory, University of Padova https://lttm.dei.unipd.it/downloads/gesture/
- IBM Research from 2017 paper, “A Low Power, Fully Event-Based Gesture Recognition System” https://www.research.ibm.com/dvsgesture/
- Kaggle data, binary https://www.kaggle.com/benenharrington/hand-gesture-recognition-database-with-cnn

Literature Review:

Rios-Soria, D. J., Schaeffer, S. E., & Garza-Villarreal, S. E. (2013). Hand-gesture recognition using computer-vision techniques. Retrieved November 01, 2020, from https://core.ac.uk/download/pdf/295558123.pdf

A. Amir *et al*., "A Low Power, Fully Event-Based Gesture Recognition System," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 7388-7397, doi: 10.1109/CVPR.2017.781.

Heintz, B. (2020, February 12). Training a Neural Network to Detect Gestures with OpenCV in Python. Retrieved November 01, 2020, from https://towardsdatascience.com/training-a-neural-network-to-detect-gestures-with-opencv-in-python-e09b0a12bdf1?gi=e620f2d30367

On-Device, Real-Time Hand Tracking with MediaPipe. (2019, August 19). Retrieved November 01, 2020, from https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html