Sean Yang

5/9/19

Pd 4

Journal report 27

In the previous presentations, I briefly went over the intuition of the methods I used. For example, when I discussed my HOG + SVM method for the face detection method, I said it extracts the features and then predicts based on the feature. Although the explanation is not wrong and it’s likely what I will say in my presentation in the interest of time, it would be nice to get a better intuition of all of the methods and just in case someone asks me about the way a particular method works.

The HOG + SVM method is fairly simple to understand. HOG stands for the histogram of the oriented gradient. Here are the steps:

1. Calculate the direction of the gradient from small cells
2. Sum up the gradient in each cell so stronger directions in cell show
3. Get HOG for the whole image
4. Get patches and slide the boxes through
5. See if the face is present using SVM. SVM divides data in a vector space into face and not face

I tried to understand the method of facial alignment. It uses the ensemble of regression trees. I understand the ensemble of regression tree; it divides data into different intervals and fit a line through the interval. The article says “face alignment can be solved with a cascade of regression functions.” However, I don’t quite understand why this works intuitively. The paper is quite math heavy and I will have to take a look at it.

I also figured out how to load images to CNN and tested the model I trained. It’s good at telling if a gesture is neutral or nonneutral but not good at distinguishing between left and right.