

CS3600 Research Extra Credit

Paras Jain (pjain67)

April 2016

1 What I did

I worked with Yun Zhang to generate training data for a project that aimed to use Point-Of-View camera data to determine when a person was looking at a television or not.

2 Project Goals

The project Yun is working on aims to determine when a person is looking at a television. This is important as the research group wants to use such a tool to determine the impact of watching TV on eating habits.

This problem is technically difficult for many reasons - one is that detecting a television is not a trivial problem. While Deep ConvNets can be an effective method to test for the presence of a TV in an image, there are many TVs in varying environments (dark rooms, flat-screen, CRT, computer, etc).

This is also a difficult problem as determining attention is difficult. Part of what Yun was trying to do was to use the time evolving nature of video to track the motion of the TV to see if that motion could be mapped to attention.

3 Machine Learning and AI learnings in the project

In order to detect the TV, a Deep ConvNet is being used. I assume this is paired with some form of a sliding window technique to pinpoint the center of the TV and to extract scale data about the TV. Pre-trained ConvNets are already very accurate at detecting TVs in common scenarios but I anticipate the researcher may have trouble in some of the less-optimal scenarios I outlined above.

Given now the position of the TV as a function of time, it seemed that Yun planned to use some variant of a HMM (Hidden Markov Model) to model the problem of attention. This makes sense given the annotation task where the TV can be in states $\{focused, moving, unfocused\}$. A HMM makes sense given these states.

4 Conclusion and Future Work

Overall, I found this project to be interesting. I am actually interested in looking into some facets of this research myself - particularly, I think this same approach can be used to say detect when a person is smoking (by tracking the movement of a cigarette) and use that data to determine if a person is a heavy smoker or a light smoker. This technique also could be used for tracking lecture effectiveness as a camera on a student could track when they are paying attention to the whiteboard.