

Project Proposal - Aerial Note-taking

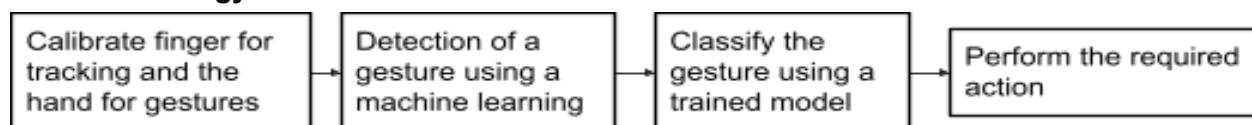
- By Justyna Kaczmarzyk <kczmrzy2@uic.edu>, Ayush Gupta <agupt69@uic.edu>

Motivation: In this age of digitization, it is now popular to use a tablet along with a bluetooth connected pencil. This tool brings in a lot of merit and helps deliver content in an efficient manner; however these devices come with a few disadvantages, such as bluetooth connectivity issues, replacement of pencil tips, and making sure it is charged.

Problem Statement: We would like to propose the idea of aerial notetaking where a user is allowed to use hand gestures to achieve the same utility that the pencil would provide. The user would be able to write or draw on a plain canvas and use a range of tools such as a pencil, marker, an eraser, etc. with a variety of hand gestures. This kind of product would be useful for students, lecturers, virtual meetings, and generally any presentation setting. By freeing the user's hands, it also has potential to be extended into other fields such as sign language recognition, virtual reality, computer games, and more.

Prior Research: Researchers at Institute for Human-Machine Communication, TU Munich, Germany came up with a methodology to improve real-time hand gesture recognition. Their idea consisted of using two 3D CNNs working in tandem. The first CNN is a lighter of the two in terms of resource consumption and is tasked to find whether a gesture was present in the current sliding window or not, once a gesture is detected it is passed to the second CNN that is required to classify the gesture. The researchers further included the Levenshtein distance as an evaluation metric as it is capable of detecting misclassifications. Meanwhile, researchers at University of Aizu in Fukushima, Japan used Kinect, a motion sensing input device produced by Microsoft, to obtain image, voice, and depth data in order to detect the joint positions of the user's body. This information then allowed them to detect users' hands and recognize their gestures. Researchers at Al-Mustansiriyah University also used a Kinect sensor along with deep learning structures such as the DenseNet Convolutional neural network.

Our Methodology:



Topics from Class:

- PCA, principal component analysis (lecture #2)
- Object detection (lecture #8)
- SIFT for object detection (lecture #11)
- Image segmentation (mentioned in syllabus)
- Convolutional Neural Network (mentioned in syllabus)

References

- <https://journals.sagepub.com/doi/full/10.1155/2014/278460>
- <http://jsju.org/index.php/journal/article/view/466/461>
- <https://www.hindawi.com/journals/tswj/2014/267872>
- <https://arxiv.org/pdf/1901.10323v3.pdf>