

EE 368 Term Project

Real-Time Sign Language Recognition on Android

Team Members: Shalini Ranmuthu [shalinir@stanford.edu], Ishan Patil [iapatil@stanford.edu], Hans Magnus Ewald [hmewald@stanford.edu]

Project Idea and Goals

For our project, we would like to develop an algorithm for human hand detection and interpretation. This means that we not only want to identify human hands in images, but also correctly model their pose and classify gestures. Although this algorithm has many varied applications such as human-machine interface commands or machine understanding of human behavior, we will use this algorithm specifically to detect and interpret American Sign Language gestures.

Specific goals for the project include (1) live video detection and conversion of sign language to text in real time and (2) mobile implementation on an Android device.

Proposed Methodology

In order to do this idea justice, we plan to organize the project into an image processing component for hand segmentation and modeling of the gesture (for EE 368), and a machine learning component for hand pose interpretation (for CS 229). We would like the image processing component of this project to detect both the position and configuration of human hands. This will make the overall algorithm more robust in terms of the camera angle, lighting conditions, etc.

The first step should be a form of segmentation that can identify human skin tones/ hand geometry. We then intend to include some redundant measures of identifying hands, in order to counter scenarios in which one method might fail to pick up the presence of a hand. We have already identified potential datasets (such as this American Sign Language Lexicon Video Dataset) to use in development.

Once a region of interest has been identified as containing a hand, the last detection step would involve creating a model of the hand pose. The features extracted from this model would then be provided to the machine learning component of the project. We will initially implement this algorithm on still images, and then tackle video streams if time permits.

References

- Kuznetsova, et.al. "Real-time sign language recognition using a depth camera"
- J. Isaacs, S. Foo. "Hand pose estimation for American sign language recognition."