



**CALIFORNIA STATE SCIENCE FAIR
2016 PROJECT SUMMARY**

Name(s) Dhruba J. Ghosh	Project Number 36449
Project Title Viseme Analysis: Implementing Contour Sequence Classification to Augment Speech Recognition	
Abstract Objectives/Goals To test the efficiency of various lip contour detection methods as well as machine learning algorithms in using a sequence of images of a speaker to determine what the speaker was saying. Methods/Materials Laptop computer with Python development kit, open source computer vision and scientific computing libraries, and dataset of videos obtained from public LILJ dataset. Implemented feature extraction algorithms and sequence classification algorithms by merging existing models and tested accuracy by running on a separate dataset. Results The final design, which linked adaptive thresholding, a support vector machine, and a hidden Markov model, predicted a spoken letter based on solely video data. It attained a 43% average recognition rate on the test data. Conclusions/Discussion I created a segmented model that took raw image data from a video and predicted the letter that was spoken. This method can be extended to cover dictionaries larger than the English alphabet, and due to its segmented nature, the extracted features can also be added to an audio-based speech recognition system. Though viseme analysis using lip contour detection and hidden Markov models cannot function professionally as a standalone program, it can be effective in filtering out noise in existing speech recognition systems.	
Summary Statement I devised an algorithm to implement speech recognition through lip reading.	
Help Received All code was written by me using open source libraries. Publicly available datasets were downloaded from the University of Surrey website.	