# SANTA CLARA UNIVERSITY
## DEPARTMENT OF COMPUTER ENGINEERING

Date: April 9, 2015

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY

**Alex DeBoni**
**Jesse Harder**

ENTITLED

# GPU-Accelerated Lip-Tracking Library

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF

BACHELOR OF SCIENCE IN COMPUTER SCIENCE AND ENGINEERING

_____
Thesis Advisor

_____
Thesis Reader

_____
Department Chair

# GPU-Accelerated Lip-Tracking Library

by

Alex DeBoni
Jesse Harder

Submitted in partial fulfillment of the requirements
for the degree of
Bachelor of Science in Computer Science and Engineering
School of Engineering
Santa Clara University

Santa Clara, California
April 9, 2015

# GPU-Accelerated Lip-Tracking Library

Alex DeBoni
Jesse Harder


Department of Computer Engineering
Santa Clara University
April 9, 2015

## ABSTRACT

A major part of having correct pronunciation when learning a new language is moving your lips in the correct way. One solution to this is software which will track a student's lip movements and provide feedback. This paper describes how we are creating a C++ library to accurately track lips in provided images. Further, this library will attempt to use a CUDA-enabled GPU implementation to improve the algorithm's performance. It will fall back on a CPU implementation if such a GPU is not found. As a result, the lip tracking library will run on Windows, Linux, and OS X, as well as Android devices.

# Table of Contents

# List of Figures

# Chapter 1

# Ethical Analysis

## 1.1 Ethical Justification for the Product

The world is a big place with many people living in it. Every day, however, advances in technology are increasing the connectivity between people and cultures around the world. This technology provides a means by which anyone around the world could conceivably connect with anyone else. One of the inherent barriers to this process, however, is the fact that not all people speak the same language. In fact, there are thousands of different languages spoken around the world.

The focus of our Senior Design Project is to develop GPU accelerated lip tracking software that can run reliably and accurately for real-time video processing. We are doing this primarily so that the software can then be used in language learning software that is being developed by the computer engineering and language departments at our university. The language program will use the data obtained with our lip tracking software will provide in addition to audio information from videos of users speaking in front of a camera. This information can then be used to more effectively instruct the user as to how they might improve his or her pronunciation.

The main ethical justification for our product is based on its use with the language learning software for which we are creating it. Ideally, the way our product is used with the language learning application will help people to learn to improve their ability to speak additional languages. This increase in linguistic ability is important because being able to speak multiple languages increases a person's capacity for connecting with others in the world. More people becoming connected in this way fosters a greater sense of global community, which in term can lead to greater global moral awareness. Our hope is that the software we are making can help to bring humankind one step closer to a more socially connected world.

It's use in this language learning software is by no means the only application for our lip-tracking software. There are many other ways that lip-tracking software that is both accurate and fast could

be useful. For example, one might use it to design lip reading software that tries to interpret what someone is saying just from seeing his or her mouth move. With such software it would be easier to do automated captioning of video so that blind people might be able to understand what is going on. With translation software this could be expanded to work for people who do not speak the language of the video as well.

## 1.2    Team and Organizational Ethics

First and foremost, we place a high priority on being ethical to each other by each performing a fair share of the work. It is unfair for one person to have to do a majority of the project by himself. We intend to clearly plan out which of us will take the lead on which portions of the project so that this will not happen. Should one of us become ill or otherwise indisposed, the other will obviously step in to pick up the slack, but the sick member will do his best to make up for the lost time once he is feeling better.

In addition to our ethical obligations to each other, the two of us also have an ethical obligation to our advisor and the rest of the team working on the full product. We intend to make sure to keep our advisor well-informed of our progress throughout the year. This way she can better aid us if we should run into trouble along the way. We also will not lie about our progress if we fall behind, as this will only hurt us, as well as everyone else involved in the project.

Another ethical issue is plagiarism. It is easy to copy code off the internet and pass it off as one's own without citing the author. We will not do this as it is not only ethically wrong but incredibly easy to avoid. We will simply cite the author of any useful library or snippet of code that we should find and utilize for our lip-tracking code library.

Finally, we will avoid wasting Santa Clara University's resources. Any funding that we should obtain from the school will be appropriately budgeted to make sure none goes to waste and that all of it is beneficial to our project.

Lastly, all of our actions will comply with the IEEE/ACM code of ethics. This means that we will do work in the best interests of our advisor, as long as it is also in the public's best interest. Our code will be written in the highest quality possible and will not be copied unjustly from others. And finally, by making our code open-source, we will allows others to use and learn from our work for their own projects.

## 1.3   Product and Society

Overall, our project should prove helpful to society by serving to increase people's linguistic diversity. We can, unfortunately, see a nefarious use for our lip-tracking library, which is its use in surveillance applications. It would technically be possible for our project to be used to lip-read people through CCTV cameras and figure out what they are saying. The Criminal Law Handbook states that we should not expect privacy in a public place. Most people, however, expect that their conversations will be private when others are not around them. Despite this, we do not believe that the ethical fault lies within our product, but rather in the way that people might choose to use it. After all, GPU-accelerated lip-tracking algorithms don't spy on people, people spy on people.

Since our product handles images of people's faces, a concern some people may have is that it might collect and send the photos to a malicious third-party. To combat this, our project's code will also be completely free and open-source. This means that people using our library can feel safe that our code won't do anything undesirable. For example, our code will not collect images or other data of the user and it will not send any data to a remote server. Additionally, open-sourcing the code allows others to benefit from our code by using it for their own projects.

# Bibliography

[1] Ahmad B. A. Hassanat. Visual speech recognition. *CoRR*, abs/1409.1411, 2014.

[2] Jian Li, Yuqiang Lu, Bo Pu, Yongming Xie, Jing Qin, Wai-Man Pang, and Pheng-Ann Heng. Accelerating active shape model using gpu for facial extraction in video. In *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*, volume 4, pages 522–526, Nov 2009.

[3] Sbastien Stillittano, Vincent Girondel, and Alice Caplier. Lip contour segmentation and tracking compliant with lip-reading application constraints. *Machine Vision and Applications*, 24(1):1–18, 2013.

[4] Jinwei Wang, Xirong Ma, Yuanping Zhu, and Jizhou Sun. Efficient parallel implementation of active appearance model fitting algorithm on gpu. *The Scientific World Journal*, page 13, 2014.

[5] Mau-Tsuen Yang, Zhen-Wei You, and Ya-Chun Shih. Lip contour extraction for language learning in vec3d. *Mach. Vision Appl.*, 21(1):33–41, October 2009.