

NATURAL EYE CONTACT USING GENERATIVE ADVERSARIAL NETWORKS
LIMITATIONS OF VIDEO CONFERENCING AND ITS EFFECTS ON SOCIETY

An Undergraduate Thesis Portfolio
Presented to the Faculty of the
School of Engineering and Applied Science
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By

David Chen

May 6, 2021

SOCIOTECHNICAL SYNTHESIS

The forced integration of video conferencing to almost every aspect of society has had significant impacts on society. The STS research paper analyzes the effects that the technology of video conferencing has on society. It reveals flaws and gives motivation for the technical project. The technical portion of the research works to improve the technology of video conferencing by creating a new method of simulating natural eye contact in video calls.

The goal of the technical project was to find a way to provide the most natural form of video calling and simulate in-person conversations. This technology will then allow for more natural and intimate person-to-person communication and remove distracting barriers to communication. In order to carry out the project, I used a conditional generative adversarial network to perform image-to-image translation. The model I trained was then able to convert images into slightly different images which appear as if they were taken from a lower angle. This angle difference would then account for the distance between the webcam and the eyes of the video caller on the screen.

After 8 hours (100 epochs) of training using 10 minutes of video, the model was able to successfully convert images to a different angle. Since the purpose of the technical project was to prove that such an approach was viable, the data that the model received was simple and had a smaller resolution of 256 x 256 pixels compared to the standard 1280 x 720 pixel resolution that is normally used. However, because of the

successful results of the project, future work can be done to improve and adjust the model.

The STS research conducted sought to analyze the limitations of video conferencing and its effects on society. By using the Actor-Network Theory framework, several groups of actors including hosts, participants, engineers, and institutions were identified. A case study was performed to explore how the use of video conferencing in language learning caused several barriers that hindered the learning process.

While conducting the research I discovered that video conferencing had a lack of eye contact. I also found that an unequal distribution of access to technology existed, causing certain groups in society to be more severely affected than others. The lack of eye contact, among other factors like more opportunities for distraction, causes a disconnect between video conferencing technology users. The research concluded that society had to better adapt to the use of technology and that technology had to improve, through better connection and the development of eye contact correction technologies.

This technical research concluded that conditional generative adversarial networks and image-to-image translation are a viable and potentially a better approach to current methods of eye contact correction in video conferencing. This improvement to video conferencing is one of the ways we can more effectively use the technology of video conferencing in society. Overall, it was not only beneficial but also an enlightening process to search for ways to combat the limitations of video conferencing and further improve the lives of all technology users.

TABLE OF CONTENTS

SOCIOTECHNICAL SYNTHESIS

Natural Eye Contact Using Generative Adversarial Networks

Technical advisor: Mark Sherriff, Department of Computer Science

LIMITATIONS OF VIDEO CONFERENCING AND ITS EFFECTS ON SOCIETY

STS advisor: Catherine D. Baritaud, Department of Engineering and Society

PROSPECTUS

Technical advisor: Mark Sherriff, Department of Computer Science;

STS advisor: Catherine D. Baritaud, Department of Engineering and Society