

YouTube Mirror

Sihwa Park
York University
4700 Keele St.
Toronto, ON, Canada
shpark@yorku.ca

1. DESCRIPTION

YouTube Mirror is an interactive, audiovisual AI installation that generates images and sounds in response to the images of the audience captured through a camera in real time as a concept of data mirror. YouTube Mirror uses a cross-modal generative machine model that was trained in an unsupervised way to learn the association between images and sounds obtained from the YouTube videos I watched. The machine will see the world only based on this audio-visual relationship, and the audience can see themselves through the machine-generated images and sounds. YouTube Mirror is an artistic attempt to simulate my unconscious, implicit understanding of audio-visual relationships that can be found in and limited by the videos I watched. YouTube Mirror also attempts to represent the possibility of the machine's bias caused by implicit bias inherent in video recommendation algorithms as well as a small set of personal data.

We try to understand the relationships between images and sounds when we watch videos. Along with the popularity and impact of video-based social media platforms such as YouTube, we watch a plethora of videos, and our video consumption affects how we see other people and the world. What videos we watch are not only determined by our choices but also hugely affected by the recommendation algorithms of the platforms that are designed to make their users remain longer on their platforms. The "watch-to-next" videos suggested by the algorithms are, in general, based on the user's previous watch history and other metadata related to the videos. Since these data could be implicitly biased or wrongly reflect the user's behavior or preference, the recommendation models could cause a feedback loop effect, narrowing down the choices of videos the user can find [1, 2]. This feedback loop will affect our understanding of audio-visual relationships that we unconsciously find in the videos we watch. This project tries to make a machine simulate these audio-visual associations and represent the world through the relationships the machine learned.

The YouTube Mirror project lies at the intersection of personal data, machine vision, and data art in an attempt to explore how data-driven audiovisual art can represent the machine's understanding of audio-visual relationships limited by a small set of personal data and the social media platform.

2. VIDEO DOCUMENTATION

YouTube Mirror was exhibited as an installation at the Media Arts and Technology Program 2022 End of Year Show at the University of California, Santa Barbara (See Figure 1).

The video documentation of the installation is available at <https://sihwapark.com/YouTube-Mirror>.

3. SPACE REQUIREMENTS AND FLOOR PLAN

Figure 2 illustrates the floor plan for a YouTube Mirror installation. The installation is designed to enable the audience to interact with the work in real time via a webcam. The machine-generated images will be displayed on a portrait mode monitor, and the generated sounds will be played through 2-channel stereo speakers. This installation plan includes a video grid as a visualization of the video data used for training the model via the background projection. The video grid will be looped on the wall screen with no sound. The installation is more suitable for an indoor space, such as an indoor plaza or hallway in a University building, and a gallery.

4. LOGISTICAL REQUIREMENTS

The YouTube Mirror installation usually takes less than an hour to set up. The work does not have any networked needs.

5. EQUIPMENT REQUIREMENTS

5.1 An Equipment List from the Author

A list of equipment I will bring as below:

- A PC
- A webcam
- A portrait-mode monitor with an arm
- An audio interface

5.2 An Equipment List from the Conference

A list of equipment the conference organizers need to provide is as below:

- 2-channel stereo speakers
- A projector
- A wall screen
- A pedestal
- A box to hide the pc and the audio interface



Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

NIME'23, 31 May–2 June, 2023, Mexico City, Mexico.



Figure 1: YouTube Mirror installation

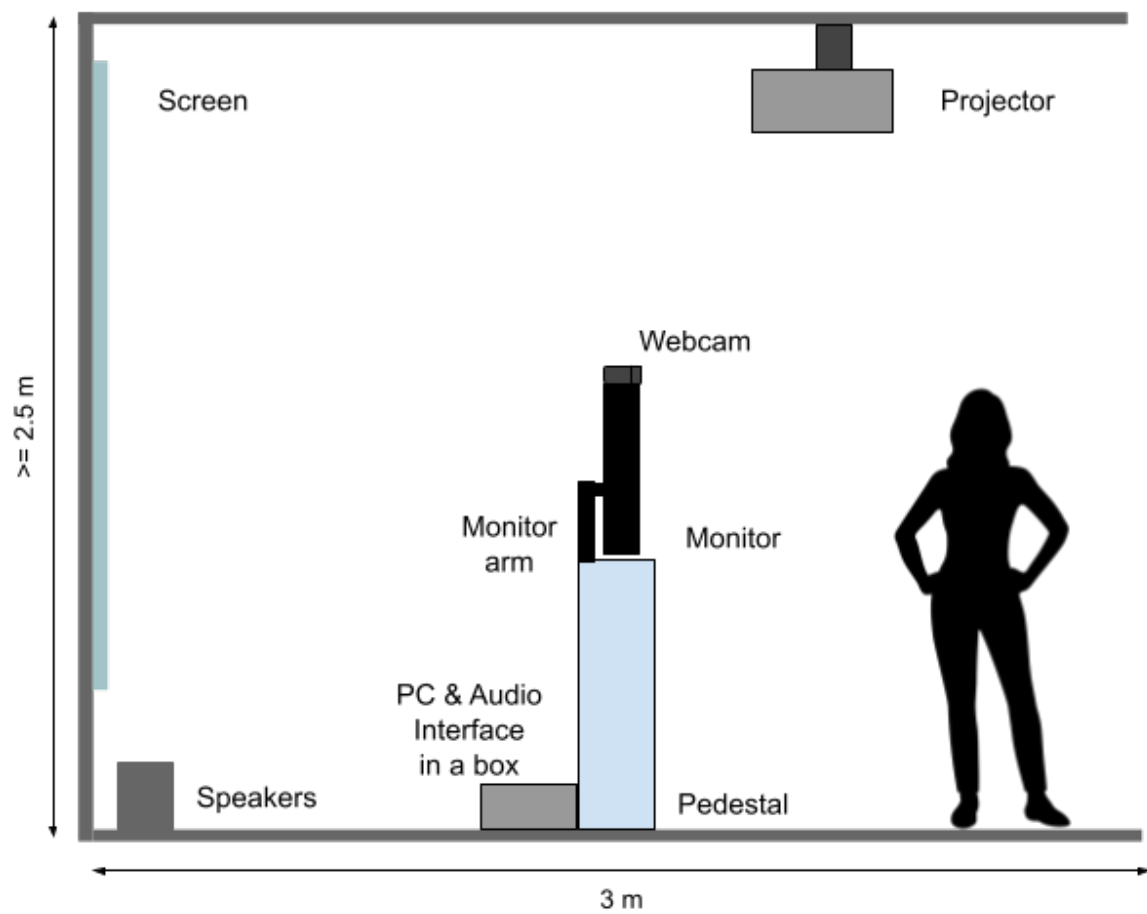


Figure 2: YouTube Mirror floor plan

6. REFERENCES

- [1] P. Covington, J. Adams, and E. Sargin. Deep Neural Networks for YouTube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 191–198, 2016.
- [2] Z. Zhao, L. Hong, L. Wei, J. Chen, A. Nath, S. Andrews, A. Kumthekar, M. Sathiamoorthy, X. Yi, and E. Chi. Recommending what video to watch next: a multitask ranking system. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 43–51. Association for Computing Machinery, 2019.