Real-time Projection of Lip Animation onto Face Masks using OmniProcam

Luna Takagi
Tokyo Institute of Technology
Tokyo, Japan
takagi.l.aa@m.titech.ac.jp

Toshiki Sato
Japan Advanced Institute of Science and Technology
Ishikawa, Japan
tsato@jaist.ac.jp

Shio Miyafuji
Tokyo Institute of Technology
Tokyo, Japan
miyafuji@c.titech.ac.jp

Hideki Koike
Tokyo Institute of Technology
Tokyo, Japan
koike@c.titech.ac.jp

Figure 1: An OmniProcam System which enables 360-degree horizontal projection. (Left) A configuration diagram of the OmniProcam System. (Middle) An overview of the tabletop system in which the OmniProcam is embedded. (Right) Images are projected on multiple screens at arbitrary positions in 3D space.

ABSTRACT

This paper describes an OmniProcam system, which enables 360 degree horizontal projection by a fisheye lens with a coaxial procam in which the optical axes of the camera and projector are exactly matched. Combined with 2D marker recognition, the OmniProcam can display images onto screens at arbitrary positions in 3D space. As an example application, we developed a system which projects lip animation onto the user's face masks for better communication at the physical meeting in current COVID-19 situation. The system recognizes the user's speech, generates the lip animation using Lipsync, and projects the animation onto the user's face masks.

CCS CONCEPTS

• **Human-centered computing → Displays and images**; *Mixed / augmented reality*.

KEYWORDS

projector-camera system, lipsync, spatial augmented reality, 3D display

SIGGRAPH '21 Posters, August 09-13, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8371-4/21/08.
https://doi.org/10.1145/3450618.3469171

ACM Reference Format:
Luna Takagi, Toshiki Sato, Shio Miyafuji, and Hideki Koike. 2021. Real-time Projection of Lip Animation onto Face Masks using OmniProcam. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '21 Posters), August 09-13, 2021.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3450618.3469171

1 INTRODUCTION

Projector-camera systems (procams) are often used in spatial augmented reality researches. A procam is a combination of a camera as an input device to recognize the real world and a projector as an output device to project images onto the real world. However, since most conventional procams have narrow projection angle and fixed projection direction, it is required to use multiple procams to cover larger area such as a room. For example, RoomAlive [Jones et al. 2014] and MeetAlive [Fender et al. 2017] used multiple procam units to project images to the entire room. Since they are using multiple projectors and cameras, they require calibration of a camera and a projector as well as that of multiple procams.

In this study, we developed an OmniProcam system, which enables 360-degree horizontal projection with a single procam unit by using a fisheye lens and a coaxial procam, in which the optical axes of the camera and projector is exactly matched. The OmniProcam system can capture a 360-degree scene in the horizontal direction and project images onto a screen placed at an arbitrary position in 3D space by combining 2D marker recognition.

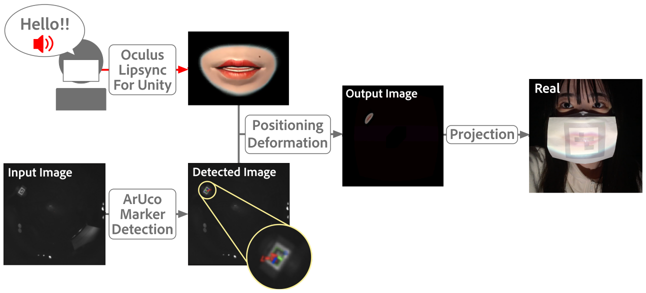

Figure 2: A processing flow of the real-time projection of lip animation onto face masks.

2 OMNIPROCAM SYSTEM

Fig. 1(left) illustrates a configuration diagram of OmniProcam we developed and Fig. 1(middle) shows a tabletop system in which the OmniProcam is embedded. As is indicated by blue arrows in Fig. 1(left), the 360-degree infrared image around the OmniProcam passes through a fisheye lens (Opteka, OPT-0.2X-37), is reflected by a half-mirror (Edmund Optics B/S 3T NIR 20R/80T 75x75), is captured by a camera (FLIR Grasshopper3, 90FPS), and is input to a PC (Lenovo ThinkPad E560). Infrared LED lights (uxcell 30 LED) are placed around the fisheye lens. The input image is sent to Unity through FlyCapture SDK. Unity uses OpenCV for Unity's ArUco Marker Detection [1] to calculate the position and orientation of 2D markers in 3D space. The 2D marker is made of retroreflective material. As is indicated by yellow arrows in Fig. 1(left), the output image generated by Unity is transformed so that it is projected to the marker position, and then output from a projector (HITACHI LP-WU6500J, 5000lm). The image from the projector passes through the half-mirror and fisheye lens again, and is projected to 360-degree horizontal direction. Fig. 1(right) shows an example of image projection on multiple screens placed around the table.

3 APPLICATION: PROJECTION OF LIP ANIMATION ONTO FACE MASKS

The developed OmniProcam can recognize markers at arbitrary locations in 3D space and project corresponding images onto them. As one application using this feature, we projected a lip image onto a mask worn by the user. Due to the influence of COVID-19, online meetings are now common. In such online meetings, if each participant is participating from his or her own home, there is no need to wear a mask, and the entire face image can be transmitted. On the other hand, it is also common to see hybrids in which some people physically gather in a room to participate in an online conference. In such physical rooms, each participant needs to wear a mask to reduce the risk of infection. However, it has been reported that masks can interfere with communication because they cover most of the face, making it difficult to see facial expressions [Claus-Christian 2020]. As an example of OmniProcam applications, we developed a prototype application that projects an animated image of lips onto a mask worn by the user. Fig. 2 shows the processing flow of the system. A 2D marker made of a retroreflective material is attached to the mask, and when the participant wearing the mask speaks, the speech is converted into an animation of lips using Oculus Lipsync for Unity [2]. The image is projected accurately onto the mask using OmniProcam as shown in the most right photo in Fig. 2.

4 DISCUSSION

Our current implementation uses infrared light and an infrared camera for easy recognition of 2D markers. However, since the infrared light is reflected at the objects near the infrared LED, it is difficult to detect markers far from the system. It is also difficult to recognize objects without markers in the real world. Therefore, we are planning to use an RGB color camera instead of an infrared camera in our next implementation. When the image is in color, it is possible to use existing deep-learning based recognition systems such as object/face/hand recognition, and so on.

Although we used a projector which has relatively high brightness (5000lm) and high resolution (1920x1200), the brightness and resolution of the projected image on the screen is low since the image is projected to omni-direction through the fisheye lens. In a daytime environment, it becomes difficult to see the projected contents, which limits the range of application. We believe the development of devices will solve this problem.

Another issue to be concerned is the focus of the projected image. Most of the existing projectors have a fixed focus of projection. This issue could be solved by using focus-free laser projectors.

5 CONCLUSIONS

This paper proposed a coaxial 360-degree procam system, called OmniProcam. By using a fisheye lens, the system enables omni-directional projection by a single procam unit. Furthermore, by matching the optical axes of the camera and projector, the calibration of the camera and projector becomes much easier. OmniProcam recognizes 2D markers in 3D space and projects images based on the information obtained from the 2D markers. As an application of OmniProcam, we developed a real-time projection of lip animation onto face masks for better communication in physical meeting during COVID-19 situation. Although we have not yet conducted any formal user studies, the projected lip animation helps other users to understand the speaker's emotion.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI (#JP20H04221).

REFERENCES

Carbon Claus-Christian. 2020. *Wearing Face Masks Strongly Confuses Counterparts in Reading Emotions.* Vol. 11. 2526 pages.

Andreas Rene Fender, Hrvoje Benko, and Andy Wilson. 2017. MeetAlive: Room-Scale Omni-Directional Display System for Multi-User Content and Control Sharing. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (Brighton, United Kingdom) *(ISS '17)*. Association for Computing Machinery, New York, NY, USA, 106–115. https://doi.org/10.1145/3132272.3134117

Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 637–644. https://doi.org/10.1145/2642918.2647383

[1]https://docs.opencv.org/master/d5/dae/tutorial_aruco_detection.html

[2]https://developer.oculus.com/documentation/unity/audio-ovrlipsync-unity/