International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2016)
D. Reiners, D. Iwai, and F. Steinicke (Editors)

Superimposing Questioner on Presentation Screen Using Microphone with Whole-Sky Camera

Yuki Kobayashi[1] and Kenji Funahashi[1]

[1]Nagoya Institute of Technology

Abstract

Academic conferences are usually held at a large hall. It is difficult that audiences watch the facial expression and gesture of a presenter there. Sometimes the presenter is shot by video camera and projected onto another screen, but it involves costs. At the same time someone usually asks a question there. However conference site is large as mentioned above, it is also difficult that audiences watch the questioner. Of course a conference staff can shoot the questioner and projects onto another screen, but additional manpower is required. Although it is not easy to install a camera at a conference site, a questioner generally uses a microphone. Then we consider projecting a questioner onto a screen using the camera built-in a microphone. In this paper we describe the pilot system to verify the effectiveness of this idea.

Categories and Subject Descriptors (according to ACM CCS): H.4.3 [Computer Graphics]: INFORMATION SYSTEMS APPLICATIONS/Communications Applications—Computer conferencing, teleconferencing, and videoconferencing

1. Introduction

Academic conferences are usually held at a large hall. It is difficult that audiences watch the facial expression and gesture of a presenter there. Sometimes the presenter is shot by video camera and projected onto another screen, but it involves costs. We already have developed a presentation support system which superimposes a presenter onto the screen displayed a presentation slide [FN14]. At the same time someone usually asks a question there. However conference site is large as mentioned above, it is also difficult that audiences watch the questioner. Of course a conference staff can shoot the questioner and projects onto another screen, but additional manpower is required. There are related studies for this problem. One is that a fisheye camera and a lot of microphone are set on the ceiling and a speaker position is estimated using microphone levels [SKY*04]. Another is that infrared ray emitter is installed in a microphone and a camera is paned and zoomed for a speaker using the infrared ray irradiation [SMK08]. Although it is not easy to install a camera at a conference site, a questioner generally uses a microphone. Then we consider projecting a questioner onto a screen using the camera built-in a microphone. In this paper we describe the pilot system to verify the effectiveness of this idea.

2. Extract Questioner with Whole-Sky Camera

You fundamentally use a microphone near the face, and in unspecified direction. So we tried to install two fish-eye cameras on a microphone, as one is back to back with another one. In this study we used RICOH THETA S that has two fish-eye cameras as a whole-sky camera, and has a microphone for videography. Figure 1 shows the example image that is taken by the camera. After converting one hemisphere image of figure 1 to figure 2, a person is extracted by face recognition. We use following equations for the conversion;

$$x = \frac{r(\frac{X-w}{2})}{\sqrt{D^2+(\frac{X-w}{2})^2+(\frac{Y-h}{2})^2}} + \frac{w}{2}, \tag{1}$$

$$y = \frac{r(\frac{Y-h}{2})}{\sqrt{D^2+(\frac{X-w}{2})^2+(\frac{Y-h}{2})^2}} + \frac{h}{2}, \tag{2}$$

where x and y are the coordinate of the pixel before conversion, X and Y are the coordinate of the pixel after conversion, w and h are the image width and height, D is the distance from the lens to the projection surface, and r is the radius ($r = w/2 = h/2$). Incidentally the questioner is not always in only one hemisphere image. So any directional hemisphere image is generated from whole spherical image. First the axes x, y and z are defined as shown in Figure 3. We describe y axis rotation by 90 degrees for example. The left semicircle of left hemisphere and the right semicircle of right hemisphere are integrated to one hemisphere as a circle (Figure 4). The RGB value at (y, z) in Figure 4 is calculated by the RGB value at (x, y) in Figure 3, where assuming the point (x, y, z) is projected by equidistant projection method. This is proceeded by 30 degree

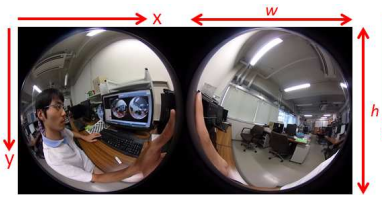

Figure 1: *Captured image*

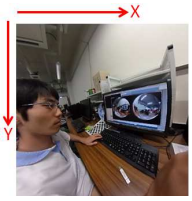

Figure 2: *Converted image*

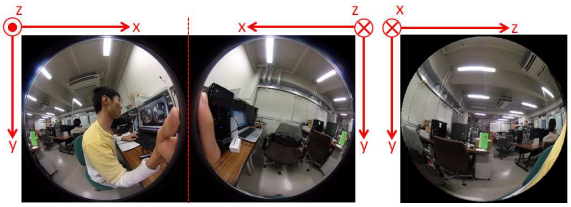

Figure 3: *Captured image (2)*

Figure 4: *Rotated image*

Figure 5: *Result (0 degree rotation)*

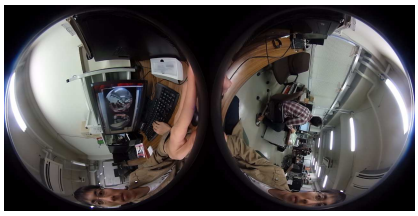

Figure 6: *Result (horizontal 90 degrees)*

Figure 7: *Result (vertical 60 degrees)*

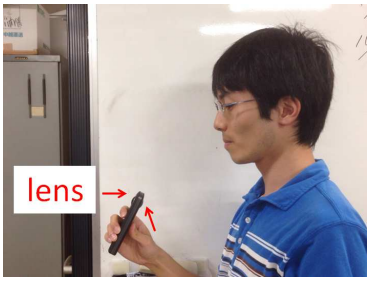

Figure 8: *Situation of figure 7*

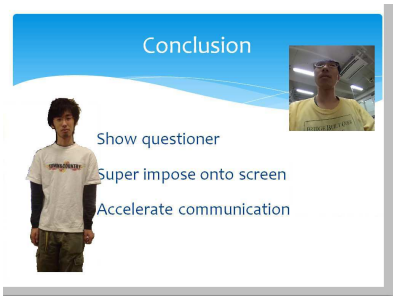

Figure 9: *Screen image*

step for a person extraction. When some people are extracted, the largest face area is determined as the questioner.

3. Experiment and Result

We used PC and whole-sky camera; RICOH THETA S to develop the system we proposed. It should be noted that the camera was connected to PC with USB cable for the experiment. Figure 5 shows the captured image and converted image at the situation where one lens of the camera was just in front of the face. Figure 6 shows the situation where the microphone (camera) was just upright and in front of the face, but two lenses were oriented 90 degrees from the face. Figure 7 shows the situation as shown in Figure 8; the microphone was tilted, and a part of head was shot by one lens and other part of face was shot by another lens. The presentation screen image is shown in Figure 9.

4. Conclusion

In this paper we proposed new approach to show a questioner to audiences. The questioner uses camera-attached microphone and is superimposed onto a presentation screen. It is expected to accelerated their communication since the presenter can see the questioner, and also audiences understanding. As future tasks, we should have an experiment in the large hall to verify the effect of this system. The microphone should be also configured with wireless network. We would like to estimate the front face image from the face that is looked up.

References

[FN14] FUNAHASHI K., NAKAE Y.: Getting yourself superimposed on a presentation screen. *Proceedings of the 2nd ACM symposium on Spatial user interaction* (Sep 2014), 138–138. 1

[SKY*04] SATOSHI N., KAZUHIDE H., YOSHINARI K., KOH K., MICHIHIKO M.: Audio and visual information integration for speaker's localization in automatic shooting of lecture (in japanese). *IEEJ Trans. EIS 124*, 3 (Jun 2004), 729–739. 1

[SMK08] SATOSHI I., MASAKI F., KAZUYA T.: Tracking questioner with infrared-ray-emitting microphone (in japanese). *IPSJ journal 49*, 1 (Jan 2008), 141–149. 1