CONCERT VIEWING HEADPHONES

Sawako Miyashita Masatoshi Hamanaka Seunghee Lee
University of Tsukuba
1-1-1 Tennodai Tsukuba Ibaraki 305-8577 Japan
sawako@music.iit.tsukuba.ac.jp

ABSTRACT
Concert Viewing Headphones, which are equipped with a projector, an inclination sensor, and a distance sensor for zoom control, have been developed. These headphones enable a user to view and listen selectively to specific performers in a group performance recorded on film, and they have both image and sound processing functions. The image processing extracts a portion of the image indicated by the user and projects it free of distortion on the front and side walls. The sound processing generates virtual microphones that allow a listener to focus on select performers even in the absence of individual, physical microphones.

Categories and Subject Descriptors
D.3.2 [**Sound and Image processing**]: Miscellaneous;

1. INTRODUCTION
We have developed a device that enables someone viewing a musical event on film to select particular performers to listen to. It is based on a pair of headphones and comprises a projector, an inclination sensor, and a distance sensor in addition to the basic headphones, as shown in Figure 1. This *Concert Viewing Headphones* enable images and sound sources to be viewed and listened to selectively by natural movement of listening sounds. The two sensors detect the orientation of the user's head, and an image corresponding to the orientation is projected on the screen.

If the user cups an ear to hear better, the device zooms in on the image at which the user is looking. This zooming is done on the basis of the distance between the user's hand and ear, as determined by the distance sensor. Moreover, the sounds are remixed in accordance with the distance between the hand and ear. Therefore, a user watching the film of a concert can enjoy the performance of a particular performer by moving his or her head and by cupping an ear.

With our device, the sounds are remixed when the image is changed. In conventional live image media like DVD, the sounds are mixed in advance and cannot be remixed. A previous system called *Sound Scope Headphones* changes the sound mixing on the basis of the user's head direction. One of the aims of the system is

Figure 1: CONCERT VIEWING HEADPHONES.

to master playing skills of instruments. However, the system does not handle images. It was difficult to only listen to a particular part for novices.

2. DESCRIPTION
Concert Viewing Headphones comprise two systems: image processing and audio processing.

2.1 Image Processing System
The image processing system comprises image extraction, keystone distortion correction, and image output, as illustrated in Figure 2. The first step in image extraction is to detect the user's head direction by using the elevation and azimuth measured with the inclination sensor. Next, the gaze orientation, i.e., what the user is looking at, is determined on the basis of the distance measured with the distance sensor.

2.2 Audio Processing System
The audio processing system comprises *imaginary microphone creation, distance calculation,* and *mixing*. Figure 3 shows the system's overview flowchart.

In this study, we used a large number of sound sources created by positioning microphones at many points on the stage while filming with a fisheye-lens camera. Ideally, every performer would have a sound source, but this is unrealistic. To compensate for the missing sound sources, we defined imaginary microphones for every performer without a real one. Thus, the sounds are mixed by calculating the distance between the positions of the imaginary microphones and the center of the image. Therefore, the sounds can be remixed each time the images are changed.

For those performers not in the line of sight, the sound volume is zero. That is, their sound is muted. For performers who are in the line of sight, their sound is emphasized so as to approach the center of the image. We configure some functions that determine whether or not the mixing rate is loud enough to diminish the distance calculated by the calculation part. Each performer's sound is multiplied by the mixing rate, the sounds are added together, and they are output. We prepared five mixing rate functions, as shown in Figure 4.

Concert Viewing Headphones change the image and sound mixing in accordance with the orientation of the user's head.

ACE'10, 17-NOV-2010, Taipei, Taiwan.

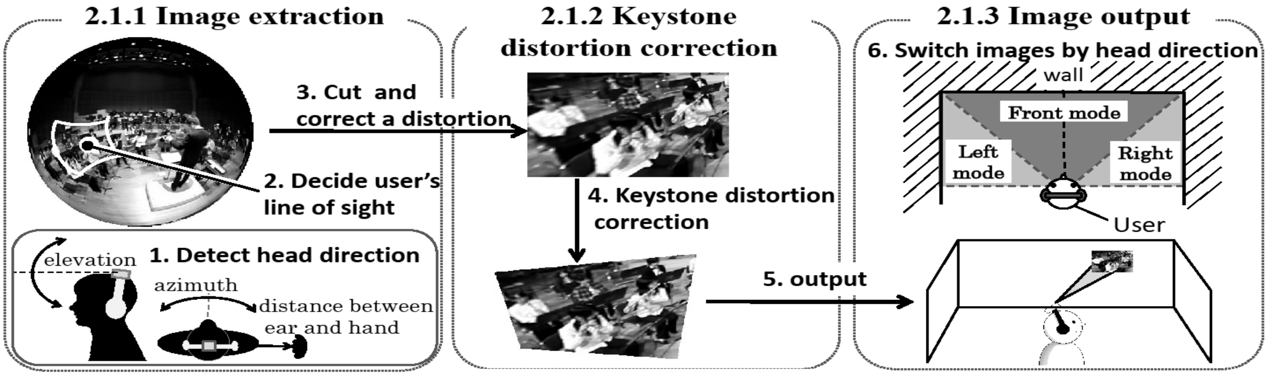

Figure 2: Image processing system

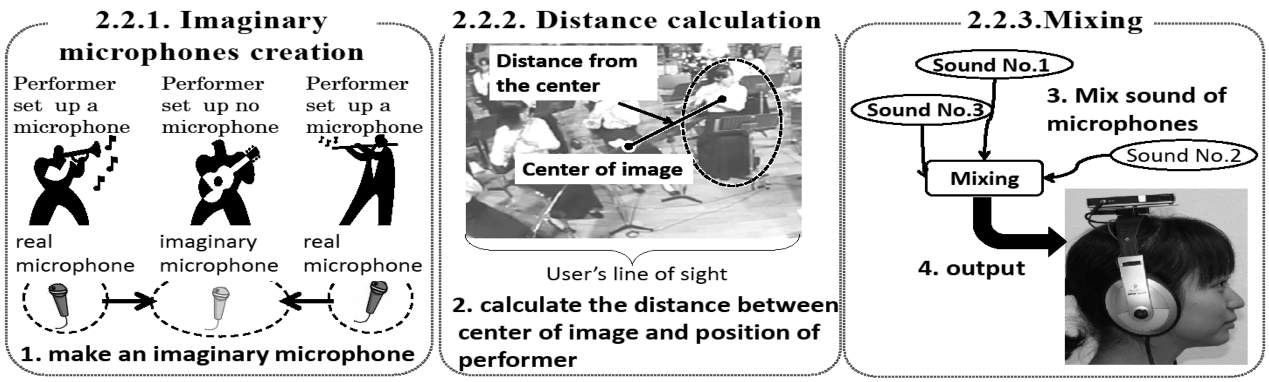

Figure 3: Audio source processing system

When the user operates the zoom function by cupping an ear, performers that are at the edge of the image move out of the zoomed image or move away from the center of the zoomed image. As a result, the sounds of the performers at the center of the line of sight are emphasized.

3. EVALUATION

We evaluated our *Concert Viewing Headphones* by first creating an image and sound source. We filmed and recorded a University of Tsukuba Symphonic Band concert. Microphones were placed on the music stands. The entire stage with all the performers visible was recorded with a fisheye-lens camera. We experimented with the mixing functions to identify the natural correspondence between the image and mixing changes. Specifically, we located a point at which one could hear a 440-Hz sine wave at random positions in the image. Ten evaluation participants attempted to

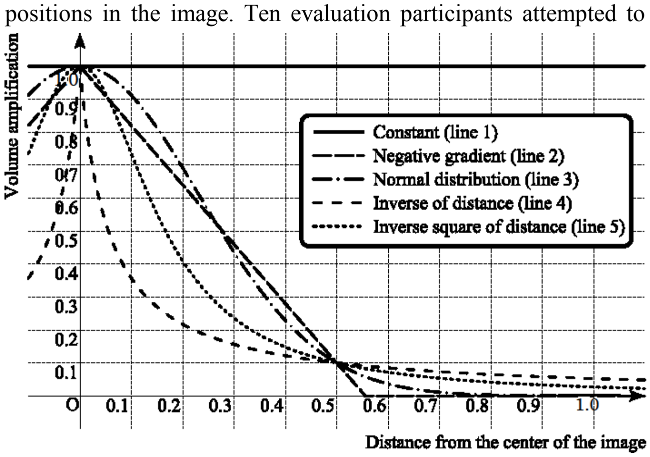

Figure 4: Mixing functions.

locate this point with only their ears. We recorded the time it took to do this. Each time a participant located the point, the sine wave relocated to another point at random. This was repeated five times for each function. We considered that a function where examinees can find a sine wave faster than any other function is the best one. The time was the shortest for the inverse square function and with the zoom function.

4. CONCLUSION

Our Concert Viewing Headphones enable a user to view and listen to specific performers selectively in a group performance recorded on film. Image processing extracts the portion of the image indicated by the user and projects it free of distortion on the front and side walls. Sound processing creates imaginary microphones for these performers so that the user can hear the sound from any performer. Testing with images and sounds captured using a fisheye-lens camera and microphones showed that sound localization was fastest when an inverse square function was used for the sound mixing.

5. REFERENCES

[1] Apple. *QuickTime Tutorials*. http://www.apple.com/quicktime/tutorials/.

[2] Masatoshi Hamanaka, Seunghee Lee. Sound Scope Headphones: Controlling an Audio Mixer through Natural Movement, *Proceedings of the International Computer Music Conference*, pages 155-158. New Orleans, 2006.