

ARMixer: Live Stage Monitor Mixing through Gestural Interaction in Augmented Reality

Weihan Huang

Stephanie Bourgeois

yuiwong@kmd.keio.ac.jp

sbourgeois@kmd.keio.ac.jp

Keio University Graduate School of Media Design
Yokohama, Japan

Yun Suen Pai

Kouta Minamizawa

pai@kmd.keio.ac.jp

kouta@kmd.keio.ac.jp

Keio University Graduate School of Media Design
Yokohama, Japan

ABSTRACT

Existing stage monitor mixing systems are inefficient and cannot accommodate the communication between the musicians and sound engineers. We introduce ARMixer, which allows musicians to perform self-stage monitor mixing through gestures in augmented reality to provide an intuitive mixing experience. We performed two usability tests and found that ARMixer is acceptable to the user and has excellent psychoacoustic intuitiveness in terms of mixing parameter controls by gestures and identifying mixing target.

ACM Reference Format:

Weihan Huang, Stephanie Bourgeois, Yun Suen Pai, and Kouta Minamizawa. 2021. ARMixer: Live Stage Monitor Mixing through Gestural Interaction in Augmented Reality. In *SIGGRAPH Asia 2021 Posters (SA '21)*, December 14–17, 2021, Tokyo, Japan. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3476124.3488632>

1 INTRODUCTION

Audio mixing plays an important role in a music production, and using the visual metaphor of an audio mixing interface (AMI) can help users familiarise with the interface layout, recall mixing knowledge, and complete a mixing task. Compared to a conventional channel strip metaphor which is used with several knobs and faders, the stage metaphor is based on the concept of “deep mixing” to create a virtual stage that presents each audio channel through spheres, providing an intuitive mixing experience [2019]. However, stage metaphor is not widely used in the AMI field because the AMI becomes cluttered and difficult to be manipulated when multi-channels (spheres) need to be processed [2017; 2016]. Additionally, current stage monitor mixing systems have various problems, for instance, the sound from wedge monitors would be “collected” by microphones, once it is too loud, there is a risk of whistling. Also, it is a high cost for small venues to use in-ear monitors. More importantly, the communication efficiency of mixing between musicians and sound engineers is also a challenge [2012]. [2015] proposed a mobile application to allow performers to mix monitors, but switching the interface between the instrument and phone screen may

interfere with the performers’ ongoing task. Therefore, we introduce ARMixer, a system which uses the stage metaphor for its interface, allows musicians to perform an in-situ self-stage monitor mixing through gestures in augmented reality (AR) and provides the intuitive mixing experience.

2 DESIGN & IMPLEMENTATION

Each musician with an instrument on stage can actually be seen as a single audio channel. We designed ARMixer so that the virtual AMI corresponds to the position and placement of other surrounding musicians. Only one audio channel is processed per mix, meaning the interface is simple and the stage metaphor is appropriate to be used in this monitor mixing scenario. Furthermore, compared to a tangible interface that may affect the spectacle of the stage, the interface leveraging AR is personalized because the stage monitor serves the individual musician.

2.1 Hardware & Software

As shown in Fig 1, we simulated the video see-through AR HMD by assembling the *Oculus Quest 2* VR headset with the *Zed Mini* stereo camera. The *Leap Motion* hand controller was also attached in front of the VR headset for compatibility with natural gesture interaction. In addition, we used an audio interface (*Focusrite Scarlett 2i2*) with a microphone or other instruments connected as the real-time audio input and output devices in the program. Finally, all gesture interactions and interface were edited in *Unity*.

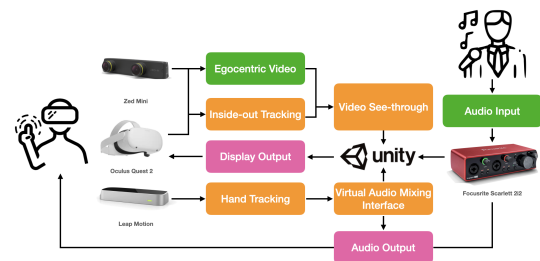


Figure 1: ARMixer System Overview

2.2 User Interface

For the AMI elements, the single audio channel is represented by a blue sphere following the stage metaphor paradigm. ARMixer provides four mixing parameters: volume, pan, reverb, and equalizer (high 12kHz, mid 2.5kHz, and low 80Hz frequencies). As shown

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SA '21, December 14–17, 2021, Tokyo, Japan
© 2021 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-8687-6/21/12.
<https://doi.org/10.1145/3476124.3488632>

