International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2022)
R. Kondo and T. Teo (Editors)

Lyrics onto Cityscape: A Mixed Reality System for Appreciating Lyrics Projected onto a Cityscape While Walking

R. Izumi[1], S. Kodama[1] and T. Takahashi[1,2]

[1]Tokyo Denki University, Japan
[2] ASTRODESIGN, Inc., Japan

Abstract

High-performance mobile devices have made it easy to listen to the music anytime, anywhere that matches listener's mood or to listen to the specific song at the specific place. We propose a new experience that allows people to carry visual expressions matched to the music. In this paper, we propose and develop a mixed reality system that dynamically projects the lyrics of a song users are listening to as animation of lyrics onto the cityscape where they are walking around. By using our system, the listeners will be able to appreciate visual expressions that match both the music and the landscape wherever they wish to enjoy the music.

CCS Concepts

• ***Computing methodologies*** → ***Mixed / augmented reality;***

1. Introduction

In recent years, the act of "carrying the music" has become common to enjoy the music due to the high-performance of mobile devices. For instance, the mobile devices make people easier to listen to the music according to their desires such as "I want to listen to this song that fits my current mood" or "I want to listen to this song in that place".

It has been pointed out that the combination of music and visual expression enhances audiovisual impressions mutually and enriches the musical experience [Iwa92]. With the spread of video sharing sites and SNS, music videos and short movies have been widely distributed. There have been increasing the opportunities to appreciate the music with videos. However, there have been proposed no idea that enables users to take along visual expressions that matches the situation in which they are listening to music, just as they do with music.

As a first step, we propose a new music experience by using mixed reality (MR) technology that allows users to appreciate the music video-like visual expressions which match the cityscape when they walk around according to their desires to "listen to this song at that place". We develop an MR system that dynamically projects and displays animation of the lyrics of a song onto the cityscape. The user can be viewing animation of the lyrics projected on the walls of the buildings when walking outdoors while listening to the music, actually, on the user's HoloLens2 [Hol] display. The lyrics are projected according to the movement of the user's viewpoint, so the user can smoothly continue listening to the music and appreciating the projected lyrics.

2. Related Work

Yamada et al. [YGM20] proposed and developed a prototype of "Words Street," an MR system that superimposes lyrics text on the floor or wall where a user wearing the HoloLens, an optical see-through head mounted display, is looking. The "Words Street" system showed the possibilities of the experience to enjoy lyrics projected onto the user's surrounding environment. However, its operating range was limited to simple environments such as corridors. This is because their system uses SLAM (simultaneous localization and mapping) and ToF (time of flight) sensor of the HoloLens to estimate the user's position and map their surrounding environment.

There is "an AR lyrics music video" that uses augmented reality (AR) to place lyrics in outdoor space [HOW]. The work is a similar expression of our approach but the lyrics are placed manually by the creator.

3. Proposed System

3.1. Processing Flow of Proposed System

We develop an MR system that dynamically projects and displays animation of the lyrics of a song onto the cityscape. Its processing flow is shown in Figure 1.

First, 3D shape models are prepared in advance that approximate complicated shaped buildings to simple polygonal shapes in the area the user walks around. We use the 3D city model data "PLATEAU" provided by the Ministry of Land, Infrastructure, Transport and Tourism of Japan as 3D shape models [PLA]. Next,

DOI: 10.2312/egve.20221288

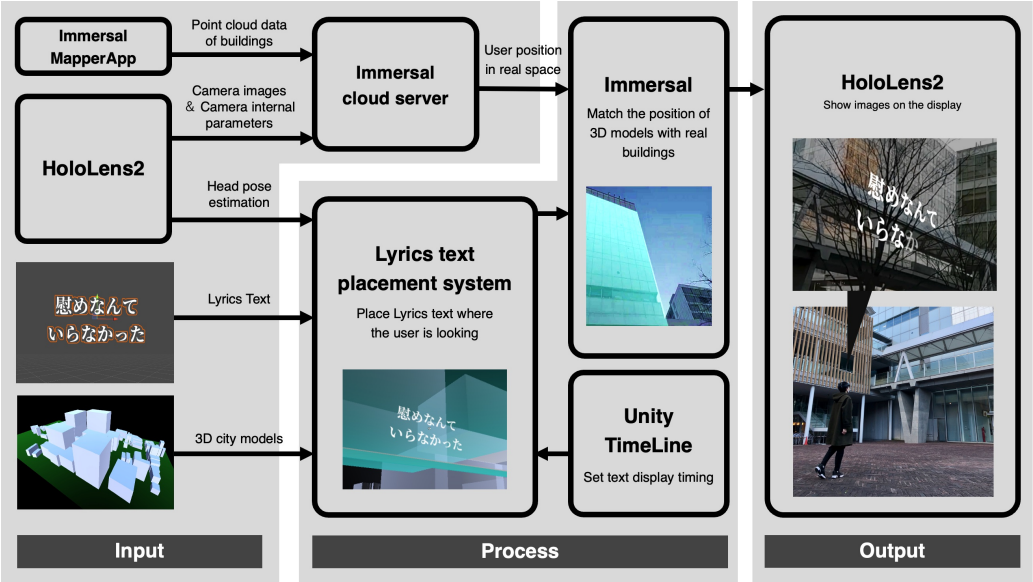

Figure 1: *Overview of our system.*

the view direction of the user is determined by attitude information of HoloLens2. Each ray is traced in the direction of the user's sight, and intersection points of the ray and 3D models of the PLATEAU are calculated. Each piece of lyric information is segmented into several pieces of text information such as phrase. They are converted into several images, then placed at intersection points. Timing and synchronization of lyrics placement and projection is determined in advance using Unity's TimeLine function. Finally, our system matches the projection coordinates with the positions of the building in real space. In order to project the lyrics accurately onto a large-scale building, it is necessary to always match the position of the 3D city model with the position of the building as seen from the user by means of a method that can measure a wider area than using a small depth sensor of HoloLens2. The MR system refers Immersal [Imm] to estimate self-position at 5 seconds intervals based on still images obtained from the HoloLens2 camera.

3.2. Functions to improve the experience

We develop the following two functions for appreciating the visual expressions of lyrics. (1) The lyrics text is projected onto huge building walls not as still images but as animation to enhance the entertainment values. The font size is automatically changed according to the distance between the user and the projected wall for visibility. (2) Projection targets are assumed either of buildings, the sky, or the ground, then the attitude of the projected lyric text is determined. If the lyric text is out of the user's field of view, the lyrics are displayed floating in the center of the field of view so that they cannot be missed.

4. Implementation and Pilot Study of MR System

In order to verify our ideas, we have implemented an MR system on Unity, which projects the lyrics text to huge building walls. The subjects were 8 university students. All subjects were asked to walk around a square (approximately 30m x 25m) freely for about 1 minute and 30 seconds while listening to the music. Lyrics were projected onto the walls of buildings around the square synchronizing to the music. After the verification, we conducted a questionnaire about the impressions and usability of our MR system. As a result, our system was highly rated on the ease of use.

5. Conclusion

When we walk through busy cityscapes and beautiful tourist spots, we wish to enjoy music with visual expressions that match the scenery. We have proposed and developed the first prototype of a mixed reality system that creates and produces such a new music listening experience. The limited number of places that support our system is an issue for the future. We need to increase the number of subjects and evaluate the safe walking of using the system.

References

[Hol] Microsoft Corporation: Microsoft HoloLens 2. https://www.microsoft.com/en-us/hololens, [Online; accessed 2022-09-16]. 1

[HOW] Haru.Robinson: HOWL. https://www.youtube.com/watch?v=3EklwJYeBWM, [Online; accessed 2022-10-31]. 1

[Imm] Immersal: Digital and reality ar ready to merge. https://immersal.com, [Online; accessed 2022-09-16]. 2

[Iwa92] IWAMIYA S.-I.: The effect of the matching of sound and video on the interaction between auditory and visual processing in communication via audio-visual media. *The Journal of the Acoustical Society of Japan 48*, 9 (1992), 649–657. (in Japanese). 1

[PLA] MLIT Japan: PLATEAU. https://www.mlit.go.jp/plateau/, [Online; accessed 2022-09-16]. 1

[YGM20] YAMADA T., GAO Y., MIYATA K.: Words street: A mixed reality system for enjoying lyrics while walking. *Entertainment Computing Symposium 2020* (2020), 1–7. (in Japanese). 1