Conference Poster

Spatial Audio for Human-Object Interactions in Small AR Workspaces

Author(s):
Yang, Jing; Sörös, Gábor

Publication Date:
2018

Permanent Link:
https://doi.org/10.3929/ethz-b-000270325

Originally published in:
http://doi.org/10.1145/3210240.3210811

Rights / License:
In Copyright - Non-Commercial Use Permitted
ABSTRACT
While spatial audio has been an essential component in Virtual Reality, it has been rarely applied to Augmented Reality. We propose a concept and a prototype to enhance human-object interactions in daily life with 3D audio. We augment real objects in a small workspace around the user with spatial audio notifications.

CCS CONCEPTS
• Human-centered computing → Auditory feedback; Sound-based input/output; Ubiquitous and mobile devices;

1 INTRODUCTION
The auditory channel gives us an immediate 360° sense of space, time, and presence. An appropriate 3D soundscape alone enables us to perceive the environment without visual elements. Given this advantage, people usually render authentic spatial audio in VR.

Figure 1: 360° spatial audio notification6. Image credits: SketchUp 3D Warehouse.

We anticipate that a virtual soundscape can also enhance our interaction with real objects. We design a system to attach artificial 3D audio notifications to everyday objects by tracking the user and the objects in real time. Although the objects have no speakers, the user perceives as if the sounds are authentically coming from a particular object. We believe that such a system can help enhance lively and quick 360° interactions in navigation, prompt notification (Figure 1), audio message attachment [1], and general scenarios where visual augmentation is not desired or not feasible.

2 PROTOTYPE SCENARIO
Our prototype scenario is as follows: A user works on a PC at a wooden desk, on which there is a coffee mug and a plate of fruits. We attach virtual audio notifications to the mug and the plate, which creates the impression that they are the sound sources.

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).
MobiSys ’18, June 10–15, 2018, Munich, Germany
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-5720-3/18/06.
https://doi.org/10.1145/3210240.3210811

REFERENCES

3 PROTOTYPE IMPLEMENTATION
The scene is modeled in Unity1. We assume a dominant wooden plane (desk) and another plane (screen) in front of the user. Except for the mug and the plate, any other objects in the environment are assumed to be sufficiently far away so they do not influence sound propagation. We track the user’s head using OpenCV2 with facial landmarks, and track objects by applying the Vuforia AR SDK3 plugin. For simplicity, fiducial markers represent the objects and the user at the workplace. After mapping the real scene into the virtual model, Google’s Resonance Audio SDK4 renders the spatial audio which is delivered to the user via earphones.

In a preliminary experiment, five participants were asked to judge (1) the sound source distance (near at 15cm / far at 30cm) of four pairs of positions, (2) the sound source direction (front / back / left / right) of 16 static equidistant positions. The experiments were conducted without any visual interference. All participants could correctly tell the movement direction. As for the directions, although participants found front / back determination difficult since they did not have any visual clue and were not allowed to move the head as assistance, they could differentiate between left and right at a high accuracy of 86.7%. Plus, when the sound moved around, they could correctly tell the movement direction.

So far, we have demonstrated the feasibility of the idea. In the future, we will model indoor spaces considering boundaries as well as materials. We intend to realize the whole system with only head-mounted camera, smartphone, and earplugs. We aim at rendering spatial audio at arbitrary positions so that both the user and the objects can move freely in the environment.

https://unity3d.com/  
https://enoxsoftware.com/opencvforunity/  
https://developers.google.com/resonance-audio/