



Audio Augmented Reality Using Unity for Marine Tourism

Uipil Chong^(✉)  and Shokhzod Alimardanov 

University of Ulsan, 93 Daehak-ro, Mugeo-dong, Nam-gu, Ulsan, Republic of Korea
upchong@ulsan.ac.kr, alimardanovshohzod@gmail.com

Abstract. The development of marine tourism content is a much-needed challenge at this point. Audio augmented reality (AAR) is a new and fast-growing technology. We assume that AAR will improve the quality of services provided by marine tourism. In this research, the actual potential of AAR in marine tourism was studied by developing an AAR prototype using the Unity engine. We have used whale images with 3D sounds. The prototype can localize several 3D sounds in real-world locations. We conducted a listening test comparing the normal whale audios with the 3D whale audios generated by AAR. 81% of the participants were positive for 3D sounds and whale images. Interviewees felt better engagement in the service. They considered that the technology can be useful for receiving more information easily and making them more physically active. We conclude that the main contributions of AAR may be giving insightful knowledge about marine life in an interesting way and stimulating people to engage in nature closer.

Keywords: Audio augmented reality (AAR) · Unity · 3D sounds · ARCore · Marine tourism

1 Introduction

The tourist attraction is a major point of interest in the marine tourism industry. There are a wide variety of factors that have a direct impact on users' choice. Visitors always have a desire to find the uniqueness of the destination while having fun. Proper navigation and guidance are some of the things that could help to meet those expectations. Current systems in marine tourism rely on professional guides and hardware-dependent technology to satisfy their needs. This causes some other problems such as language barriers and high costs. Although English is the most widespread language in the world, there is always an insufficiency of English speakers in some destinations. Hiring professional guides as many as needed is costly for both consumers and sellers in the industry. The study analyzed the role of AAR as a solution to the problems.

Audio augmented reality (AAR) is a new-rising technology that augments objects with virtually spatialized sounds. The concept of AAR has been proposed a few decades ago but remained less explored in the field of augmented reality (AR) research, of which a majority has focused on visual augmentation [1–3]. With AAR, users can perceive the virtually synthesized sounds by hearing them via normal headphones.

AAR is a voice-oriented part of Augmented Reality (AR). AR resembles a bridge that connects the virtual and real worlds. It makes real life more interesting and interactive by adding extra information. People can see real-size whales wherever they want, mechanics may be given extra information about the current problem as well as safety warnings, education can bring real jungle in the classroom. In fact, we have seen AR's assistance in many areas of our lives [4].

As a result of AR's significant development over the last years, we can find its implementations in various industries. Experiments in construction [5], education [6], gaming, retail, marketing, medicine [7], tourism [8], and automotive industries [9] have noted positive improvements.

The study in reference [10] proposed a synergy between the human body development, computing machines, and neural environment. Their focus was to motivate children toward nature. Pokémon Go, which was developed by Niantic, is an AR product that would encourage exercise, but not in a heavy-handed way [11]. AR games make players engage in nature physically.

Another study [12] investigated AR's impact on sociability, entertainment, and learning. There were two groups of people using either an AR app or non-AR apps. AR app users noted the entertainment value and learning opportunities. They also acknowledged that they were better engaged in the exhibition.

After researching the impact of AR in other industries, we assume that it might also be profitable to create software that makes people more physically active by being better engaged in the service of the marine tourism industry.

We have developed an application that augments several 3D sounds on top of real-world objects. The service provided by our AAR application can be experienced using headphones and smartphones.

This paper is organized as follows. Section 2 presents background and related works. Section 3 explains important 3D sound implementation tools of Unity with C# coding algorithm. It also consists of the prototype development process. Experimental results are given in Sect. 4. Conclusions are presented in Sect. 5.

2 Background and Related Works

2.1 Augmented Reality

Augmented Reality (AR) combines virtual objects and real-world observations. Virtual objects not only come into view as part of the user's environment, but users can also interact with the objects. Typical AR systems comprise technologies, such as computing, motion tracking hardware, and an event registration interface [5].

AR is classified based on various criteria. For example, depending on the kinds of hardware that users use there are stationary AR systems, spatial AR systems, desktop AR systems, mobile AR, head-mounted displays, contact lenses.

Stationary AR systems use powerful cameras and recognize real objects and scenes. They are used in virtual fitting rooms.

Spatial AR systems are also called projection mapping and video mapping. They bring real life-sized virtual objects to real life. Spatial systems are much popular in the

automotive industry. Automakers are practicing spatial AR systems for showing their new models to their clients.

Desktop AR systems take advantage of webcams and project digital data onto real-world objects and display them on the desktop screen. Online stores are implementing desktop AR systems for virtual fitting rooms.

Mobile AR systems are the most popular AR system among others due to ubiquitous mobile devices. Usually, mobile devices receive information from GPS to extract the user's position and show specific data that has a direct connection with user locations. Mobile AR systems play an important role in navigating people in tourism cities.

Head-mounted displays' most popular representative is smart glasses. They can be either video see-through or optical devices depending on the technology behind.

Contact lenses are considered the future of AR since it is still under development.

The next classification depends on the identification of scenes or objects in the real world. There are vision-based AR and location-based AR.

Vision-based AR mainly focuses on the recognition and interpretation of images, objects, and scenes in the reality.

Location-based AR anchors virtual data to certain locations persistently with the help of GPS. In tourism, this persistently localized data helps tourists to find their destinations and collect more information about different facilities.

2.2 3D Sounds and AAR

Adding 3D sound effects with the relative listener's position will make the sound systems more realistic and suitable. A detailed description of 3D sounds and technologies behind are given in reference [13, 14].

Audio Augmented Reality (AAR) enhances the real audio environment by superimposing virtual audios [15]. It's a part of augmented reality that concentrates mainly on 3D sounds. In the research, we use AAR to generate 3D sounds.

2.3 Unity and ARCore

Developed by Unity Technologies, Unity is one of the most popular game engines for both 2D and 3D cross-platform game development. It was launched in 2005. C# and JavaScript are used as the main scripting in Unity. While installing Unity, it is required to have a Unity account. Once Unity is installed, you will be able to access the game engine from the icon [16].

ARCore is Google's platform for creating AR applications for Android and iOS. Its first version was launched in February 2018. It makes use of different capabilities, such as motion tracking, environmental location, and light estimation. Targeted at Android 7.0 and above, not all devices in the market support this technology, although this number has increased greatly since the first SDK version [17].

3 Development of Audio Augmented Reality Prototype

3.1 3D Sound Coding in Unity

The development process begins by installing the `ARFoundation` package in Unity (see Fig. 1). `ARFoundation` is a cross-platform API that helps to create augmented reality for both Android and iOS. Developed by Google and Apple, `ARCore` and `ARKit` are special platforms to build AAR projects on Android and iOS respectively. We can create one AR app and run it on both `ARCore` and `ARKit` with the help of `ARFoundation`. Since the scope of the current study centralizes mainly on Android, the android AR development process will be covered only. However, the main part of the prototype is the same, except for the deployment process on a targeted operating system. The prototype can be implemented on the iPhone without much effort.

The next stage is setting up Unity for android development. We import several components in the Game Engine.

Initially, we import `ARSessionOrigin`. It is the parent for an AR setup. `ARSessionOrigin` contains an `ARCamera` that detects features, such as planes or point clouds, and listens to 3D sounds relative to the position of a phone's Camera.

We have used Resonance Audio SDK to spatialize audios in AR. Created by Google, Resonance Audio gives AAR developers powerful technology for delivering high fidelity spatial audio at scale, to users across the top mobile and desktop platforms. One of its main script files is `ResonanceAudioListener` which is attached to `ARCamera` and listens to the 3D sounds in real-time.

To make the experience interesting and visible, we have imported two 3D whale models and attached the audios.

There is an important component inside the Unity Editor, `AudioSource`. The `AudioSource` manages sound behaviors through the AAR experience. It plays back the sound in either 2D or 3D or in mixed. We import the `AudioSource` component. It comes with several special properties. These properties include enabling sound spatialization tools and minimum/maximum distances which determine audio listening zones.

Inside the `AudioSource` component, there is a property called Spatial Blend. It takes values from 0 to 1 which calculate how much the 3D engine is utilized. Increasing Spatial Blend to 1 means the sound will be played back fully in 3D.

The other properties are min/max distances. Minimum distance specifies how close users and audio source should be to listen to the sound. Reversely, maximum distance is for specifying the borders of the virtual sound waves can reach at most. Users hear the sound if there are in this determined sphere. The unity engine simulates the sound to give the feeling of real audios.

Degrees of fading audios or increasing the volume are regulated by Rolloff Mode in the game engine. We applied linear rolloff mode. The further away we get from the sound, the less we hear it.

C# script file was created to manage audios and put them to positions in the sphere. We have defined several variables to assign sound objects as well as to keep track of the number of 3D sounds that a user can localize in the real world. The program inputs values for each variable and then waits for a user to touch the phone screen.

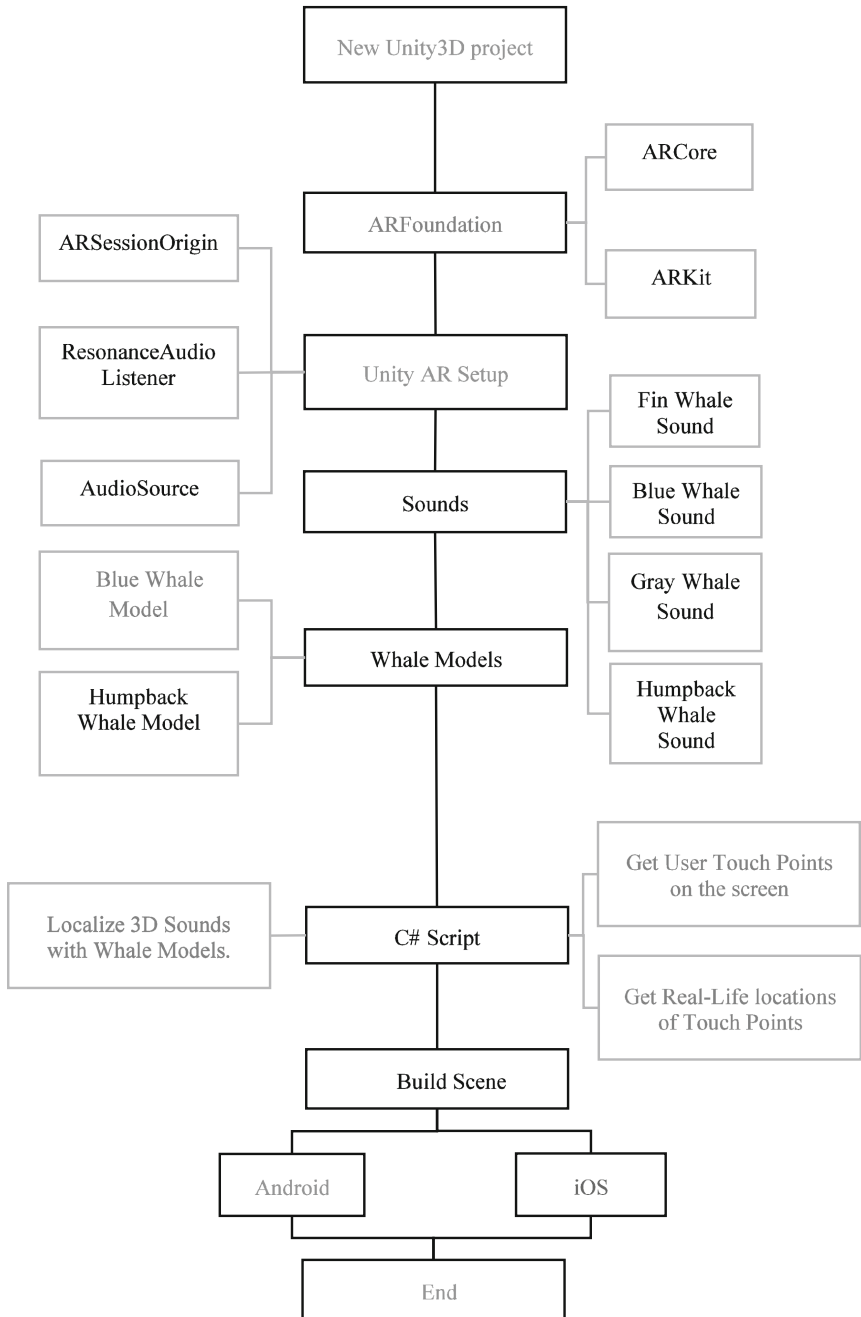


Fig. 1. The development process of AAR apps using Unity.

The Function in Algorithm 1 gives information about the user and phone interactions and checks if a user touches the screen. It stores touch positions, returns `true` when touch inputs exist. In the reverse case, the function returns `false` and gives the default value to the `touchPosition` variable. The `true/false` values are essential to make decisions for 3D sound localization. It should be clarified that the function in Algorithm 1 obtains only the touchpoint on the phone screen, not the location in the real world. Real-world positions (3-dimensional coordinates of the screen touchpoint) are taken by the `Update` function (see Algorithm 2).

```
// Get Vector2 touch positions.
BOOL TryToGetPosition(OUT Vector2 touchPosition)
    // Obtain a screen touch struct.
    IF (Screen touch phase == Beginning of screen touch)
        touchPosition = Screen touch position
        RETURN TRUE
    ENDF

touchPosition = DEFAULT
RETURN FALSE
```

Algorithm 1. Function declaration that obtains information about the user and phone interactions.

Looking more closely inside the `Update` function in Algorithm 2, there are two if-statements. The first ensures nobody is touching the device. Second if-statement casts ray and saves hit information.

```
VOID UPDATE ()

IF (!TryToGetPosition(OUT Vector2 touchPosition))
    RETURN
ENDIF

IF (raycastManager.Raycast(touchPosition, hits, TrackableType.PlaneWithinBounds)
    VAR hitPose = hits[0].pose

    IF (Placed 3D sounds < Maximum limit)
        // Instantiate 3D sounds at the position and orientation.
        Instantiate (3DSounds, hitPose.position, hitPose.rotation)
    ENDF
ENDIF
```

Algorithm 2. The `Update` function declaration that places 3D sounds with whale images.

To withdraw the real-world coordinates of the points on a mobile phone screen and implementing the above-stated ray casting methods, we have used the `Raycast()` function. `Raycast()` is a function that gets real-world positions of screen points by casting virtual rays toward real-world objects. Furthermore, it gathers hit information. The hit information contains the obstacles that the virtual ray comes across as it travels. These can be walls, the ground, chairs, etc.

In the nested if-statement in Algorithm 2, the number of placed objects is calculated and checked if it is more than the allowed limit. In the end, the program localizes 3D sounds in world space.

We used the `Instantiate()` function that takes an object, the position and rotation (see Algorithm. 2). The function clones game objects like the duplication in the editor. Duplicated game objects are placed at the location specified by position and rotation arguments.

3.2 Building the AAR Scene on Android

We used an Android phone to build the prototype. Figure 2 demonstrates the full process starting from the creation of the Android developer environment, goes through build settings inside Unity, and ends at deploying the app on a phone.

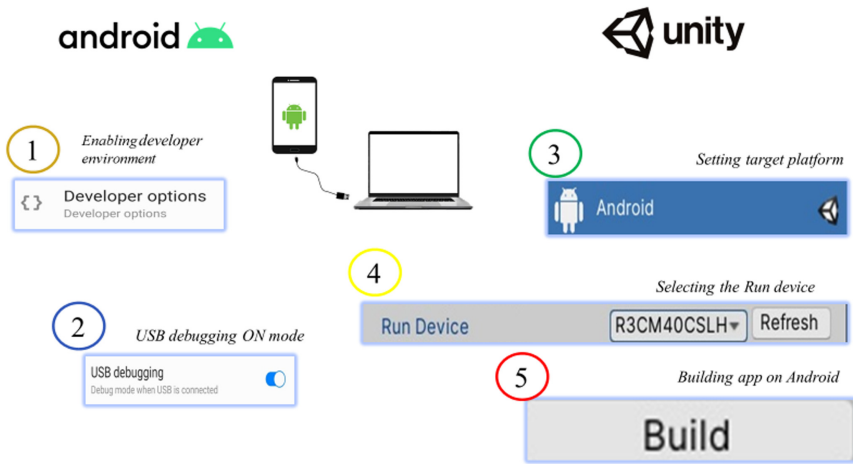


Fig. 2. Building the AAR app on an android phone.

Firstly, we enable developer options on Android. Developer options can be set to visible by going to `Settings > About > Device` and tap seven times on `Build number`. Secondly, we allow USB debugging that is located inside the developer options. We are giving access to a computer for building apps directly on the phone. The process continues with connecting a smartphone to a computer via a USB cable.

The final part is done inside Unity. We build the scene on the target android phone by implementing Unity's special building tools. Figure 2 presents configurations in Unity. The word `Android` at the third step is the target platform. It should be set to the desired type of operating system before building the scene. The fourth step is showing the device that executes the AAR application. Finally, the `build` command starts the actual building process, after which we can practice the AAR experience through the smartphone.

4 Experimental Results

The experiment was conducted successfully on a whale tour ship. We implemented a listening test on August 15, 2020, comparing the normal whale audios with the 3D whale audios. Participants in the experiment listened to 2D sounds at first. Then they listened to 3D sounds.

Thirty-six people from teens to 50 s participated in the experiment, which had a 50/50 female-male ratio. The test results of the experiment are shown in Fig. 3. 81% of people answered positively when they were asked whether 3D sounds felt better than normal sounds. 67% of people gave positive answers to the question about the potential of AAR for improving the quality of the services in marine tourism.

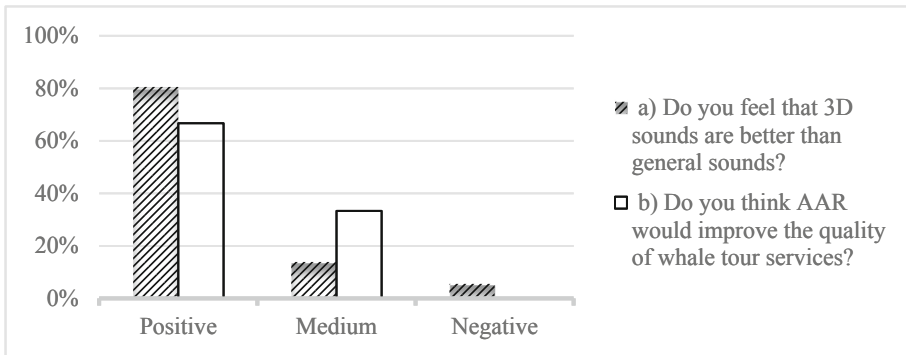


Fig. 3. The survey results on the whale tour ship.

5 Conclusions

We have tested the role and potential of AAR in marine tourism. 81% of participants confirmed positive effects. Using AAR, we may motivate people to interact with nature in a physically active way. We have confronted some limitations of the technology as well. When it comes to the surface that moves continuously on the ship, AAR has shown weak points to persistently locate 3D models.

Acknowledgments. This research is supported by the 2020 National Research Foundation (No. 2017R1D1A3B05030815) of the Korean Government.

References

1. Patil, P.P., Ronald, A.: Cross-platform application development using unity game. *Int. J. Adv. Res. Comput. Sci. Manage. Stud.* **3**(4), 19–27 (2015).
2. Bederson, B., Druin, A.: Computer augmented environments: new places to learn, work, and play. *Adv. Hum. Comput. Interaction*, **5**, 37–66 (1995).

3. Kim, K., Billinghamurst, M., Bruder, G., Duh, H.B., Welch, G.F.: Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR, pp. 2947–2962 . IEEE TVCG (2018).
4. Rovithis, E., Moustakas, N., Floros, A., Vogklis, K.: Audio Legends: Investigating Sonic Interaction in an Augmented Reality Audio Game. *Information MDPI* (2019).
5. Mutis, I., Ambekar, A.: Challenges and Enablers of Augmented Reality Technology for in Situ Walkthrough Applications. *Journal of Information Technology in Construction - ISSN 1874-4753* (2020).
6. Petrov, P.D., Atanasova, T.V.: The Effect of Augmented Reality on Students' Learning Performance in Stem Education. *Information, MDPI. ISSN 2078-2489* (2020).
7. Parekh, P., Pate, S., Patel, N., Shah, M.: Systematic review and meta-analysis of augmented reality in medicine, retail, and games. *Visual Computing for Industry, Biomedicine, and Art* (2020)
8. Williams, M., Kelvin K.K., Yao, J.R.C.: Nurse: Developing an Augmented Reality Tourism App through User-Centered Design. arXiv:2001.11131 (2020).
9. Čujan, Z., Fedorko, G., Mikušová, N.: Application of virtual and augmented reality in automotive. *Open Eng.* **10**(1) (2020).
10. Wirza, R., Nazir, S., Khan, H.U., Garcí'a-Magariño, I., Amin, R.: Augmented reality interface for complex anatomy learning in the central nervous system: a systematic review. *J. Healthcare Eng. Hindawi* (2020)
11. Bueno, S., Gallego, M.D., Noyes, J.: Uses and Gratifications on Augmented Reality Games: An Examination of Pokémon Go. *Applied Science, MDPI* (2020)
12. Savela, N., Oksanen, A., Kaakinen, M., Noreikis, M., Yu, X.: Does Augmented Reality Affect Sociability, Entertainment, and Learning? A Field Experiment. *Applied Science, MDPI* (2020)
13. Yanghan Kim, Youngjin Park: 3D Sound Manipulation: Theory and Application. *ICSV15, Daejeon, Koreaapp*, pp. 1–18 (2008).
14. Usman, M., Kamal, K., Qayyum, R., Akram, S., Mathevan, S.: 3D sound generation using kinect and HRTF. In: *IEEE 2nd ICSIP*, pp. 307–310 (2017)
15. Harma, A., Jakka, J., Tikander, M., Karjalainen, M.: Augmented reality audio for mobile and wearable appliances. *J. Audio Eng. Soc.* (2004).
16. Jing, Y., Gabor, S.: Audio Augmented Museum Experiences using Wearable-Inertial Odometry. *MUM 2019, Pisa, Italy* (2019)
17. Geronazzo, M., et al.: Creating an Audio Story with Interactive Binaural Rendering in Virtual Reality. *Hindawi Wireless Communications and Mobile Computing* (2019).