

Article

Flying a Quadcopter—An Audio Entertainment and Training Game for the Visually Impaired

Silviu Ivascu , Florica Moldoveanu , Alin Moldoveanu , Anca Morar, Ana-Maria Tugulea and Victor Asavei 

Computer Science, University Politehnica, 060042 Bucharest, Romania; florica.moldoveanu@cs.pub.ro (F.M.); alin.moldoveanu@cs.pub.ro (A.M.); anca.morar@cs.pub.ro (A.M.); ana_maria.tugulea@upb.ro (A.-M.T.); victor.asavei@cs.pub.ro (V.A.)

* Correspondence: ivascu.silviu10@gmail.com

Abstract: With the increase in the number of sensory substitution devices, the engineering community is confronted with a new challenge: ensuring user training in safe virtual environments before using these devices in real-life situations. We developed a game that uses an original sonification model, which, although not specific to a certain substitution device, can be an effective means of training for orientation in space based on audio stimuli. Thus, the game is not only a means of entertainment for visually impaired (VI) people but also one of training for the use of assistive devices. The game design and audio design are original contributions by the authors. The sonification model, which is crucial for a game dedicated to visually impaired people, is described in detail, both at the user and the implementation level. For better immersion, special sound design techniques have been used, such as ambisonic recordings and impulse response (IR) recordings. The game has been improved gradually, especially the sonification model, based on users' feedback.

Keywords: visually impaired; audio game; sensory substitution; sound design; sonification



Citation: Ivascu, S.; Moldoveanu, F.; Moldoveanu, A.; Morar, A.; Tugulea, A.-M.; Asavei, V. Flying a Quadcopter—An Audio Entertainment and Training Game for the Visually Impaired. *Appl. Sci.* **2023**, *13*, 6769. <https://doi.org/10.3390/app13116769>

Academic Editor: Masayuki Takada

Received: 23 April 2023

Revised: 30 May 2023

Accepted: 1 June 2023

Published: 2 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Vision represents a dominant sense and plays a very important role in shaping the human quality of life. According to the Report on Vision provided by the World Health Organization (WHO) in 2019 [1], there are approximately 2.2 billion people on Earth who suffer from a form of visual impairment. This work takes into account two categories of visually challenged subjects: the people who have experienced a loss of visual functions early during their lifetime and the ones with a late onset of visual impairment. Since the proposed work requires target users to possess an upper level of perception of 3D spatiality, we consider that people with a late onset of visual impairment will have a better gameplay experience.

From the point of view of education, if vision problems present themselves early in the life of a person, the applicability of classic educational methods and materials is narrowed. This usually leads to limited development of social skills, lower self-esteem, and ultimately to a lower quality of life. From a communication perspective, a visually impaired person does not receive non-verbal cues such as gestures and facial expressions. Half of the total number of visually impaired people can be helped either by addressing their problems with rehabilitation procedures or by the use of specific preventive measures [2].

A number of portable devices have been created with the purpose of helping the visually impaired to move independently through indoor and outdoor spaces, by guiding them using combinations of different sounds and vibrations [3–8]. The effectiveness of these devices depends on the prior training of the VI in order to recognize the sound and vibration patterns of the devices. Very good results have been obtained with training in a virtual environment space, as in the case of the Virtual Training Environment (VTE) [9]

for the Sound of Vision device [6]. The game described in this paper uses an original sonification model, which, although not specific to a certain device, can be an effective means of training for orientation in space based on audio and haptic stimuli. Thus, this game has two purposes: it represents a means of entertainment for visually impaired people, but it also provides a training environment for the use of audio-based assistive devices.

The game described in this paper is based on the authors' experience in developing virtual training environments for the visually impaired, using sonification models. A very important project preceding the authors' current work is the Sound of Vision (SoV) (Bujacz et al., 2016; Bujacz, Kropidlowski, Ivanica, Moldoveanu et al., 2016; Caraiman et al., 2017; Jóhannesson et al., 2016). In order to train the visually impaired for using the SoV sensory substitution device, a Virtual Training Environment (VTE) application has been developed [9]. The VTE focused on the way users perceived different objects from a virtual environment using sonification and haptic models.

Another application developed by the authors, which helped in creating the sonification model of the current game, is presented in [10,11]. The purpose of the application was to train VI users for navigating a labyrinth using a sonification model. The application was tested with the assistance of three blindfolded users for one month. During the preliminary tests, certain aspects of the application needed to be revised, and the tests had to be repeated multiple times. The main aspects that needed revision were the ones regarding the user interaction with the application. The application was designed for both desktop and mobile platforms to maximize availability. The game difficulty increased gradually during the five levels that the user had to complete, as a gamification strategy. The proprietary sonification model of the virtual environment has been developed using the Google Resonance Audio SDK [12]. The application uses the same hardware as the game described in this paper (Figure 1). By using only commercial hardware, the accessibility has widened to address the visually impaired community and the average consumers [13].



Figure 1. Hardware for the mobile application: a. mobile phone; b. headphones; c. gamepad; d. mobile phone VR glasses.

In the VTE application, the serious games were designed with a primary focus on training visually impaired individuals to utilize the SoV device. These earlier games were relatively simple in nature, aiming to provide practical training for device usage. However, in our current game, we have introduced several features that differentiate it from its predecessors. These include a different sonification model focused on navigating with a virtual vehicle in a 3D space, incorporating elements of entertainment; expanded gameplay mechanics such as altitude-based movement; and the introduction of a multiplayer mode.

Compared to the previous game in the series, our new game features a more intricate and nuanced approach to sonification, which enhances the auditory feedback provided to

the player. Additionally, the inclusion of altitude-based movement adds an extra dimension to the gameplay, allowing for more diverse and immersive experiences.

This paper presents the design and implementation of an audio-based game for visually impaired people.

Section 2 contains more sub-sections: related work (Section 2.1); game design, including game architecture (Section 2.2); and sound design (Section 2.3). Section 2.1. presents a review of previous studies that are closest to the one described in the article.

The game design is original. Section 2.2 explores the conceptualization and design considerations of the game in depth, encompassing various aspects such as gameplay mechanics, difficulty levels, characters, and sound effects. Section 2.2.5 provides a high-level overview of the game's architecture within the Unity 3D framework.

Another important original contribution is the sound design, including the conception and implementation of the sonification model, which is crucial for a game dedicated to visually impaired people. The sonification model is described in detail in Section 2.3.1, both at the user and the implementation level. Details are given for both types of sound reflection and reverb simulations implemented in the sonification model: real-time sound reflections and reverberations and baked simulation. Section 2.3.2 gives technical details about the creation of the two characters' voices and Section 2.3.3 about the recording of sounds for baked simulation.

The Results in Section 3 gives details about game testing and game improvements based on users' feedback.

The Discussion in Section 4 presents the findings of the study, incorporating evaluations and user studies, and analyzing the results in relation to the research objectives. Additionally, it outlines potential future research directions.

2. Materials and Methods

The game is exclusively auditory in nature, lacking any visual feedback provided to the users. Within the article, visual aids are shown in figures (such as Figure 2), utilized to convey the concept of spatiality to the readers.

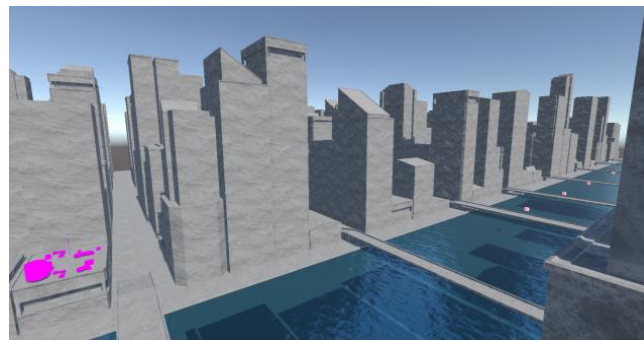


Figure 2. Game space.

2.1. Related Work

Although the literature specific to the field of this research is vast, this section only refers to the works that are closest to the one described in this article, i.e., games dedicated to visually impaired people.

The Nintendo Switch exclusive game “1–2 Switch” [14] is one of the entertainment software solutions that are accessible to the VI community. This game includes 28 mini-games that use mechanics specifically designed for the game console because they use the accelerometer in the Joy-Cons and the infrared sensor mounted inside one of the controllers. The main purpose of this game is entertainment. Thus, it is not relevant for the domain of sensory substitution devices, but it reveals interesting gamification techniques addressable to the VI such as associating real-life gestures with in-game actions.

Another important entertainment application with focus on sound design is the work of Fizek et al., Audio Game Hub [15], which features eight games that can be played on mobile and also on desktop platforms. This project underwent testing with both sighted and visually impaired players to improve its user interface and game mechanics. A total of 29 participants from different gamer groups played various modes of the game during individual face-to-face interviews. The results showed that sighted gamers preferred playing with visual mode off and enjoyed the audio-only gaming experience, while visually impaired gamers found their needs well met and praised the joy-of-use. Sound-centered gameplay mechanics received positive feedback, with an emphasis on the importance of high-quality ambient sounds and voice design for enhancing the overall gaming experience of the Audio Game Hub. RS Games [16] represents another platform with free, multiplayer board, card, and dice games, such as Monopoly, Uno, Rummy, or Bingo, that are accessible to the blind community, through text-to-speech interaction.

Recent advancements, such as the accessibility techniques for VI, can be experienced by players in “The Last of Us: Part II” [17]. These techniques have received considerable attention from both users and the media. However, adaptations for large-scale usage are still challenging to achieve, as most companies do not allocate enough resources to meet the needs of a relatively small part of their clients, such as the VI community.

Nickerson and Hermann [18] created a version of the classic Tic-Tac-Toe game that is accessible to visually impaired players. They incorporated sonic displays to make the game attractive for children. User tests were conducted with visually impaired children in order to evaluate the newly added sounds and the overall audio interface. The authors conducted an informal evaluation of their game with five players, including two musicians and one visually impaired individual, ranging in age from late twenties to mid-fifties. The overall feedback from the evaluation was positive. Players found that the game was playable with practice and enjoyed the engaging rhythms. However, there were some issues identified, such as confusion about the grid’s starting point in the sonification, players using opponents’ slider positions as an advantage, masking of higher-pitched tokens overpowering lower ones, varied playing patterns, and a lack of clarity when transitioning between columns. To address some of these issues, the authors added column delimiters to the graphical interface.

Another VI-accessible audio game is Mudsplat [19]. In this game, an avatar controlled by the player must defeat monsters who try to throw mud at him/her. The avatar has a hose that can spray water on the monsters. The player advances in the game’s levels by defeating a certain number of monsters. To win in the game, the monsters must be quickly located and “shot” before they throw mud. Other audio games by the same authors, accessible to VI, are Tim’s Journey [20] and X-Tune [19]. These games all use proprietary sonification methods. The TiM games showcase the potential of sound-based content and auditory interfaces, offering spatial freedom and encouraging movement for players. The games demonstrate the advantages of audio games, such as covering larger spaces without expensive equipment, using small hardware devices, and developing new input interfaces. Additionally, audio games provide ergonomic benefits, relieving eye strain. Mudsplat, among other TiM games, is nearing completion, and testing with visually impaired children in Sweden, the United Kingdom, and France has shown both their strong interest and their ability to navigate complex interfaces successfully.

Smith and Nayar [21] introduce an audio-based user interface for racing games. The racing auditory display contains two sonification techniques: a sound slider that allows the visually impaired to understand the car’s speed and trajectory, and the turn indicator system, which alerts the player about the properties of upcoming turns. AudioDoom [22] represents a hyperstory (a story that occurs in a 3D acoustic virtual environment) for visually impaired children. By dividing the navigable space into voxels and using 3D aural representations of the environment, AudioDoom enables players to navigate in a set of corridors and interact with virtual objects and characters. Entombed [23] is an audio role playing a rogue-like game for visually impaired. The main user interaction revolves around

audio cues (enabled by a built-in text-to-speech engine) and keyboard commands. Blind Hero [24] is an adaptation of the popular Guitar Hero rhythm game. The authors replaced the visual stimuli from the original game with haptic stimuli through the use of a haptic glove with pager motors attached to the tip of each finger.

Rovithis et al. [25] describe “Audio Legends” as an augmented reality audio game featuring an exploration phase based on auditory navigation, and an action phase, where players hit or block virtual sonic targets through gesture-based user interaction. Enclosing Dark [26] is a virtual reality auditory adventure game that takes advantage of built-in features of VR systems (such as Oculus Rift), spatial audio, and haptic feedback. Among the audio techniques used in this game, we mention narrative for explanations and instructions, spatial audio to simulate a natural environment, echolocation to transmit information about the distance to objects in the environment (for gun firing actions), and passive sonar to assist in the navigation process (sending information about the player’s distance to the walls in the environment).

The majority of downloadable games lack complexity and challenge, lacking advanced gaming techniques to engage players effectively. They also employ basic sound strategies, resulting in diminished interaction, content, and immersion. In-game interaction primarily relies on text-to-speech, which can be cumbersome. Additionally, certain games are quite expensive considering their quality. Many games face compatibility issues on supported platforms and lack portability.

2.2. Game Design

The game has been developed for Windows and Android OS. Both the mobile and PC versions require an Xbox 360 or Xbox Series X/S gamepad and a pair of headphones to facilitate gameplay interaction. The Android version of the application requires a set of Virtual Reality (VR) glasses, such as Google Cardboard. The accelerometer in the Android version of the application is used in order to take into account the rotation of the VI user’s head in computing the view transform in the virtual world.

The game has a simple, user-friendly menu, which allows the user to select one option from entering training mode, starting, or leaving the game. Each menu option is communicated to the user through a vocal message. The sonification model and user interaction have been improved iteratively, based on users’ feedback, as described in Section 2.3.1.

2.2.1. Game Story

The game facilitates the learning of the sonification model in a fun setting. The short story we created has the user as an employee of a courier company. The company has visually impaired employees who pilot virtual vehicles such as quadcopters in order to deliver light objects to virtual people.

In the game, the user receives information from two characters: the control point operator and the vehicle itself. The virtual vehicle has a futuristic type of AI, making it capable of empathy and self-awareness.

The virtual control center operator usually briefs the whole mission to the user at the beginning of the game. For example: “We are living in hard times during this isolation. Just remember, you and others like us are part of the solution, not part of the problem as other citizens are. . . . You need to fly your quadcopter to our warehouse and grab some water filters from there. Someone will be waiting for you there to help you. Go on. I will give you updates when you get there”.

The sentient vehicle gives details about individual tasks that the user should perform (such as delivering packages to people, picking up packages from one point and distributing them across multiple locations, and delivering cargo lost by other employees), helps them drive the vehicle, and informs the overall progress state of the current mission. Examples of such tasks are as follows:

- “You have X packages left to deliver”
- “You collided. Remember to use the brake! Right bumper on the gamepad”

A target is defined as the next place the player must travel to in order to fulfill a task or part of a task. Each target emits a sound that is used by the player in order to discover its position.

The game can be played by one or two players simultaneously. Players can cooperate to fulfill the tasks. For example, the mission is started by Player 1 who is mandated to reach four locations. If another player joins the game, the first player can go to locations one and two, and the second player can go to locations three and four. The multiplayer mechanics, described in Section 2.2.4, are designed to make users collaborate in order to finish the game.

2.2.2. Game Space and Difficulty Levels

The game space is a virtual city composed of multiple blocks, as shown in Figure 2. The user always starts from the middle of the city, above the river that passes through the city. The main reason for this choice in the gameplay design is to provide the players with a simple start without any close obstacles. The locations of packages are scattered throughout the city, between the buildings and at different altitudes.

The game takes advantage of the different altitude levels at which the virtual quadcopter is capable of navigating. To navigate at different altitudes, the sonification model needs to convey this information. This aspect is discussed in Section 2.3.1. The difficulty of each level is determined by the number of targets, the different altitudes where the targets are located, and the positions of the targets in the city. For example, for easier difficulty, the targets can be placed on the ground in places with few obstacles nearby, such as the middle of the street. For a higher level of difficulty, the targets can be placed at varying altitudes and in tight spaces between buildings.

2.2.3. Gameplay

To fully understand the sonification model and vehicle control mode, the players must be conferred gradual learning support. For this purpose, the game offers the training mode, which allows the user to learn how to control the vehicle and the meaning of the different sounds, such as the sound emitted by a target, the sound of confirmation when reaching the target, sonification of an obstacle, and others, as described in Section 2.3.1. Learning to control the virtual vehicle is facilitated by its simple design. It resembles the control mode of real filming drones.

After completing the training, when the users feel confident in their ability to understand the sonification model and vehicle control mode, they can start the game. The scenes in which the user actually uses the virtual vehicle to reach the targets become increasingly difficult. If a user feels the need to re-enter the training mode, they can do so at any time using the option inside the main menu of the application. To return to the main menu, the user presses the A button on the gamepad.

The gamepad (Figure 3) is used to navigate the vehicle through the virtual environment and to change between the listening mode and standard navigation mode. The left joystick is used to change the altitude of the vehicle and to rotate the vehicle clockwise or counterclockwise. The right joystick enables the user to move the vehicle to the front and back or left and right. The right bumper button is used to brake the vehicle and if the user holds the button pressed, the listening mode is activated and remains active while the button is held. The double function of this button is designed to be sure that the user is standing still while the listening mode remains active.



Figure 3. Gamepad button mapping.

While the listening mode is active, the target sounds are no longer occluded by obstacles; they can be heard by the player from any distance and the altitude sound becomes active. This way, the user can obtain information regarding the direction of the next target and the altitude difference between the virtual vehicle and the target.

The mobile version has an extra input that the player can use. As the phone is attached to the user's head, if the user rotates his/her head, the behavior of the vehicle mimics the rotation of the head as if the player is the vehicle direction vector.

The application uses the gamepad's haptic feedback motors to vibrate at different speeds and for different durations to convey the intensity of a collision.

The user navigates through the virtual city in order to deliver packages using the controls described above. An example of game strategy is as follows:

- While standing still, the user can press and hold the brake button (right bumper button) and rotate the vehicle using the left joystick until they can hear a sound. That should be the sound toward the next destination (this sound is not affected by the buildings, as mentioned in Section 2.3.1).
- After the direction toward the next target has been set, the user can navigate forward with the right joystick until an obstacle blocks the pathway or until the target is reached.
- During the navigation, the user can avoid obstacles by steering the virtual vehicle on the left or right side. The vehicle is also capable of raising the altitude so the user can avoid an obstacle by flying over it. After avoiding an obstacle, the user can set the direction again using the first step of the strategy until the target is reached.

This is only an example strategy, not the only way the game can be used, but it is the general strategy that the test subjects used so far.

2.2.4. Multiplayer Mechanics

An important aspect of the game described here is the cooperation between users in completing certain tasks. The multiplayer mode was implemented using the Photon package for Unity. Currently, the application supports two concurrent users: the "main" user, who started the level, and the "secondary" user, the one who joins the multiplayer session. Inside the game logic, if the main user leaves the multiplayer session, the secondary player becomes the main user and continues the mission. The application also provides a single beep as audio feedback for the players when another user joins the multiplayer session.

Users can communicate their positions to each other through short beeping sounds. They have the option to emit the short beeping sound frequently to indicate an urgent event or convey a sense of haste to the other player while emitting the sound sparingly is used solely to communicate their current position when the sound of the virtual vehicle's engines is not sufficient.

An important cooperation element is the gate behavior we created. Across the map, there are gates that have different behaviors depending on the number of players in the session (Figure 4).

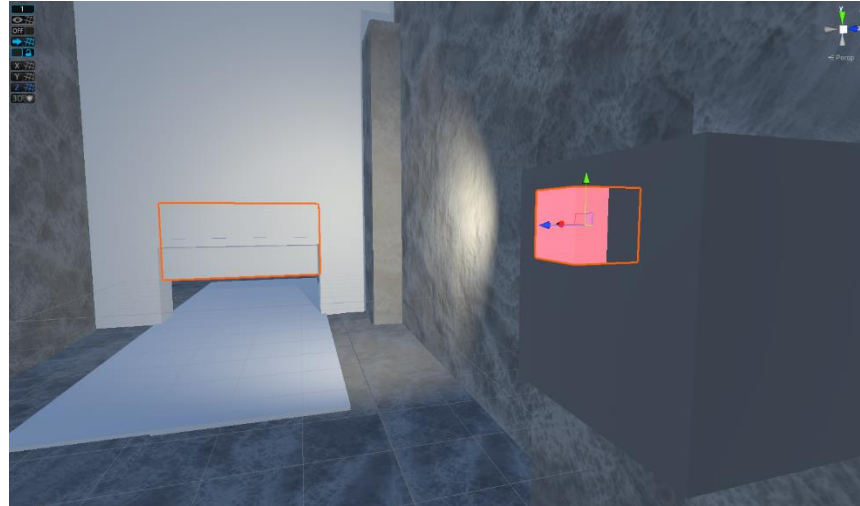


Figure 4. Highlighted in orange: multiplayer gate (Left) switch (Right).

When the session only has one user, in order to pass the gate, the user is required to activate the switch. After the switch is activated, the user has 15 s to pass through the open gate. After this time interval, the gate closes. When the session has two players, the gate behaves differently. In order to pass the gate, one player needs to activate one switch and hold their position in order to keep the switch active while the other player passes through the gate. On the other side of the gate, there is another identical switch that the other user needs to press in order to keep the gate open (Figure 5). At this point, the first player is free to pass through the gate while it remains open.

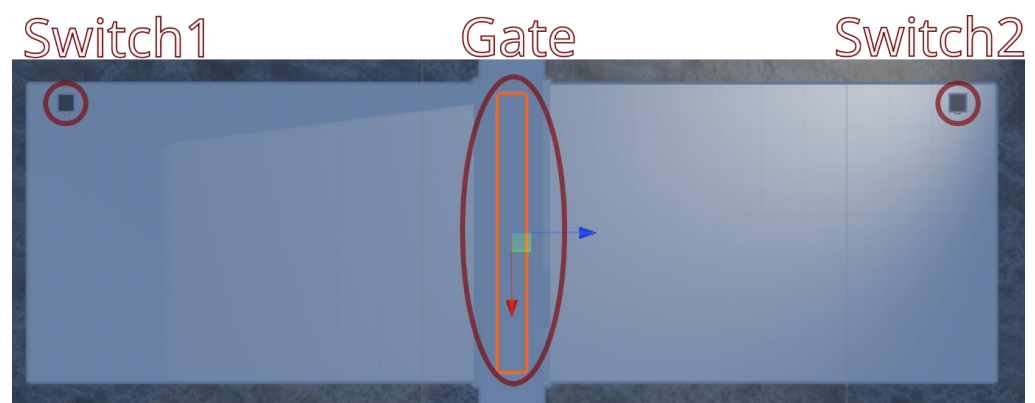


Figure 5. Gate and switches layout seen from above.

2.2.5. Game Architecture

The high-level design of the game, implemented within the Unity 3D framework, is represented in the class diagram from Figure 6. The following classes are central to the game architecture:

GameManager: This class manages the game state and coordinates various aspects of gameplay. It handles tasks such as initiating game sessions, tracking player progress, managing level transitions, and handling game events.

InputMaster: Responsible for managing players' input from both the menu and the gameplay scenes. This class handles user interactions, such as button presses, mouse

movements, or touch gestures. It translates input actions into meaningful commands and communicates them to the appropriate components.

PlayerController: Manages the behavior of the virtual vehicle during gameplay. This class encompasses physics simulations, sound effects, and the response to player input. It handles vehicle movement, collision detection, and updates the corresponding visual and audio feedback. This class also manages the sounds produced by the virtual vehicle, including engine sounds and collision sounds. It also handles the virtual audio sources associated with each character. The vehicle class coordinates the playback of engine sounds based on vehicle speed, acceleration, and user input. It triggers appropriate collision sounds when collisions occur, providing realistic audio feedback. Additionally, it manages virtual audio sources for each character, ensuring that their voices are spatialized correctly within the game environment.

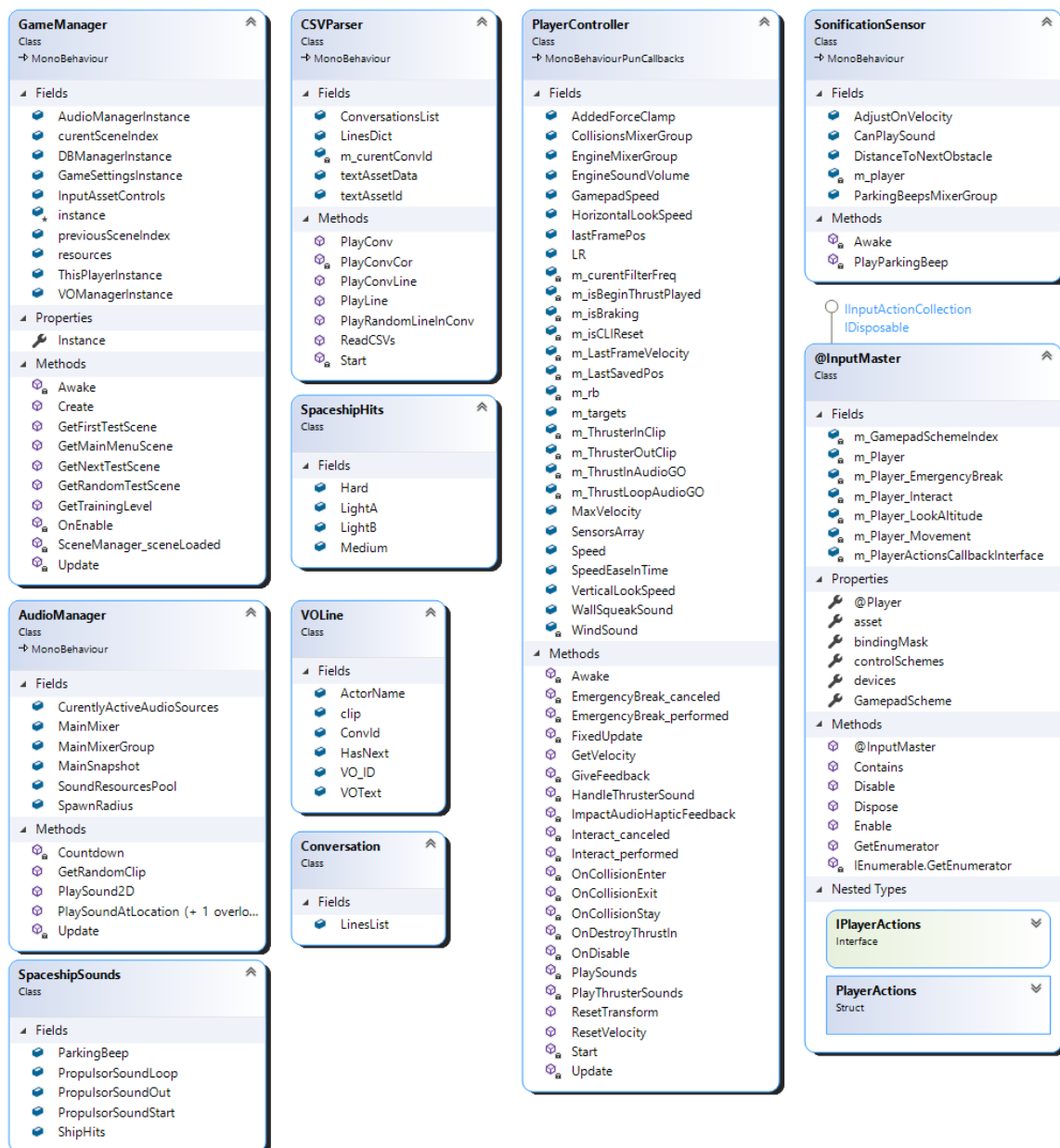


Figure 6. The main classes of the game.

MainMenu: The main menu of the game provides options such as “Start training”, “Start Game”, and “Exit Game”. This class handles user interactions with the main menu,

such as button clicks or menu navigation, and communicates the selected options to the GameManager for further processing.

Sonification: This class implements our proprietary sonification model. It is responsible for translating game events and data into auditory feedback. The Sonification class utilizes various sound synthesis techniques and algorithms to create immersive and informative audio experiences for the players. The class “SonificationSensor” in Figure 6 represents one of the raycasts from the total of five that form the “scanning cone” mentioned in Section 2.3.1.

CSVParser: This class handles the parsing of CSV (Comma-Separated Values) files used for playing the audio files of each character’s voice recording. It provides functionality to read and interpret the voice lines CSV file, extracting relevant information such as voice IDs, texts, associated characters, conversation IDs, and triggering conditions. The CSVParser class enables the GameManager to retrieve and play voice lines based on specific game conditions and events.

2.3. Sound Design

Sound plays an important role in any video game and is crucial for a powerful gaming experience. Even more so for an audio game, where the sound must replace the image.

Good sound design serves a number of functions. Sound design aids in the creation of an immersive experience for the player, giving them the impression that they are actually in the game’s world. A feeling of place and atmosphere can be created through dialogue, music, and sound effects. Good sound design can improve gameplay by delivering crucial auditory cues that guide the player through the game’s world. The sound of footsteps, for instance, can aid a player in finding an enemy, and the sound of a door opening can indicate a new location to explore. Sound design can also have a significant emotional effect on the user. For instance, music can influence a scene’s tone and atmosphere by producing tension, enthusiasm, or grief. A fully immersive and engaging gaming experience is largely dependent on effective game sound design.

We start this section with the sonification model, then describe the technique used to create the voice lines of the two characters. The next section gives details about the methods used to record the different types of sounds needed in the sonification model.

2.3.1. Sonification Model

Similar to other sonification techniques [3,27,28], our sonification model assigns audio information to objects, characters, and events in the game, such as the virtual quadcopter position, the next target’s position, and sounds related to the way the user interacts with the environment, as shown in Table 1 and detailed below.

Table 1. Sonification model.

Sound	Meaning	Rendering	Details	Activated
Obstacle	There is an obstacle in the direction the player is moving.	Short beep, with its play frequency inversely proportional to the distance between the obstacle and the player.	The scanning area is shaped as a cone that varies its angle and scanning distance based on velocity, as detailed below.	Always
Target	Direction of the next target location.	The target emits a click sound from its position.	The sound is spatialized and processed through the Google Resonance HRTE, as explained below.	Always
Altitude	It conveys the difference in altitude between the player and the next target as follows: target above the player, target same altitude as the player, and target below the player.	Target above player: rising beeps. Target below player: descending beeps. Target same altitude as player: four constant beeps.	The frequencies are specified below.	Listening mode

Table 1. Cont.

Sound	Meaning	Rendering	Details	Activated
Collision	Sound generated whenever the player collides with an obstacle.	It has three different sound recordings for three collision intensities: low, medium, and high.		Always
Engine	Conveys the vehicle's response when the player is moving it.	Mimics a real engine sound.	Intensity of the sound is directly proportional to the joystick tilt.	While vehicle is moving
Environment	It does not offer useful information for navigation. Its purpose is to increase immersion.	Recording of birds and other sounds of the city.	Ambisonic recording that takes the user's head rotation into account.	Always
Wind	Vehicle movement. Its purpose is to communicate that the vehicle is still moving even when the player is not tilting the joystick anymore. Without this sound, the player will have the impression that the vehicle instantly stops after releasing the joystick.	The sound increases in intensity as the speed of the vehicle is higher.		While vehicle is moving

The sonification model was implemented using two types of sound reflection and reverb simulations:

1. Real-time sound reflections and reverberations, which were simulated using the features of Google Resonance Audio API. This method was used to generate every sound spatialized in the Unity game engine, such as target direction and the vehicle-generated sounds, for example, collision sounds and engine sounds.
2. Baked simulation. As the sounds from the interior of the vehicle are static, the approach was different. In this case, various samples of IR (impulse response recordings) have been recorded in different rooms and baked to a higher quality reverb with the help of Ableton Live Suite. These sounds are not localized in the Unity game engine.

The obstacle sound. Similar to car parking sensors, the virtual vehicle warns the user if there is an obstacle in front of the player, and consequently, the sound becomes more pronounced if the user is closer to the obstacle. We convey the distance between the user and the obstacle this way.

After testing the VR application with a visually impaired subject who had relevant experience with similar software, a problem with the encoding of the distance between the user and the obstacle was identified.

Figure 7 shows the user moving the virtual vehicle in the direction of the blue arrow. The blue line on the left side of the figure represents the surface of a big obstacle. In this case, the “parking sensor” approach described above tells the user that the obstacle is farther away than it actually is because the collision happens with the left side of the vehicle (red circle in Figure 7) and the sensor offers information about the front of the vehicle.

The solution proposed and tested with the visually impaired user was to create an adaptive scanning cone that scans the environment with the base of the cone oriented toward the direction of the movement vector (Figure 8). The angle of the cone is inversely proportional to the velocity of the virtual vehicle. When the vehicle is at full speed, the angle of the cone is small, and when the vehicle is stationary, the cone is deactivated (the angle of the cone tends to be close to 180°). The user is interested in obstacle positions when the vehicle is moving. When the velocity of the vehicle starts to increase, the scanning cone is activated, and the direction and angle of the cone start to be adjusted as previously explained. The distance at which the cone “detects” obstacles adapts (directly proportional) to velocity. If the velocity of the vehicle increases, so does the scanning distance.

Figure 8 shows this adaptive behavior based on the velocity of the vehicle. The left side of the figure shows a stationary vehicle, and the right side shows a vehicle moving forward at full speed. An important fact not shown in the figure is that the cone also orients itself toward the direction of movement. The vehicle can move upward, downward, left, and right.

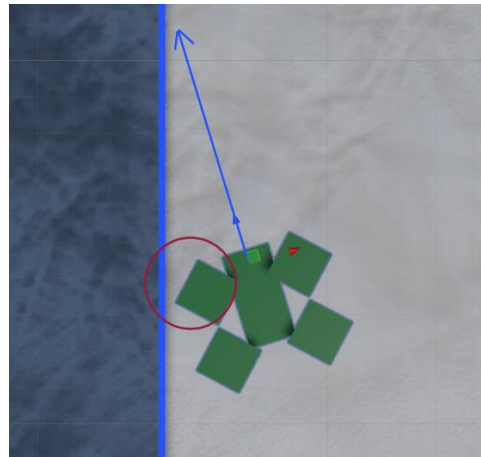


Figure 7. Problem with the encoding of the distance between the vehicle and the obstacle.

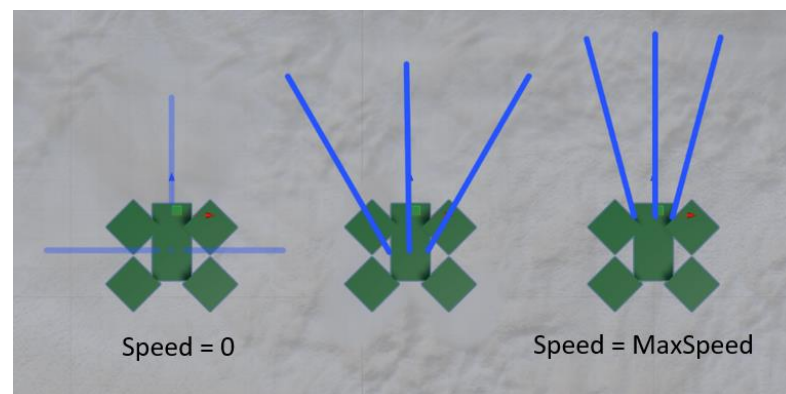


Figure 8. The adaptive scanning cone.

The target sound. The user hears the direction in which the next target is located at all times. This sound has two modes. The default mode is when the sound associated with the target is localized. In this mode, the sound has an attenuation curve; the user can only hear it when they are close to the target. The user can also hear the way the sound reflects off building surfaces. The second mode of the target sound is triggered in listening mode. In this mode, all other sounds except the target sound have a lowpass filter applied at a cutting frequency of 200 Hz. This way, the sounds that are not of interest to the player are muffled. Sound occlusion and the attenuation curve are also removed for the target sound, allowing the user to hear the location of the next target regardless of their position in the world.

The altitude sound. For the altitude guidance, a sound has been generated using the wavetable synthesizer in Ableton, which increases in frequency if the target is above the current position of the player and lowers in frequency if the player is below the target. The frequency of the altitude guidance sound increases in musical steps in order to provide a more pleasant audio experience. For example, the frequency of D3, the D musical note in the third octave on the keyboard, is 146.83 Hz. The sound that raises in pitch uses the following frequencies sequentially: 146.83 Hz (D3), 349.228 Hz (F4), 391.995 Hz (G4), and

1174.66 Hz (D6). By incrementing in these steps, instead of adding 1 Hz for each step, the sound is more melodic as it plays a G7 sus chord.

The collision sound. Different levels of collision are conveyed using recorded sounds of metal sheets being hit by different objects at different intensities. These recorded sounds are further processed using Ableton Live Suite to obtain a final sound sample that is used in the game. The collision intensity is directly proportional to the velocity of the vehicle at the moment of collision.

The engine sound. The engine sound used in the game is created using a recorded sound sample of an anti-wind lighter that is layered over a sound generated by a wavetable synthesizer. The sound generated by the wavetable synthesizer has its base pitch modulated by a lower-frequency oscillator, similar to an internal combustion engine. The sound is played when the user tilts the joystick to move the vehicle in the intended direction.

The environment sound. This sound was obtained by processing a recording of approximately 10 min of outside sound, obtained using specialized equipment for recording ambisonics. The 10 min recording was processed in order to remove unwanted noises such as equipment maneuvering noises or other sounds that are not supposed to be heard in the city at high altitudes. The recordings were made in places with very few people in proximity. After the unwanted sounds were removed, the resulting 3–4 min recording was made to be loopable (repeatable) in order to be used in the game.

The wind sound: This sound was obtained using an omnidirectional microphone and made loopable using Ableton Live. Its volume is directly proportional to the speed of the virtual vehicle.

2.3.2. Character Voices

To record the voice lines of the two characters during the story, Ableton Live 11 was used as the main digital audio workstation (DAW). The lines were read by a single voice actor and the vocal variations were obtained by applying various filters (Figure 9) and popular sound design techniques that are detailed below.

For the control point operator character, we used two settings for the audio processing, depending on the sound source, radio, or verbal communication.

- For verbal, face-to-face communication, we applied basic low-pass and high-pass filters to the raw recording in order to reduce the noise introduced by the recording equipment, a sound compressor to reduce the dynamic range of the recording, and a noise gate, which reduces the signal volume to 0 when the recording amplitude is below a certain level (usually, when the actor makes a long pause in the vocal performance, the default noise of the microphone is heard).
- For radio communication, we applied more drastic low-pass and high-pass filters to simulate ham radio speakers that usually lack fidelity in reality and are able to reproduce a narrower frequency spectrum.

For the vehicle character, we used similar filters. The main difference is that the pitch of the recording has been raised, and then multiple delayed signals of the recording have been added in order to create a more robotic type of voice.

All the voice lines, which are heard through the ham radio speakers, pass through a convolution reverb module in the DAW. This helps simulate the reverb from certain rooms. We recorded sound samples from different rooms and chose the most appropriate sample for the interior of a quadcopter command room. These recordings are used for the convolution reverb to simulate room reverberations, and they are called impulse response (IR) recordings.

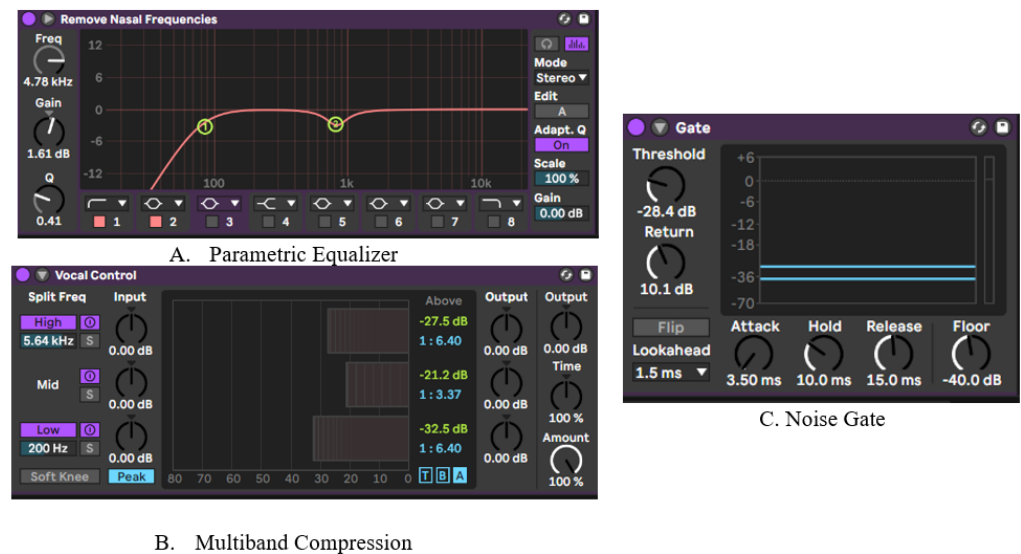


Figure 9. Sound processing chain for vocal processing (signal flows from A to C).

Impulse response recordings are used for convolution reverb to capture the sound characteristics (mainly reverb) of a particular acoustic space, such as a concert hall, recording studio, or cathedral. An impulse response recording is made by playing a very short, sharp sound, such as a balloon pop, hand clap, or starter pistol, in the space being captured and recording the sound that is produced by the space in response to the initial sound. This recording captures the unique acoustic characteristics of the space, including the size, shape, and materials of the walls, ceiling, and floor, as well as any objects or furniture in the space. This recorded impulse response can then be used to create a convolution reverb effect, which applies the captured sound characteristics of the space to any audio signal that is processed through it, making it sound as if the audio was recorded in that space.

Our game includes a tool for playing the audio files of each voice recording. This tool was used to play both predictable voice lines for the main story and dynamic voice lines, such as the quadcopter reactions to the user's progress. For example, the vehicle gives feedback if the user has collisions or when the user progresses through the mission (see Table 2). We created a CSV parser with the following header:

Table 2. Voice lines CSV parser example.

VO_ID	Text	Character	Conversation_ID	Moment of Play	Other Observations
VO_AI_01_04	You have 2 more packages to deliver.	Vehicle	3	When you deliver package #3	This could have a trigger chance
VO_AI_00_13	OW! Use the brake please. Right bumper on the gamepad. CTRL on the keyboard.	Vehicle	29	Collision level 3	

This tool allows the use of recorded voice lines while designing the level. It can play individual voice lines and associate the sound clip with the proper virtual sound source in the game with minimal implementation requirements. Each voice line is part of a conversation. By using the Conversation_ID assigned to each voice line, this tool can also play a series of voice lines one after another, each with its own corresponding virtual sound source.

2.3.3. Sound Recording

Based on the directions from which the sounds are picked up, the microphones can be classified into four main categories, as illustrated in Figure 10. An omnidirectional microphone (Figure 10a) picks up sounds coming from all around its position. Omnidirectional

microphones should not be confused with 360 microphones or ambisonic microphones. They are usually used for field recording or recording the ambiance of a room. Cardioid microphones (Figure 10b) and hypercardioid microphones (Figure 10c) are usually large-diaphragm, condenser studio microphones that are best capable of recording voices and instruments. They have the advantage of recording the sound coming from a specific direction. Bidirectional/“Figure 8” microphones are best for interviews, podcasts, or situations where sounds need to be recorded from two opposite directions. A very important observation is that all of these microphones record mono. Patterns can vary between different use cases, and there are other types of patterns that derive from the main patterns. For example, the unidirectional/shotgun microphone has a variation of the hypercardioid pattern and is characterized by a very small angle of capturing sound. They are used in outdoor situations where the subject is surrounded by unwanted noise. Other patterns include subcardioid and supercardioid.

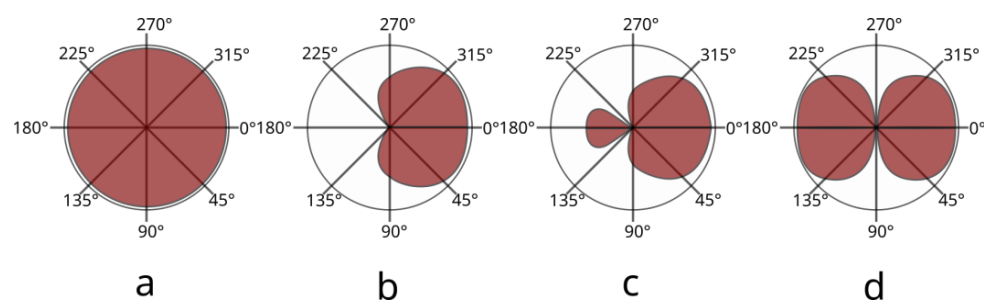


Figure 10. Microphone polar patterns: (a) omnidirectional, (b) cardioid, (c) hypercardioid, and (d) bidirectional/Figure 8).

We used mono recordings (one microphone) to create most of the sounds needed for our sonification model (collision, engine, and wind). The only recordings with multiple microphones were the ambisonic recordings, which were used for the environments sound.

An ambisonic recording, in simple terms, is very similar to a 360° video [29–31]. To record ambisonic audio, three bidirectional/Figure 8 microphones and one omnidirectional microphone can be used. The omnidirectional microphone placed in the center is labeled W (the white sphere in Figure 11). The bidirectional microphones are to be placed in the center (X: red, Y: green, and Z: blue in Figure 11).

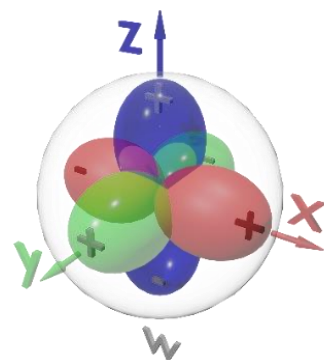


Figure 11. First-order ambisonics.

W is the sound pressure, the equivalent of the “Mid” channel in M/S recordings. The X, Y, and Z components represent the variation in sound pressure on the three axes. The encoder takes two parameters into consideration: θ (azimuth) and φ (elevation). In our case, the azimuth and elevation angles are given by the rotation of the virtual vehicle as well as the rotation of the head of the user if the application is played using a mobile phone. This format is called Ambisonic B [32,33]. In order to encode a mono sound to

an ambisonic format, we can use the following formulae, “SignalValue”, representing the mono signal recorded.

$$X = \text{SignalValue} \times \cos \theta \times \cos \phi \quad (1)$$

$$Y = \text{SignalValue} \times \sin \theta \times \cos \phi \quad (2)$$

$$Z = \text{SignalValue} \times \sin \phi \quad (3)$$

$$W = \text{SignalValue} \times \frac{1}{\sqrt{2}} \quad (4)$$

The encoding described with Equations (1)–(4) was not used in our game due to the fact that we used a specialized microphone for better results. Due to physical limitations, it is not possible to place all the four microphones in the same position; therefore, the sound waves cannot be recorded accurately using the Ambisonic B format. The standard in the industry used to record is the Ambisonic A format. This is conducted using a tetrahedral array of microphones. The minimum number of microphones used is four, as seen in Figure 12. The microphones are named FLU—Front Left Up; FRD—Front Right Down; BRU—Back Right Down; and BRU—Back Right Up. A possible solution for converting from Ambisonics A to Ambisonics B format (proposed for the TetraMic microphone) [29,32,34–37] can use the formulae below, with each variable representing the signal recorded by each microphone (FLU, FRD, BLD, and BRU).

$$X = \text{FLU} + \text{FRD} - \text{BLD} - \text{BRU} \quad (5)$$

$$Y = \text{FLU} - \text{FRD} + \text{BLD} - \text{BRU} \quad (6)$$

$$Z = \text{FLU} - \text{FRD} - \text{BLD} + \text{BRU} \quad (7)$$

$$W = \text{FLU} + \text{FRD} + \text{BLD} + \text{BRU} \quad (8)$$

We recorded ambisonics using an ambisonic microphone [38,39] (Figure 12). Environment recordings are processed using Ableton Live 11 Suite in order to make a loopable recording, and after this process, the recording was converted to Ambisonic B format (Ambix). Due to the fact that Ableton Live is not capable of processing four audio channels, extra steps were needed using third party software (audacity and any HEX editor to modify the .wav header of the recording).



Figure 12. Ambisonic microphone (left); tetrahedral microphone array (Right).

3. Results

The aim of this paper is to present the most significant results of our research, but not exhaustively. In this chapter, we give some details about game testing, users' feedback, and application improvements based on the feedback.

The game is purely auditory; it does not offer visual feedback. That is why we tested the game with a group of three blindfolded subjects and one visually impaired subject. The tests were performed during the period of COVID-19 restrictions; therefore, the number of people involved in the tests was reduced.

The tests lasted approximately four months, with time allocated for modifications based on unsatisfactory test results. The blindfolded user group consisted of a user with no notable experience in using video games and two medium-experienced users. Their ages were 22, 29, and 53, respectively. With the blindfolded users, the application was successfully tested 25 times with each user. After the tests with blindfolded users, the game was tested in three sessions with a visually impaired subject.

The multimodal testing included a form that was completed by each user before using the game in order to receive information on the experience the user had with the technology and a feedback form after each test session. The feedback forms helped improve the quality of the sounds and the level of satisfaction with different aspects of the application, such as the main menu, the training levels, and vehicle control features. They were composed of Likert-scale questions.

3.1. Vehicle Control and Scene Navigation

The first questions targeted the intuitiveness of the vehicle control and scene navigation. The following figure (Figure 13) illustrates the results of the question regarding the difficulty of the vehicle control ("How difficult did you find to control the vehicle?").

75 responses

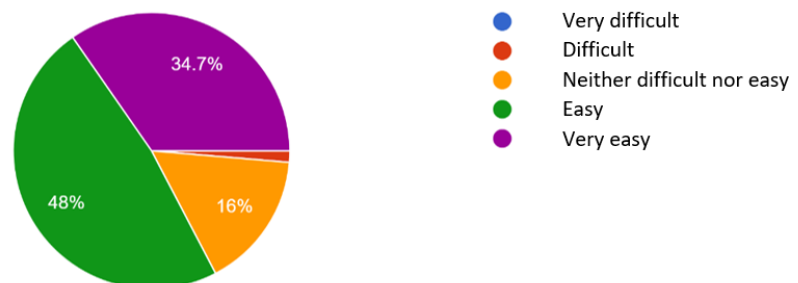


Figure 13. Answers to the difficulty of vehicle control question.

The answers to the question "How difficult did you find to adjust the altitude?" proves (Figure 14) that the altitude adjustment mechanisms are intuitive.

75 responses

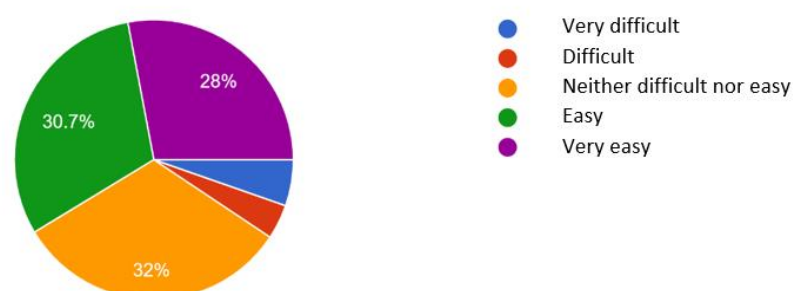


Figure 14. Answers to altitude adjustment question.

The answers to the question “How often have you used the brake?” proves (Figure 15) that the brake is useful.

75 responses

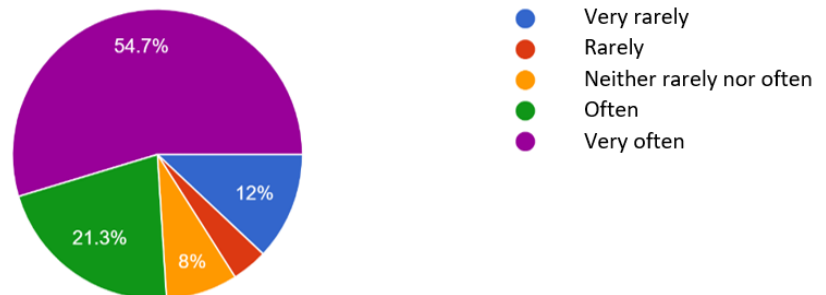


Figure 15. Answers to brake usage question.

3.2. Sonification Model

For the blindfolded users, we asked a general question: if the sonification model is useful (“How do you consider the quality of the sonification model?”). The percentages of the blindfolded users’ responses are presented in Figure 16.

We also asked an open-ended question. The answers to “What changes would you bring to the sonification model? (optional)” were as follows:

- I am starting to understand how the listening mode works if I keep the brake pressed.
- It is hard for me to remember what each sound means.
- It is very useful if someone reminds me what each sound means.
- It seems very useful but it will take time for me to learn.
- I do not get lost anymore if I use the listening mode.
- I would not bring any changes but sometimes I forget the meaning of some sounds.
- For the moment I would not bring any changes.

The answers of the VI user to the question regarding the quality of the sonification are synthesized in Figure 17.

75 responses

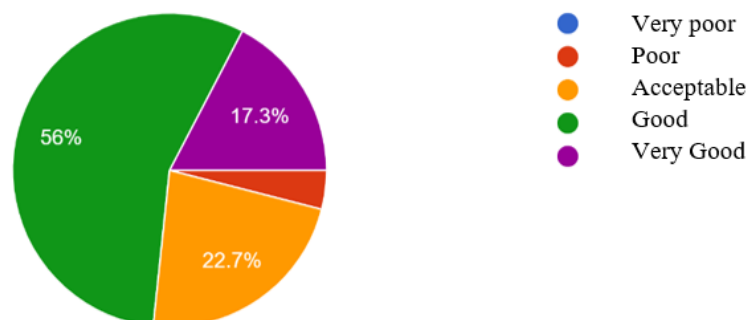


Figure 16. Blindfolded users’ answers to sonification quality question.

The visually impaired respondent was much more interested in the details of our application. As a result, we added more detailed questions about the sonification model. The answers to the questions related to the sonification methods for the direction and distance to the obstacle, the altitude, and the target (including in the listening mode) had a 100% positive response rate. Regarding the optional question “What changes would you bring to the sonification model?”, the VI user responded: “I need information about obstacles on the left, right, up, or down because I can hear obstacles only from in front

of me but I am hitting obstacles with the side of the vehicle". This feedback helped us improve the initial sonification model, as described in Section 2.3.1.

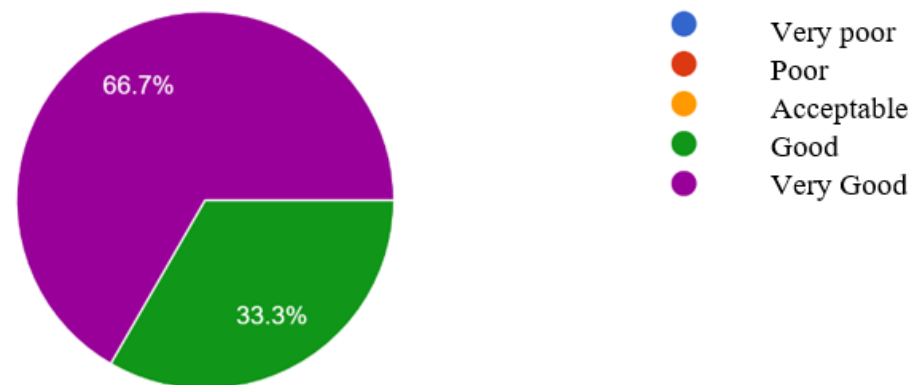


Figure 17. VI user's answers to sonification quality question.

4. Discussion

During the tests, the application was modified to meet the needs of visually impaired and blindfolded participants. This involved adjusting certain parameters, such as the maximum velocity and acceleration of the vehicle, and modifying the sounds' volume. After receiving feedback, the adaptive scanning cone described in Section 2.3.1 was introduced in the obstacle sound implementation.

Certain sounds were modified during the tests by consulting the users. Other sound changes emerged from the open-answer questions regarding explicit reductions in volume or sound changes. For certain sounds, the more prominent frequencies were reduced, and other sounds were replaced altogether. After altering the sounds, we had positive results in reducing early fatigue during testing. There were repeated cases where the user was unable to complete any level, leading to the interruption of tests to address the issues.

The proposed sonification model has the potential to be used for the training of visually impaired children in sound localization, as it can help them understand the location and distance of objects and people based on the sounds they hear.

For this project to become a marketable product, it needs intensive testing for quality assurance with more VI users, especially in the multiplayer mode. For example, some issues that could come up when using the multiplayer play mode could manifest when one of the users disconnects while the other user is holding the gate button pressed. Another issue could happen when a user connects while the other user interacts with the gate. Additionally, disconnecting and reconnecting issues need to be addressed. For example, if the main user disconnects, the second player becomes the "main" user. If the disconnected user reconnects, who should be the main user? These are the first issues that come to mind that need to be fixed before this software can become a product.

The testing scenarios had gradually increasing difficulty, but in the case of the inexperienced user, the need for increased granularity was felt. This would have been possible by increasing the number of scenarios, but taking into account the time required for their development and the time required for retesting, the considered decision was to continue with the current number of scenarios. The game requires additional development, and the authors consider that there is still room for improvement.

Author Contributions: S.I. made the system design, implementation, testing, and writing the first version of the paper. F.M. and A.M. (Alin Moldoveanu) had substantial contributions to the design of the system and evaluation method, paper structuring, reviews, and manuscript improvements. A.M. (Anca Morar), A.-M.T. and V.A. contributed to system improvement and manuscript reviews and improvements. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by Romanian National Authority for Scientific Research (UEFIS-CDI), Program PN-III-P1-1.1-TE-2021-1505 (contract number: TE 147/2022): VIP SpaceNav, Smart buildings for the visually impaired—Platform for low cost, rapid configuration and safe navigation in indoor spaces.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the project being currently under development.

Conflicts of Interest: The authors declare no conflict of interest.

References

- World Health Organization. World Report on Vision. *World Health Organ.* **2019**, *214*, 1–160.
- Available online: <https://www.who.int/En/News-Room/Fact-Sheets/Detail/Blindness-and-Visual-Impairment> (accessed on 16 February 2023).
- Bujacz, M.; Kropidowski, K.; Ivanica, G.; Moldoveanu, A.; Saitis, C.; Csapo, A.; Wersenyi, G.; Spagnol, S.; Johannesson, O.I.; Unnthorsson, R.; et al. Sound of Vision-Spatial Audio Output and Sonification Approaches. In Proceedings of the International Conference on Computers Helping People with Special Needs, Linz, Austria, 13–15 July 2016; pp. 202–209.
- The Sound of Vision Project. Available online: <http://soundofvision.net> (accessed on 27 November 2022).
- Caraiman, S.; Morar, A.; Owczarek, M.; Burlacu, A.; Rzeszotarski, D.; Botezatu, N.; Herghelegiu, P.; Moldoveanu, F.; Strumillo, P.; Moldoveanu, A. Computer Vision for the Visually Impaired: The Sound of Vision System. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops, Venice, Italy, 22–29 October 2017.
- Jóhannesson, Ó.I.; Balan, O.; Unnthorsson, R.; Moldoveanu, A.; Kristjánsson, Á. The Sound of Vision Project: On the Feasibility of an Audio-Haptic Representation of the Environment, for the Visually Impaired. *Brain Sci.* **2016**, *6*, 20. [CrossRef] [PubMed]
- Gao, Y.; Kwok, K.-W.; Chandrawanshi, R.; Squires, A.; Nau, A.C.; Tsz, Z.; Tse, H. Wearable Virtual White Cane: Assistive Technology for Navigating the Visually Impaired. *J. Med. Devices* **2014**, *8*, 020931. [CrossRef]
- Buchs, G.; Haimler, B.; Kerem, M.; Maidenbaum, S.; Braun, L.; Amedi, A. A Self-Training Program for Sensory Substitution Devices. *PLoS ONE* **2021**, *16*, e0250281. [CrossRef] [PubMed]
- Moldoveanu, A.D.B.; Ivascu, S.; Stanica, I.; Dascalu, M.I.; Lupu, R.; Ivanica, G.; Balan, O.; Caraiman, S.; Ungureanu, F.; Moldoveanu, F.; et al. Mastering an Advanced Sensory Substitution Device for Visually Impaired through Innovative Virtual Training. In Proceedings of the IEEE International Conference on Consumer Electronics—Berlin, ICCE-Berlin, Berlin, Germany, 3–6 September 2017; Volume 2017, pp. 120–125.
- Ivascu, S.; Moldoveanu, A.; Moldoveanu, F.; Morar, A.; Balan, O. Virtual Interaction for Visually Impaired and Sighted People. *Elearning Softw. Educ.* **2019**, *1*, 209–214.
- Balan, O.; Moldoveanu, A.; Moldoveanu, F.; Dascalu, M.-I. Audio Games—a Novel Approach towards Effective Learning in the Case of Visually-Impaired People. In Proceedings of the ICERI2014 Proceedings, Seville, Spain, 17–19 November 2014; pp. 6542–6548.
- Resonance Audio. Available online: <https://github.com/resonance-audio> (accessed on 2 March 2023).
- Ibáñez, M.L.; Romero-Hernández, A.; Manero, B.; Guijarro, M. Computer Entertainment Technologies for the Visually Impaired: An Overview. *IJIMAI* **2022**, *7*, 53–68. [CrossRef]
- Available online: <https://www.nintendo.com/store/products/1-2-switch-switch/> (accessed on 20 February 2023).
- Fizek, S.; Woletz, J.; Beksa, J. Playing with Sound and Gesture in Digital Audio Games. In *Mensch und Computer 2015—Workshop*; De Gruyter: Berlin, Germany, 2015; pp. 423–430. ISBN 9783110443332.
- RSGames.Org. Available online: <https://rsgames.org/> (accessed on 16 April 2023).
- Leite, P.D.S.; Almeida, L.D.A. Extended Analysis Procedure for Inclusive Game Elements: Accessibility Features in the Last of Us Part 2. In Proceedings of the International Conference on Human-Computer Interaction, Málaga, Spain, 22–24 September 2021; pp. 166–185.
- Nickerson, L.V.; Hermann, T. Interactive Sonification of Grid-Based Games. In Proceedings of the Audio Mostly Conference—A Conference on Interaction with Sound, Piteå, Sweden, 22–23 October 2008.
- Tim. The Tim Project, Tactile Interactive Multimedia. Available online: <http://inova.snv.jussieu.fr/tim/> (accessed on 27 November 2022).
- Friberg, J.; Gärdenfors, D. Audio Games: New Perspectives on Game Audio. In Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, Singapore, 3–5 June 2004; pp. 148–154.
- Smith, B.A.; Nayar, S.K. The RAD. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; ACM: New York, NY, USA, 2018; pp. 1–12.

22. Lumbreras, M.; Sánchez, J. Interactive 3D Sound Hyperstories for Blind Children. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems the CHI Is the Limit—CHI '99, Pittsburgh, PA, USA, 15–20 May 1999; ACM Press: New York, NY, USA, 1999; pp. 318–325.
23. Entombed—An RPG Roguelike Game for the Blind and Visually Impaired. Available online: <http://www.blind-games.com/entombed.aspx#:~:text=Entombed%20is%20a%20large%20dungeon,the%20ultimate%20dungeon%20crawling%20group> (accessed on 17 March 2023).
24. Yuan, B.; Folmer, E. Blind Hero: Enabling Guitar Hero for the Visually Impaired. In Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility, Halifax, NS, Canada, 13–15 October 2008; pp. 169–176.
25. Rovithis, E.; Moustakas, N.; Floros, A.; Vogklis, K. Audio Legends: Investigating Sonic Interaction in an Augmented Reality Audio Game. *Multimodal Technol. Interact.* **2019**, *3*, 73. [CrossRef]
26. Gluck, A.; Brinkley, J. Implementing ‘The Enclosing Dark’: A VR Auditory Adventure. *J. Technol. Pers. Disabil.* **2020**, *8*, 149–159.
27. Presti, G.; Ahmetovic, D.; Ducci, M.; Bernareggi, C.; Ludovico, L.A.; Baratè, A.; Avanzini, F.; Mascetti, S. Iterative Design of Sonification Techniques to Support People with Visual Impairments in Obstacle Avoidance. *ACM Trans. Access. Comput. (TACCESS)* **2021**, *14*, 1–27. [CrossRef]
28. Balan, O.; Moldoveanu, A.; Moldoveanu, F. Navigational Audio Games: An Effective Approach toward Improving Spatial Contextual Learning for Blind People. *Int. J. Disabil. Hum. Dev.* **2015**, *14*, 109–118. [CrossRef]
29. Anderson, J. Introducing... the Ambisonic Toolkit. In Proceedings of the Ambisonics Symposium 2009, Graz, Austria, 25–27 June 2009.
30. Arteaga, D. *Introduction to Ambisonics*; Escola Superior Politècnica Universitat Pompeu Fabra: Barcelona, Spain, 2015.
31. Gerzon, M.A. Practical Periphony: The Reproduction of Full-Sphere Sound. In Proceedings of the Audio Engineering Society Convention 65, London, UK, 25–28 February 1980.
32. Nachbar, C.; Zotter, F.; Deleflie, E.; Sontacchi, A. Ambix-a Suggested Ambisonics Format. In Proceedings of the Ambisonics Symposium, Lexington, KY, USA, 2–3 July 2011; p. 11.
33. Enzner, G.; Weinert, M.; Abeling, S.; Batke, J.-M.; Jax, P. Advanced System Options for Binaural Rendering of Ambisonic Format. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 251–255.
34. Heller, A.; Lee, R.; Benjamin, E. Is My Decoder Ambisonic? In Proceedings of the Audio Engineering Society Convention 125, San Francisco, CA, USA, 2–5 October 2008.
35. Tsang, P.W.-M.; Cheung, K.W.K.; Leung, A.C.-S. Decoding Ambisonic Signals to Irregular Quad Loudspeaker Configuration Based on Hybrid ANN and Modified Tabu Search. *Neural Comput. Appl.* **2011**, *20*, 983–991. [CrossRef]
36. Zotter, F.; Pomberger, H.; Noisternig, M. Ambisonic Decoding with and without Mode-Matching: A Case Study Using the Hemisphere. In Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics, Paris, France, 6–7 May 2010; Volume 2.
37. Ge, Z.; Wu, X.; Qu, T. Improvements to the Matching Projection Decoding Method for Ambisonic System with Irregular Loudspeaker Layouts. In Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Melbourne, Australia, 12–17 May 2019; pp. 121–125.
38. Hemingson, D.; Sarisky, M. A Practical Comparison of Three Tetrahedral Ambisonic Microphones. In Proceedings of the Audio Engineering Society Convention 126, Munich, Germany, 7–10 May 2009.
39. Rychtáriková, M.; Jedovnický, M.; Vargová, A.; Glorieux, C. Synthesis of a Virtual Urban Soundscape. *Buildings* **2014**, *4*, 139–154. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.