

LABORATÓRIO DE ÁUDIO PARA ENSINO UTILIZANDO REALIDADE VIRTUAL

AUDIO LABORATORY FOR TEACHING USING VIRTUAL REALITY

Dinis, V.¹, Ponte, P.¹, Paulo, J.^{1,2}, César, C.^{1,2}

¹ ISEL-Instituto Superior de Engenharia de Lisboa, Portugal

² LAA-Lab. de Áudio e Acústica do ISEL, Portugal

a46327@alunos.isel.pt, a46354@alunos.isel.pt, joel.paulo@isel.pt, carlos.rodrigues@isel.pt

PACS: no. 43.55.Br, 43.60.-c.

Keywords:

Processamento de sinais, acústica de salas, som 3D, medições acústicas, auralização

ABSTRACT.

The use of virtual reality currently allows to create highly immersive environments in terms of 3D image and spatialization of audio, and can, with obvious advantages, be used, in Mrst approach, in the teaching of audio and acoustics. In this paper, the authors describe the study and development of an application that models the Immersive Sound Room of ISEL for hearing with headphones (binaural) or with sound speaker system using the Unity game engine and the Steam Audio plugin for spatialization and transmission. Furthermore, for evaluating the acoustic parameters of the virtual Immersive Sound Room a 1st-order Ambisonic virtual microphone was developed. To validate the obtained results from virtual model, acoustical measurements with 3D sound capture was performed, using an artificial head and an Ambisonic microphone, before and after the requalification of the Immersive Sound Room of ISEL, to perceptually analyse the influence of the acoustic materials used in the coating of the surrounding surfaces as well as their respective location inside the room. The results showed some deviation in terms of acoustic comfort parameters, mainly because the Steam Audio does not consider the diffusion effects of the acoustic panels.

RESUMO.

A abordagem de realidade virtual permite hoje em dia criar ambientes altamente imersivos em termos de imagem e espacialização de áudio, com a vantagem de poder ser utilizada, em primeira aproximação, no ensino da temática do áudio e acústica. Este trabalho visa o estudo e desenvolvimento de uma aplicação que modela a Sala de Som Imersivo do ISEL para audição com auscultadores ou sistema de colunas de som permitindo representar o campo sonoro 3D com recurso ao motor de jogos Unity e ao plugin Steam Audio, para espacialização de áudio. No sentido de avaliar os parâmetros acústicos da sala, desenvolveu-se um microfone virtual Ambisonic de 1 ordem. Posteriormente, realizam-se um conjunto de medições acústicas com captação de som 3D usando uma cabeça artificial e um microfone ambisonic, antes e depois da requalificação da sala, para analisar perceptualmente a influência dos tipos de materiais acústicos e sua localização para validação de resultados. A utilização da ferramenta de modelação virtual permitiu em pré-projeto ter uma noção mais realista da acústica no interior da Sala de Som Imersivo, na utilização de diferentes materiais de revestimento das superfícies envolventes. No entanto, os resultados apresentam ainda algum desvio em termos dos parâmetros de conforto acústico, devido essencialmente ao Steam Audio não contabilizar os efeitos de difusão dos painéis.

1. INTRODUÇÃO

A abordagem de som imersivo tem tido uma grande projeção nos últimos anos devido essencialmente à popularidade da VR/AR e pelo cinema com som 3D (i.e. Dolby Atmos) e também pelos custos de produção terem baixado. Assim, é viável construir-se instalações de reprodução sonora fixas com sistemas multicanal de colunas de som (dezenas de altifalantes colocados à volta de uma sala), criando uma experiência sonora completamente imersiva. Este tipo de recursos tem igualmente muito interesse para fins de investigação e desenvolvimento.

O ISEL está a construir uma sala com estas características seguindo as recomendações técnicas, nomeadamente a norma Rec. ITU-R BS.1116-1, e terá um sistema de som multicanal com 40 altifalantes, como é representado no modelo virtual da Figura 1. Neste contexto, estamos a desenvolver este projeto utilizando ferramentas para criação de modelos virtuais de arquitetura e acústica, permitindo realizar uma experiência virtual de ambientes sonoros.

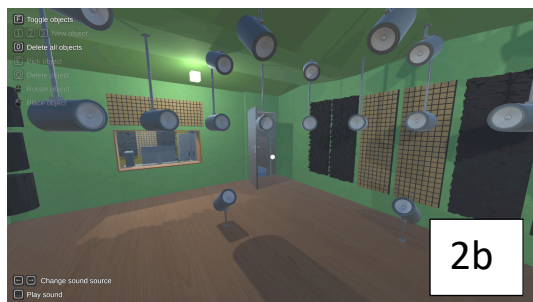
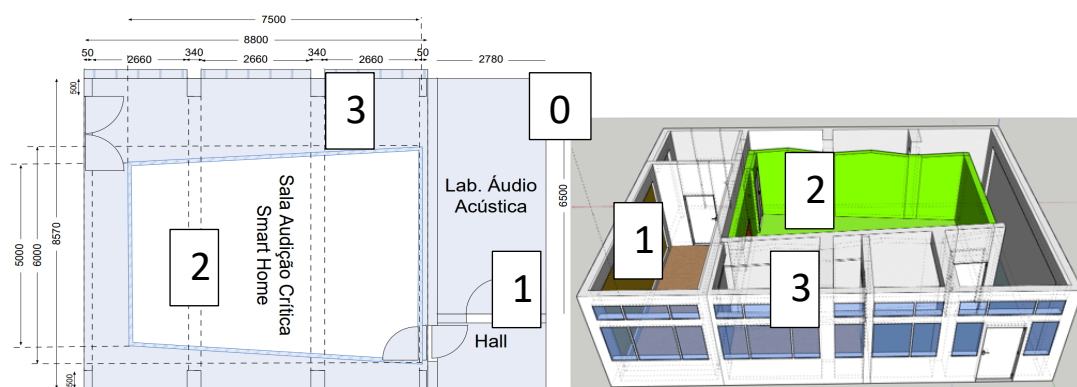


Figura 1 – Laboratório de Áudio e Acústica do ISEL e espaços destinados às atividades relacionadas com o som. Modelo virtual dos espaços (0), Laboratório de Áudio e Acústica (1), Sala de Audição Crítica e de Som Imersivo, com e sem sistema de som multicanal (2a e 2b) e Sala de pós-produção de áudio (3).

1.1. Objetivos

A ideia deste projeto surgiu da necessidade de criar uma ferramenta que permita fazer o ensino da área da acústica e áudio remotamente, no contexto da pandemia relacionada com a Covid-19. Desta forma, optou-se por desenvolver ambientes virtuais sonoros que possam ser avaliados online ou apenas em computador, com recurso a óculos de realidade virtual, VR. Assim, o ensino da área do som pode ser mais intuitivo, fazendo experiências percetuais e também com possibilidade de se realizarem ensaios acústicos virtuais.

1.2. Trabalhos relacionados

O advento de ferramentas de VR/AR potenciou o desenvolvimento de diversos projetos nas áreas do som. O projeto PATerSoN- Prototipagem acústica de ambientes médicos para o desenho de novos sinais de aviso auditivos [1] permite testar alarmes hospitalares em salas simuladas em ambiente de realidade virtual. Neste trabalho está a ser utilizada a ferramenta Steam Audio num ambiente virtual, de modo a simular a sua acústica de ambientes hospitalares e desenhar a planta de alarmes. A abordagem EAVR [2] tem o objetivo de ajudar a perceção acústica de espaços para educação com base numa experiência de som em ambientes virtuais, com VR.

O RTWH [3] é uma framework de auralização em tempo real que permite um controlo muito fino em implementações de renderização de áudio espacial. O código está em C++, voltado principalmente para plataformas Windows, e também oferece um pacote para usar as interfaces C++ de Virtual Acoustics do motor de jogo Unity, por meio de uma série de ligações C que podem ser facilmente anexadas a GameObjects, a classe base para todas as entidades numa cena de Unity3D. A estrutura também oferece a oportunidade de implementar um algoritmo personalizado de auralização em tempo real.

Wang et al. [4] revisitaram em detalhe 66 artigos sobre o uso de tecnologias de VR para educação, identificando cinco categorias principais de tecnologia: VR de desktop, VR imersiva, VR baseada em jogos 3D, VR de modelagem de informações de construção e realidade aumentada. Os autores notaram que os aplicativos de VR são usados principalmente em visualização arquitetural e educação em design, treino de segurança na construção, treino de equipamentos e tarefas operacionais e educação em análise estrutural.

EVERTims é uma framework de código aberto para auralização em tempo real em acústica arquitetural e realidade virtual [5]. Geometrias e materiais arbitrários podem ser alterados em tempo real, assim como a posição da fonte e do receptor.

2. METODOLOGIA

O motor de espacialização utilizado neste projeto é o Steam Audio, baseado no espacializador Phonon3D, desenvolvido pela empresa Impulsonic a qual os direitos foram comprados pela Valve Corporation. Trata-se de um motor de espacialização open-source, que mira proporcionar uma experiência de áudio imersiva robusta a jogos e aplicações VR. Nos últimos dois anos, devido a várias atualizações, tornou-se muito popular no meio, sendo considerada, atualmente, como um marco nas soluções de áudio imersivo.

De entre outros espacializadores de estado da arte, este provou-se ser a opção que mais se alinhava com os objetivos do projeto. Pois além de ter a vantagem de ser de utilização gratuita, ainda se destaca pela sua relativa liberdade de customização, capacidade de simular a transmissão em materiais e a possibilidade de realizar alterações no ambiente, por modo a mudar a resposta do som, em tempo real de execução [6 e 7]. A estes pontos, acrescenta-se o facto de possuir uma comunidade forte, aberta e em constante crescimento, sendo suportado em múltiplas plataformas, como Windows, macOS, sistemas Linux e Android. Contudo, o Steam Audio não tem suporte para iOS, limitando o número de dispositivos móveis nos quais aplicações que usem este motor possam abranger. Não só, devido à sua relativa novidade e por ainda estar em desenvolvimento, existe uma forte possibilidade de encontrarem-se problemas para os quais não existe solução. Isto reflete-se na documentação, que em determinados pontos é demasiado simplista e não apresenta de forma clara o funcionamento interno dos componentes descritos.

Efetivamente, o Steam Audio disponibiliza vários componentes, com os quais aplica o algoritmo de ray tracing para simular o percurso real que as ondas sonoras fariam nas situações do ambiente simulado. Internamente, o motor tenta traçar várias linhas do emissor, a fonte sonora Steam Audio Source, até ao recetor, o ouvinte Steam Audio Listener. É dada ainda a possibilidade de serem editados os parâmetros internos do algoritmo de ray tracing.

3. DESENVOLVIMENTO

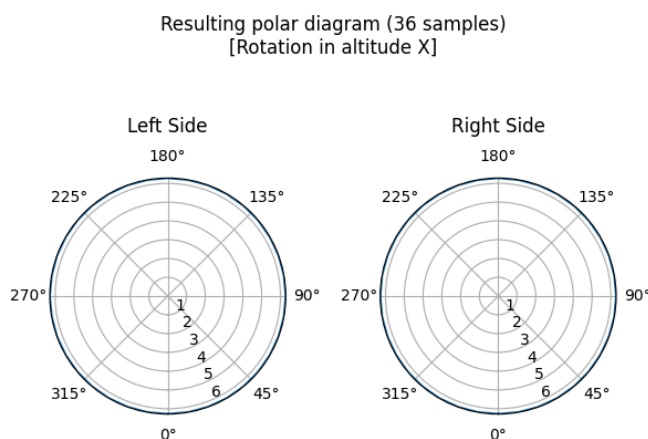
3.1 Construção e espacialização do ambiente

Foi replicado virtualmente o ambiente das salas do Laboratório de Áudio e Acústica, LAA, do ISEL – Laboratório, Sala de Som Imersivo e salas de apoio. A modulação deste espaço foi concretizado em SketchUp (Figura 1), sendo posteriormente exportado para Unity e aplicada a espacialização com recurso aos componentes Steam Audio.

3.2 Implementação de um microfone omnidirecional

Grande parte dos regulamentos relacionados com procedimentos para medições/ensaios acústicos utiliza microfones com diagramas de diretividade do tipo omnidirecional. Assim, é também importante para a realização dos ensaios acústicos em ambientes virtuais utilizar o mesmo tipo de microfone para se poder fazer comparações de desempenho.

Como o Steam Audio, nativamente, não possui um componente que funcione como microfone de referência com este tipo de diagrama polar, desenvolveu-se um dispositivo deste tipo a partir de um componente de Unity de tipo Audio Listener. Para perceber como o Steam Audio / Unity usa os microfones (Audio Listener), realizaram-se testes de diretividade a este componente, para os planos X, Y e Z. O resultado esperado da análise dos diagramas polares para o sinal captado em cada canal (ouvido esquerdo e direito) construídos seriam que, quando sobrepostos, tivessem o formato de um cardioide 3D. Isto é, dois dos planos de rotação resultariam no formato de um cardioide e o terceiro seria um círculo. Contudo, ao observarem-se os diagramas obtidos, constatou-se que a diretividade da componente Steam Audio Listener não apresentava o formato esperado. Observa-se que apenas dois dos planos apresentam os formatos expectados. O plano de Y corresponde efetivamente a um cardioide e o plano de X tem o formato de um círculo. No entanto, no plano de Z apresenta um formato semelhante ao que se pode chamar de uma "cúpula", onde apenas o segmento circular do formato cardioide esperado foi bem captado. Estes resultados são apresentados na Figura 2.



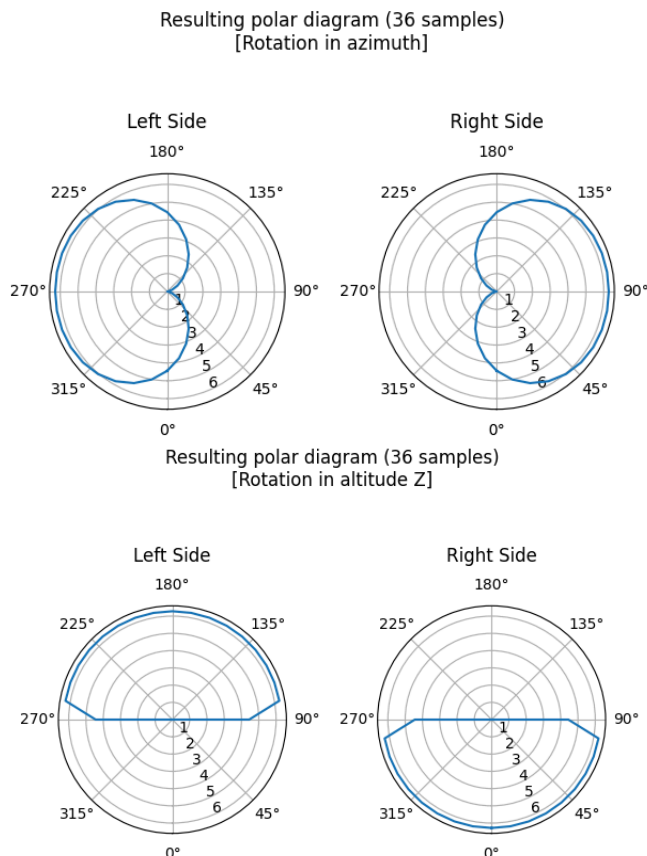


Figura 2 – Diagramas de diretividade obtidos no ambiente virtual, para os planos X, Y e Z. Os ensaios foram realizados para 36 posições de medição à volta do Audio Listener.

O microfone implementado em Unity corresponde ao Ambisonic de 1ª ordem, uma vez que a soma das 4 cápsulas (que têm diagrama polar do tipo cardioide) gera um microfone com diagrama polar omnidirecional. Atendendo ao facto de que o Unity não permitir ter vários Audio Listener em cena simultaneamente, implementou-se um script em C# (linguagem de programação nativa do Unity) responsável por controlar o processo de gravação do áudio e por comutar a cápsula ativa entre quatro posições pré-definidas, ao longo da gravação, de acordo com a geometria de posicionamento das cápsulas do esquema mostrado na Figura 3.

Quando é feita uma gravação, a cápsula única do microfone grava o áudio numa posição inicial e quando termina passa à próxima, repetindo o processo quatro vezes no total. A ordem de gravação é: FLU, FRD, BLD, BRU. No final, é criado um ficheiro wave com 4 canais, um por posição da cápsula.

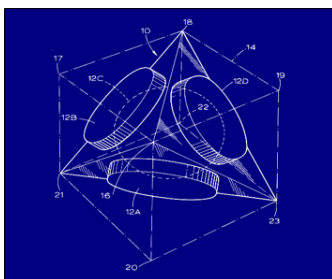


Figura 3 – Esquema de um microfone ambisonic de 1ª ordem

Para validar o funcionamento do microfone ambisonic desenvolvido, realizaram-se vários ensaios de diretividade. De entre os quais destacam-se os ensaios em espaço semelhante a sala anecoica e em sala normal.

No primeiro, foi colocado o microfone no centro de um ambiente vazio, que apresenta um comportamento semelhante ao de uma sala anecoica, onde existia uma fonte sonora que girava à sua volta, segundo os três eixos de rotação do Unity, como se observa na Figura 4.

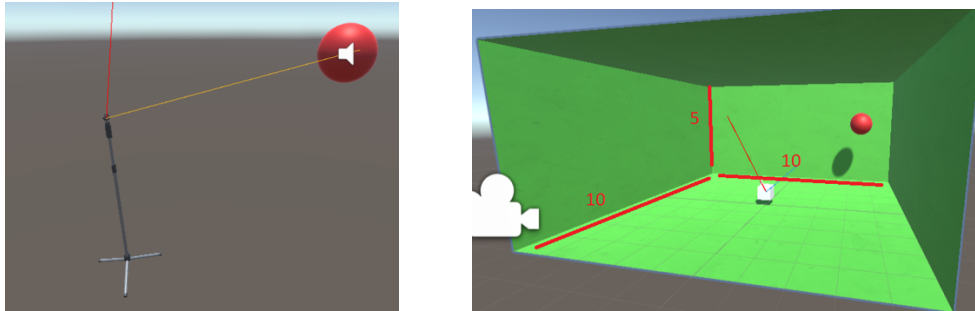


Figura 4 – Configuração de um ensaio de diretividade ao microfone Ambisonic de 1ª ordem. Medição em ambiente vazio (espaço livre) com microfone à esquerda e fonte sonora à direita (lado esquerdo) e em sala normal com espacialização de som e com características dos materiais (lado direito).

No segundo, o microfone foi colocado no centro de uma sala espacializada, onde se fizeram várias capturas com a fonte sonora em várias posições, sempre à mesma distância do microfone.

Para ambos os casos, com uma análise dos resultados com recurso a scripts Python, permitiu concluir-se que o microfone, para a componente W, apresenta um comportamento normal, diagrama polar omnidirecional.

3.3 Parâmetros utilizados

Nas configurações do Steam Audio, foram utilizados os valores configurados por omissão. Em relação aos materiais usados fez-se a seguinte abordagem: consultando os valores tabelados pelos fabricantes, interpolaram-se os valores de absorção (absorption), espalhamento (scattering) e transmissão (transmission) para as bandas de frequências do Steam Audio. Os restantes parâmetros, valores de níveis das Steam Audio Source e de materiais que não se encontravam tabelados, foram adaptados de forma livre, por modo a obter um resultado semelhante ao experimentado no espaço real.

Ao parâmetros da absorção e transmissão podem ser definidos através de constantes com valores entre 0 e 1, o comportamento do material para três bandas de frequências distintas: baixas, médias e altas. Segundo a documentação do Steam Audio [SteamAudio, 2022], na secção sobre os materiais está escrito que é considerado internamente que as baixas frequências correspondem à banda até aos 800 Hz, as médias frequências à banda dos 800 Hz aos 8 kHz e finalmente as altas frequências à banda superior aos 8 kHz. A dispersão/difusão traduz-se numa constante que toma valores entre 0 e 1. Onde 0 significa que são refletidos os sons de uma forma perfeita, como acontece num espelho, e 1 que reflete os sons de forma aleatória para todas as direções. A Figura 5 mostra a janela de configuração destes parâmetros.



Figura 5 – Painel de configuração dos parâmetros Absorção, Difusão e Transmissão dos materiais.

3.4 Aplicação Final

A versão final da aplicação desenvolvida conta com o modelo do Laboratório de Áudio e Acústica do ISEL implementado e devidamente espacializado, isto é, configurado com os efeitos de oclusão, reflexão e reverberação. Para testar, foram construídas 3 tipos de aplicações diferentes: uma onde o utilizador pode-se deslocar no ambiente do Laboratório e interagir com objetos, para computador (Figura 6); uma para o utilizador perceber a diferença entre o sistema de 32 altifalantes desenvolvido e compará-lo com componente Steam Audio Ambisonic Source, para computador (Figura 7); finalmente, uma onde o utilizador pode visitar o espaço e interagir com os objetos no ambiente, para telefone (Figura 6). A versão de telefone conta com três variantes, onde uma contém o ambiente completo e as outras duas utiliza apenas duas das salas que o compõem.

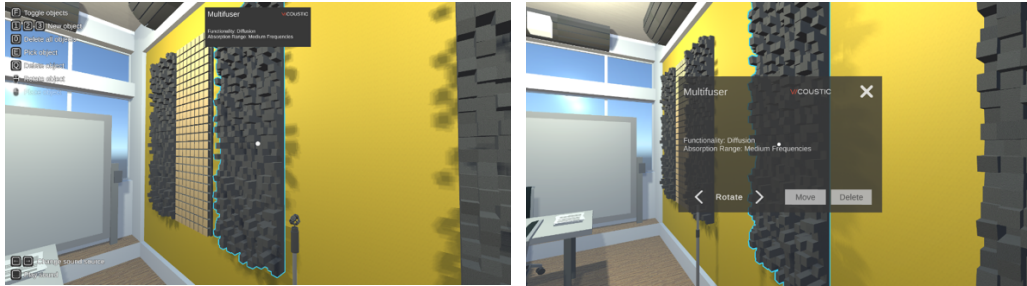


Figura 6 – Interface da aplicação para computador (versão de interação com o ambiente) e para sistema Android.



Figura 7 – Interface da aplicação para computador (versão com o sistema de altifalantes).

Para interagir com o ambiente, o utilizador tem de olhar para o determinado objeto antes de executar as ações pretendidas. Na versão de computador, as ações estão listadas no ecrã e os botões que podem ser usados são apresentados consoante o estado atual do sistema. Na versão mobile, é apresentado um menu flutuante onde estão presentes todas as ações possíveis de serem realizadas e, para executá-las, o utilizador tem de olhar para a opção pretendida durante um determinado tempo para ativá-la.

A aplicação permite utilizar sons do tipo ambisonic, até 3ª ordem, e sons multicanal já espacializados.

Os sons utilizados para aferir a aplicação foram selecionados de discografia pessoal e Speech corpus - Harvard [8], de Anechoic recordings of a Cantus Firmus [9] e AMBISONIC SOUND LIBRARY da empresa RODE.

5. CONCLUSÕES

O objetivo deste projeto foi desenvolver uma ferramenta que simule a acústica dentro de uma sala – as salas do Laboratório de Áudio e Acústica do ISEL.

Este trabalho encontra-se em evolução, desta forma, os resultados e conclusões obtidos referem-se essencialmente a uma análise perceptual. Verificaram-se discrepâncias entre o som

percebido em ambiente real e virtual. Contudo, o utilizador consegue ter uma experiência imersiva bastante realista, em termos de ambientes visuais e sonoros.

Foi desenvolvida uma aplicação para computador e para dispositivo móvel (android) com a vertente de VR que coloca o utilizador dentro de um ambiente virtual do laboratório. Nesta aplicação foram integrados os componentes do espacializador de áudio que permitem simular a percepção da audição no interior de uma sala e alguns fenómenos associados à propagação do som em espaços fechados. Deste modo, é possível oferecer ao utilizador uma experiência mais imersiva, pois conta tanto com o sentido visual como o auditivo.

Os ensaios de comparação do ambiente simulado com o ambiente real não foram concretizados por este encontrar-se ainda em construção, não permitindo assim validar os parâmetros de espacialização atribuídos ao ambiente. Contudo, a percepção auditiva experimentada no ambiente virtual traduz boas semelhanças com a realidade (em salas similares), em termos de reverberação, oclusão de fonte sonora e transmissão sonora entre paredes.

Desenvolveu-se um microfone ambisonic de 1ª ordem para realizar medições acústicas em ambiente virtual. Embora existam discrepâncias nas respostas de diretividade, a componente W apresenta uma forma omnidirecional, o que permite ser usado para ensaios acústicos.

Como trabalho futuro pretendemos desenvolver: um algoritmo de parametrização dos coeficientes de absorção sonora no Steam Audio utilizando uma abordagem algoritmos genéticos ou de inteligência artificial; a escolha dos materiais atribuídos aos objetos foi feita de forma iterativa, pois apesar dos valores serem baseados em valores tabelados estes apresentavam intervalos diferentes dos utilizados pela ferramenta Steam Audio. Deste modo propomos a elaboração de um algoritmo que realize este cálculo e apresentasse os parâmetros para o resultado mais próximo da realidade.

Pretendemos elaborar de uma versão da aplicação compatível para a web. Esta versão ficará acessível no portal de internet do Laboratório de Áudio e Acústica, LAA, e servirá como uma ferramenta de ensino, permitindo a realização de aulas remotamente, pois não implica que os participantes se desloquem para o laboratório, e deste modo, possibilita a realização de experiências no ambiente simulado.

No sentido de melhorar a experiência imersiva, a utilização de controlo por voz para interagir com a aplicação na versão para dispositivo móvel (telemóvel), apenas para VR, para complementar a interação por gestos, tem vantagens. Para ser implementado um sistema de controlos por voz é proposto o uso da classe KeywordRecognizer, disponibilizada pelo Unity que permite associar um dicionário de palavras chave, às quais o sistema vai estar à escuta.

Está também previsto a elaboração de um caso de utilização que consiste no sincronismo da aplicação virtual com o sistema multicanal de altifalantes da Sala de Som Imersivo. Ou seja, este caso é uma alternativa à utilização da aplicação na versão para smartphone, onde o equipamento de reprodução sonora é o conjunto de altifalantes em vez dos auscultadores, ou seja, estando o utilizador no interior da sala é possível comparar a sensação auditiva no caso simulado e real. Esta opção recorre ao protocolo OSC, que é um protocolo típico de troca de mensagens em tempo real entre aplicações e equipamentos remotamente para comunicar com um dispositivo ligado ao sistema de som da sala, por via de uma DAW, capaz de responder a estes pedidos.

AGRADECIMENTOS

Este trabalho está a ser suportado pelo Laboratório de Áudio e Acústica do Instituto Superior de Engenharia de Lisboa, ISEL, LAA

<https://acusticaudiolab.isel.pt>

Agradecimentos especiais para Jorge Castro da empresa Artnovion pelo projeto acústico preliminar e pela disponibilização dos modelos virtuais dos painéis acústicos para a sala de som imersivo.

REFERÊNCIAS

- [1] Projeto PATErSoN – Prototipagem acústica de ambientes médicos para o desenho de novos sinais de aviso auditivos, 2019
<https://ccg.pt/my-product/paterson-prototipagem-acustica-de-ambientes-medicos-para-o-desenho-de-novos-sinais-de-aviso-auditivos/>
- [2] Alessia Milo 1, Maarten Hornikx, Acoustic Virtual Reality as a Learning Framework for Built Environment Students, EuroNoise, 25-27 October, Madeira, Portugal, 2021
- [3] Institute for Hearing Technology and Acoustics, RWTH Aachen University. Virtual Acoustics – A real-time auralization framework for scientific research. <http://www.virtualacoustics.org/>. Accessed on 2021-08-12.
- [4] Peng Wang, Peng Wu, Jun Wang, Hung-Lin Chi, and Xiangyu Wang. A critical review of the use of virtual reality in construction engineering education and training. International journal of environmental research and public health, 15(6):1204, 2018.
- [5] David Poirier-Quinot, Brian Katz, and Markus Noisternig. Evertims: Open source framework for real-time auralization in architectural acoustics and virtual reality. In 20th International Conference on Digital Audio Effects (DAFx-17), 2017.
- [6] Josep Llorca, Ernest Redondo, and Michael Vorländer. Learning room acoustics by design: a project-based experience. International Journal of Engineering Education, 35(1):1–7, 2019.
- [7] Marcin Gorzel, Andrew Allen, Ian Kelly, Julius Kammerl, Alper Gungormusler, Hengchin Yeh, and Francis Boland. Efficient encoding and decoding of binaural sound with resonance audio. In AudioEngineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio. Audio Engineering Society, 2019.
- [8] Philippa Demonte. Speech corpus - harvard - edited and end-pointed audio. https://salford.figshare.com/articles/media/Speech_corpus_-_Harvard_-_edited_and_end-pointed_audio/7862465/1, Mar 2019.
- [9] Dario D’Orazio. Anechoic recordings of a cantus firmus. <https://data.mendeley.com/datasets/mjn6fvxfxb/1>, May 2019.
- [10] AMBISONIC SOUND LIBRARY by RODE, <https://library.soundfield.com>
- [11] S. P. Lipshitz, J. Vanderkooy, Experiments in Direct/Reverberant Ratio Modification, AES 79th Convention, Paper 2301 (1985)