

DESENVOLVIMENTO DE UM FONE DE OUVIDO COM AUDIO IMERSIVO UTILIZANDO HRTF

Luciano Neves dos Santos Junior (IC) e Paulo Batista Lopes (Orientador)

Apoio: PIVITI Mackenzie

RESUMO

As tecnologias de computação gráfica utilizadas na área de realidade virtual e realidade aumentada têm avançado de forma acelerada nos últimos anos, proporcionando uma experiência visual de muita qualidade ao usuário deste tipo de sistemas. A parte sonora, infelizmente, não acompanhou este desenvolvimento e só permite ao usuário uma sensação tridimensional por meio da utilização de um número significativo de alto-falantes. Portanto, faz-se necessário um estudo que proporcione uma maior ambientação também nos canais auditivos. Com este objetivo, a função de transferência relacionada à cabeça (do inglês, head related transfer function, HRTF) foi estudada. A HRTF modela, no domínio da frequência, as modificações sofridas pelas ondas sonoras ao se propagarem da fonte sonora ao ouvido de uma pessoa. Estas modificações são dependentes da anatomia da pessoa e de sua posição corporal. Por meio de técnicas de redes de sensores sem fio e Internet das Coisas, este estudo implementou um sistema para determinar os fatores que determinam a HRTF a ser utilizada. A HRTF assim obtida é implementada computacionalmente a fim de se obter um efeito de áudio tridimensional em um fone de ouvido comum. A técnica desenvolvida permite que o usuário tenha uma experiência sonora tão realística quanto a visual em sistemas de RA e RV.

Palavras-chave: Áudio 3D; Processamento de Sinais de Áudio; HRTF

ABSTRACT

The technology of CG used in the area of virtual reality (VR) and augmented reality (AR) have advanced very quickly in the last years, providing a great visual experience for users of that system. The sound design didn't develop as fast, being unable to provide the feeling of a tridimensional space without the use of a significant number of speakers. Therefore, it is necessary a study, that provides a greater ambience in the sound area. The Head Related Transfer Function (HRTF) is an alternative to reach that goal. The HRTF shapes, on the frequency domain, the modifications suffered by the sound waves when it propagates the signal from the sound source to the ear of a person. The modifications are dependents of the anatomy of the person and the body position. With the technique of networks wireless and Internet of Things, the study implemented a system to determine the factors that set the best HRTF to use. That way, the HRTF is implemented computationally in order to obtain a three-

dimensional audio effect in a single earphone. The developed technique allows the user a greater sound experience as realistic as the systems of AR and VR.

Keywords: 3D Audio; Audio Signal Processing; HRTF

1. INTRODUÇÃO

Sistemas baseados em Realidade Virtual (RV) ou realidade aumentada (RA) são utilizados em inúmeras aplicações: treinamento, jogos digitais, simulações, etc. Enquanto a parte visual destas aplicações progrediu rapidamente e as animações são extremamente realísticas, o áudio não acompanhou esta evolução e continua ser uma versão estéreo (dois canais) que não proporciona uma experiência imersiva proporcional à parte visual destes programas de RA ou RV (Johansson, 2019).

Para uma experiência auditiva ser realisticamente imersiva, o(a) usuário(a) deve poder localizar as fontes sonoras como se ele ou ela estivesse pessoalmente naquele ambiente quer seja um campo de batalha, uma sala de concerto, uma fábrica com máquinas, a cabine de um avião, etc. A aplicação deve transmitir a percepção de uma ambientação na qual as fontes sonoras sejam distinguíveis e localizáveis em um espaço tridimensional: o som 3D.

Sistemas de som 3D difundidos nos chamados equipamentos de *Home Theater* nos quais um conjunto de 5 ou 6 alto-falantes posicionados ao redor do ouvinte transmitem a sensação de imersão sonora. O uso deste número de alto-falantes, todavia, impede a utilização deste tipo de sistemas em caso nos quais não é possível preparar um ambiente de audição apropriado. Na maioria das aplicações, fones de ouvido são utilizados em dispositivos móveis ou portáteis. E criar uma ambientação sonora 3D em fones de ouvido constitui um desafio tecnológico a ser abordado neste projeto (Johansson, 2019).

Neste contexto, é importante mencionar que o ser humano possui a capacidade de identificar a localização de uma fonte sonora sem movimentar a cabeça, utilizando os dois ouvidos. Como isso é realizado? As ondas acústicas são modificadas pela cabeça, tronco e orelha ouvinte antes de atingir a parte interna do ouvido humano. Estas modificações são influenciadas pela direção de propagação destas ondas sonoras. A função de transferência relativa à cabeça (HRTF, da expressão inglesa *head related transfer function*) representa as alterações físicas da onda sonora no domínio da frequência devido a estes efeitos. Infelizmente, a HRTF é individualizada e, por esta razão, ainda não é utilizada nas aplicações que demandem áudio tridimensional, exceto em situações de âmbito restrito como jogos eletrônicos.

Sistemas baseados em Realidade Virtual (RV) ou realidade aumentada (RA) são utilizados em inúmeras aplicações: treinamento, jogos digitais, simulações, etc.

Enquanto a parte visual destas aplicações progrediu rapidamente e as animações são extremamente realísticas, o áudio não acompanhou esta evolução e continua ser uma versão estéreo (dois canais) que não proporciona uma experiência imersiva proporcional à parte visual destes programas de RA ou RV (Johansson, 2019).

Para uma experiência auditiva ser realisticamente imersiva, o(a) usuário(a) deve poder localizar as fontes sonoras como se ele ou ela estivesse pessoalmente naquele ambiente quer seja um campo de batalha, uma sala de concerto, uma fábrica com máquinas, a cabine de um avião, etc. A aplicação deve transmitir a percepção de uma ambientação na qual as fontes sonoras sejam distinguíveis e localizáveis em um espaço tridimensional: o som 3D.

Sistemas de som 3D difundidos nos chamados equipamentos de *Home Theater* nos quais um conjunto de 5 ou 6 alto-falantes posicionados ao redor do ouvinte transmitem a sensação de imersão sonora. O uso deste número de alto-falantes, todavia, impede a utilização deste tipo de sistemas em caso nos quais não é possível preparar um ambiente de audição apropriado. Na maioria das aplicações, fones de ouvido são utilizados em dispositivos móveis ou portáteis. E criar uma ambientação sonora 3D em fones de ouvido constitui um desafio tecnológico a ser abordado neste projeto (Johansson, 2019). Uma possível solução é o emprego do conceito das HRTFs em uma implementação que considere a posição da cabeça do ouvinte que, neste caso, é o jogador.

1.1 PROBLEMA DE PESQUISA

Funções de transferência são comumente utilizadas na análise de sistemas lineares no domínio da frequência. Elas representam a relação entre a saída e a entrada de um sistema linear, geralmente invariante no tempo no domínio da frequência (transformadas de Laplace, Fourier ou z).

No caso das HRTFs, elas representam as modificações causadas pelo meio ambiente e pela anatomia do ouvinte em uma onda sonora no percurso entre a fonte e o interior do ouvido. Elas dependem, claramente, além dos itens já mencionados do posicionamento do ouvinte, em especial do posicionamento de sua cabeça. Portanto, as HRTFs são individualizadas e, para serem aplicadas em situações cotidianas seria necessário dispor de informações biométricas e de posicionamento do audiente, o que evidentemente não é possível obter rapidamente. Por esta razão, as HRTFs não são

utilizadas em ambiente que não sejam controlados, como em um laboratório de pesquisa em Acústica

Desta observação, surge a questão cuja resposta foi pesquisada. Será que, em uma situação especial, como por exemplo em jogos que utilizam Realidade Virtual (RV) ou Realidade Aumentada (RA) é possível utilizar sensores para determinar o posicionamento da cabeça do jogador, simular uma HRTF com base nesta informação e propiciar uma experiência de áudio 3D por meio de fones de ouvido a este ouvinte?

Portanto, o produto esperado desta pesquisa é um sistema que determine a posição da cabeça de uma pessoa, transmita esta informação para um computador central que selecionará uma HRTF, aplicará esta função de transferência a um som do jogo eletrônico e possibilitará ao jogador uma experiência de áudio tridimensional na qual será possível identificar a direção na qual o som chega aos ouvidos.

1.2 OBJETIVOS

1.2.1 OBJETIVO GERAL

Implementar um sistema de som computadorizado 3D que possa ser utilizado em fones de ouvido, com baixo custo, utilizando acelerômetros para determinar a posição da cabeça do(a) usuário(a) e conexão Bluetooth para transmitir esta informação para que a HRTF apropriada seja selecionada.

1.2.2 OBJETOS ESPECIFICOS

Como objetivos específicos, podem ser citados:

- produzir um fone de ouvido de baixo custo com capacidade de reproduzir áudio imersivo;
- utilizar as técnicas de Processamento Digital de Sinais em uma aplicação prática;
- ampliar a capacidade de produzir imersão dos sistemas de RA e RV;
- apreender os conceitos de Head Related Transfer Function (HRTF).

1.3 JUSTIFICATIVA

Programas que utilizam realidade virtual (RV) ou realidade aumentada (RA) têm sido utilizados em aplicações de entretenimento, treinamento, simulações, etc. Neles, é criado um ambiente no qual o(a) usuário(a) é inserido. Este ambiente pode ser totalmente criado artificialmente no caso de RV ou conter elementos do entorno no qual o(a) usuário(a) se encontra no caso de RA. A imersão no ambiente virtual deve ser a maior possível, segundo a teoria do fluxo (Csikszentmihalyi, 1999), fazendo o (a) usuário(a) dedicar toda a sua concentração às tarefas propostas pelo programa, gerando uma sensação de prazer conectada à maneira pela qual se pratica estas tarefas.

Esta imersão total é usualmente aproveitada para cativar os praticantes de jogos digitais, um mercado mundial de US \$ 137.9 B em 2018 (Wijman, 2019), focar a atenção de um piloto que é treinado para conduzir um novo modelo de avião, habilitar um jovem que está em processo de obtenção da carteira nacional de habilitação (CNH), ajudar um aluno que tenta entender uma reação nuclear em um simulador computacional, dar suporte a um pesquisador que analisa a proliferação de bactéria em um reator de batelada, etc. Pode-se dizer que quanto mais realista e cativante for um ambiente em RA ou RV, maior o aproveitamento do usuário.

Impulsionada pela adoção dos jogos digitais, a parte gráfica dos sistemas em RA e RV evoluiu drasticamente desde os primeiros jogos (GONZALES-CAMPOS; ARNEDO-MORENA; SANCHEZ-NAVARRO,2018) (ALLEY T.,2018). Há bibliotecas de livre acesso, ferramentas de desenvolvimento, arquiteturas avançadas de processadores gráficos, etc tornadas disponíveis aos desenvolvedores de aplicações de RA e RV de forma paga ou gratuita para tornar a experiência visual do(a) usuário(a) muito realística, proporcionando a tão procurada imersão total.

Embora necessário para tornar ainda mais realística a imersão no mundo virtual, o áudio não evoluiu da mesma forma, talvez pela dificuldade em tornar disponível um ambiente de sonoridade 3D em dispositivos com número restrito de alto-falantes a um custo acessível. Um ambiente de áudio imersivo ou áudio 3D consiste na emissão de sons de forma que a pessoa se sinta dentro do ambiente no qual ela deva estar imerso. Som 3D indica que o(a) usuário(a) pode localizar fontes sonoras em um espaço tridimensional, inferindo a direção e a distância dele às fontes sonoras. As vantagens de sistemas RA ou RV com sonorização 3D são citadas por Mora-Lumbreras (2012). Em resumo, o som 3D

- provê um incentivo extra para o usuário(a) localizar objetos enquanto navega;
- ajuda a produzir um nível elevado de imersão no ambiente virtual;
- auxilia a interpretar distâncias entre objetos;
- facilita uma interação mais natural porque é similar aos sons no mundo real;
- provê informação adicional à parte gráfica, ajudando o(a) usuário(a) a absorver as informações disponíveis no ambiente virtual sem esforço extra.

Para ajudar a popularização dos sistemas de sonorização 3D, é necessário fazer um produto compacto e de baixo custo, já que os dispositivos de áudio existentes no mercado são antiquados e caros. Isto é possível por meio de técnicas de Processamento Digital de Sinais (Diniz et al., 2003), cuja implementação já é feita de forma eficaz nos processadores embutidos nos dispositivos portáteis e móveis. Contudo, todo algoritmo requer uma validação antes de poder ser utilizado de forma segura. Neste caso, o software Matlab da empresa MathWorks será usado para validação dos algoritmos. Este software reúne uma interface de desenvolvimento rápida e um interpretador de uma linguagem própria, porém muito próxima da linguagem C++. Como ele é muito popular em aplicações de modelagem matemática de sistemas físicos, existem muitas funções, programas e ferramentas disponibilizadas gratuitamente ou não para ajudar no desenvolvimento de algoritmos.

Com isso, almeja-se melhorar a experiência do usuário nas áreas de entretenimento, como o mercado educacional e de jogos digitais.

2. REFERENCIAL TEÓRICO

O ser humano é capaz de distinguir a direção e a distância dele até uma fonte sonora por utilizar as informações adquiridas através de seus dois ouvidos. A figura 1 ilustra este conceito.

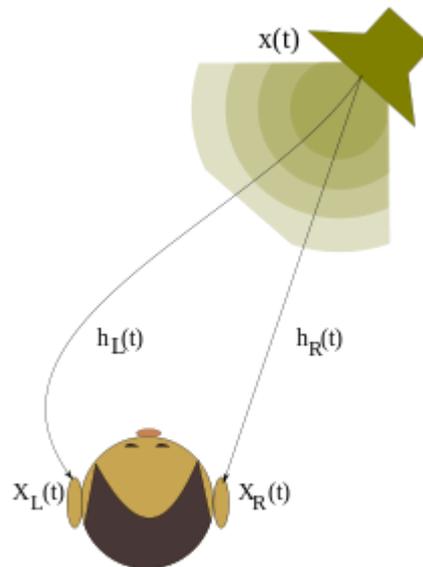


Figura 1: A percepção de sons 3D nos seres humanos (HEAD-RELATED... 2021)

O conceito ilustrado por esta figura é que os trajetos percorridos pelas ondas acústicas partindo de uma fonte sonora até cada um dos ouvidos de uma pessoa são diferentes. Esta diferença em trajetos faz com que o mesmo som, ao chegar a cada um dos ouvidos, seja modificado e apresente características de potência, conteúdo em frequência, reverberação, etc distintas (Carlile, 1996). Estas características evidentemente dependem do ambiente, da posição da cabeça em relação à fonte sonora, do conteúdo espectral do som emitido, de fatores psicoacústicos do ouvinte, etc. A modelagem de todos estes fatores de forma matemática é extremamente complexa e pode custar muito em termos computacionais. Neste cenário, o conceito de função de transferência relacionada à cabeça (HRTF, sigla derivada do termo em inglês *head related transfer function*) é muito importante.

Em sistemas lineares invariantes no tempo, função de transferência é uma forma de relacionar a entrada e saída no domínio da frequência. Para tal tipo de sistema, a função de transferência é igual a razão entre a transformada de Fourier do sinal de saída e a transformada de Fourier do sinal de excitação (Haykin and Veen, 2001). A função HRTF busca representar as mudanças espectrais que ocorrem quando a onda acústica se

propaga da fonte até um dos ouvidos de um ser humano. Estas modificações ocorrem devido ao ambiente acústico e, também, devido à difração da onda que ocorre na cabeça, no ombro, no dorso e na parte externa do ouvido. (Usman et al., 2017).

As HRTFs (uma para o ouvido direito e outra para o ouvido esquerdo) são função de 4 variáveis espaciais independentes (abstraindo-se a dependência da anatomia do ouvinte): a frequência e três variáveis que especificam a posição da fonte sonora em relação à cabeça do ouvinte. Quando a distância é superior a 1 m, a fonte é dita estar no campo distante e as HRTFs passam a depender, em coordenadas esféricas, da frequência e dos ângulos de azimute e de elevação. A figura 2 ilustra este conceito.

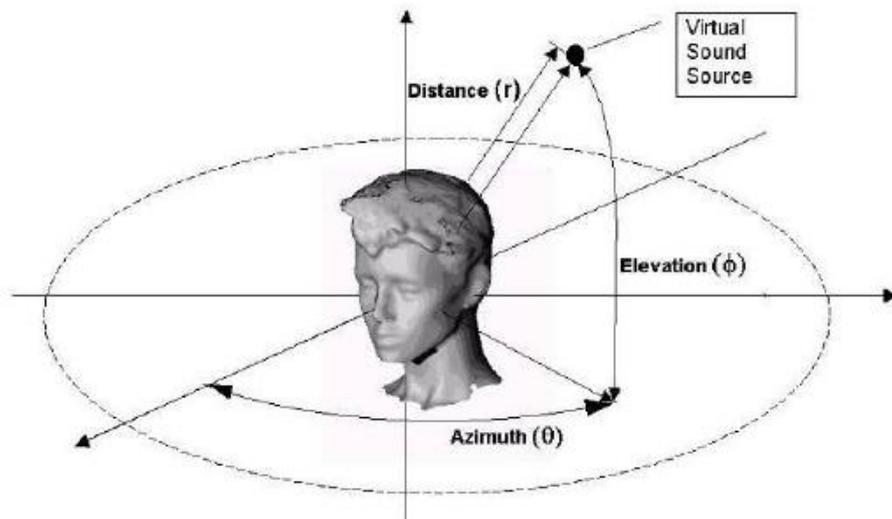


Figura 2: Coordenadas esféricas para a HRTF (Faller and al., 2006)

Neste contexto, é importante definir o sistema de coordenadas a ser utilizado. Considerando-se um sistema tridimensional de coordenadas cartesianas cuja origem está posicionada no centro da cabeça do ouvinte, azimute é definido como o ângulo entre o plano vertical equidistante entre os dois ouvidos e o plano vertical que passa pela fonte sonora. Elevação, por sua vez, é o ângulo entre o plano horizontal que abrange os dois ouvidos e a reta que contém a origem do sistema cartesiano e a fonte sonora. Conseqüentemente, as HRTFs são tremendamente influenciadas pela posição e pelo movimento da cabeça do ouvinte e qualquer alteração desta posição modifica as HRTFs.

Considerando-se o exposto anteriormente, o processamento requerido pela geração de som 3D pode ser resumido de forma macroscópica pelo fluxograma da figura 3. Inicialmente, a posição da cabeça é determinada para que as HRTFs possam ser calculadas. As amostras do arquivo sonoro são filtradas de acordo com as HRTFs selecionadas e a saída é reproduzida em um fone de ouvido.

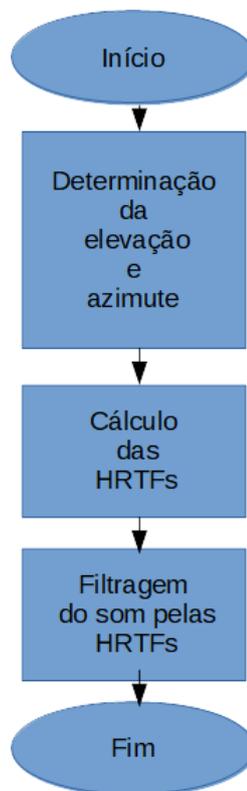


Figura 3: Processamento de som 3D

A implementação desse diagrama simplificado, todavia, requer o desenvolvimento de alguns algoritmos bastante sofisticados. Por exemplo, a determinação da posição da cabeça necessita de hardware e software de sensoriamento do azimute e da elevação. A seleção das HRTFs é baseada nestes parâmetros e depende de medições antropométricas e psicoacústicas para ser executada.

3. METODOLOGIA

É relevante entender não somente o 3D, mas também suas características, elementos e outros aspectos, sob a óptica de Processamento Digital de Sinais (PDS). Os conceitos de amostragem, transformada z, função de transferência discreta, estimação espectral, etc serão utilizados esta pesquisa (Diniz et al., 2003). Conseqüentemente, esta pesquisa iniciou-se por uma revisão dos conceitos de PDS aplicados a som 3D. Esta revisão será feita com base em artigos disponíveis nas seguintes bases de dados: IEEE, Web of Sciences, Scopus, ACM, etc.

A proposta para determinar a posição da cabeça baseia-se no uso de acelerômetros (Carrogi-Vianna, 2017). Acelerômetro é um circuito integrado capaz de medir as acelerações a que ele é submetido em três direções ortogonais, normalmente expressas em coordenadas cartesianas. A leitura destas acelerações indica a posição angular da cabeça do ouvinte, desde que uma posição de referência seja convencionalmente previamente. O dispositivo utilizado foi o circuito integrado MPU6050 que incorpora um acelerômetro e um giroscópio pelo seu custo-benefício e fácil utilização.

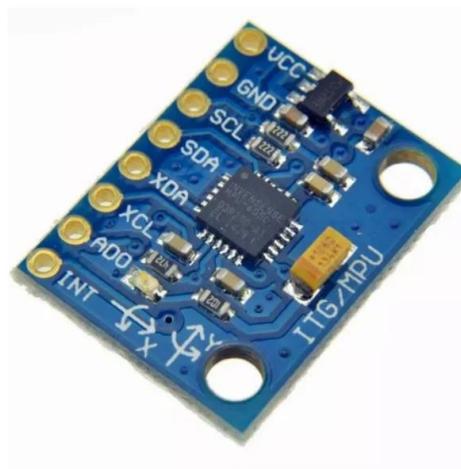


Figura 4: Acelerômetro MPU6050

Devido a alguns artefatos possíveis de acontecerem na coleta dos valores das acelerações, o uso de acelerômetros requer um tratamento estatísticos destes sinais (Kay, 1993), utilizando técnicas bayesianas (Algazi et al., 2001), implementado por meio do Matlab. A conexão do acelerômetro ao computador no qual o software Matlab está instalado é feita por meio de um *link* Bluetooth estabelecido por um microcontrolador ESP-32, muito utilizado em projetos de Internet das Coisas por ter um baixo consumo energético, boa conectividade Bluetooth e facilidade de programação através do IDE Arduino. Uma de suas grandes vantagens consiste em já ter as pilhas de Wifi e Bluetooth

embarcados, o que facilita a obtenção nos dados e a montagem de um hardware com sua inclusão.



Figura 5: microcontrolador ESP-32

Os dispositivos MPU6050 e o ESP32 são utilizados para fazer a aquisição da posição da cabeça do ouvinte, conforme mostrado na figura 6.

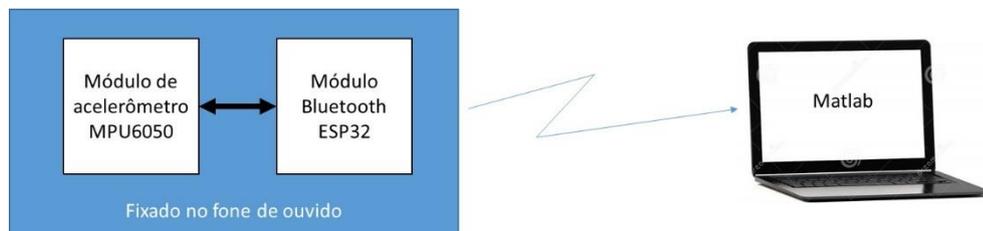


Figura 6: Diagrama de blocos do hardware de aquisição da posição da cabeça

Neste diagrama, os módulos contendo o acelerômetro e a interface Bluetooth são fixados no fone de ouvido. Os sinais provenientes do acelerômetro são transmitidos via Internet das Coisas para um computador no qual o programa de simulação Matlab está instalado. Todos os algoritmos são executados neste ambiente computacional. A reprodução do som 3D é feita em um fone de ouvido conectado à placa de som deste computador.

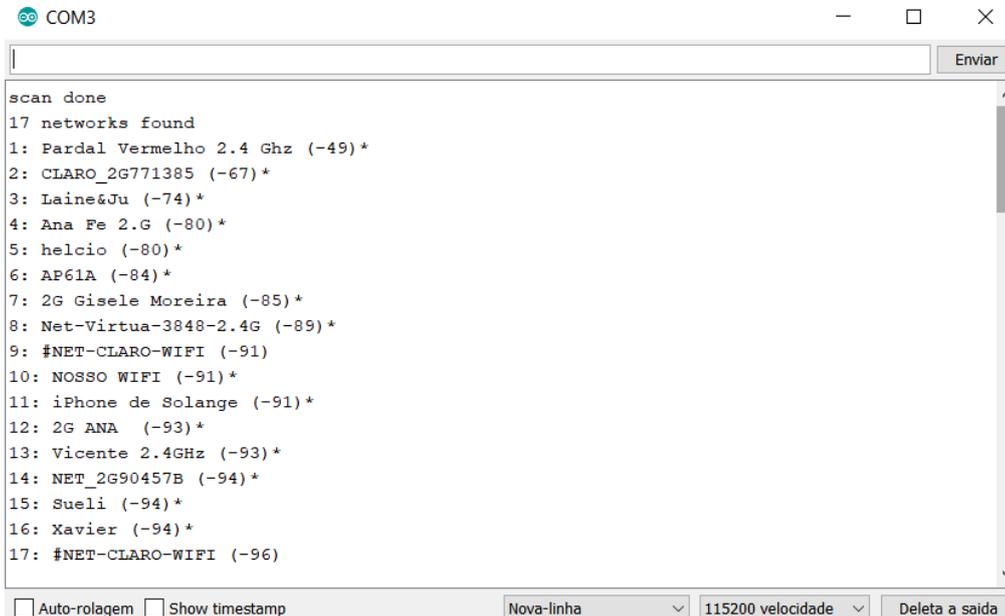
Tendo determinada a posição da cabeça, faz-se a seleção da HRTF que mais se aproxima daquele azimute e elevação. Existem base de dados com estas funções tabuladas para valores de azimute e elevação variando a cada 5°. (Lanz e Brunelli, 2007). Todavia, a troca abrupta de valores da HRTF conforme o usuário mova a sua cabeça faz com que o som 3D soe artificial e, algumas vezes, até mesmo desconfortável para o

ouvinte. Para evitar esta situação, faz-se necessário utilizar alguma técnica de suavização destas mudanças na HRTF. Como ainda não há um consenso na literatura sobre qual a melhor maneira de implementar esta suavização, uma parte da pesquisa deveria abranger simulações e teste de um algoritmo capaz de interpolar de forma suave, sugerindo continuidade as HRTFs da base de dados. A princípio, dependendo do desenvolvimento da pesquisa, esta seleção indicaria duas HRTFs pré-catalogadas manipuladas no domínio das transformadas de Fourier para a obtenção de uma terceira HRTF, calculada nesta etapa da implementação, que será efetivamente utilizada para a implementação do som 3D. Infelizmente, o atraso na execução do projeto causado pela pandemia, não permitiu a realização desta parte da pesquisa

A plataforma para validação dos algoritmos é o popular programa de simulação chamado Matlab, software amplamente usado em várias áreas da engenharia para diversas aplicações, sendo algumas delas o processamento digital de sinais, cálculos matemáticos, entre outros.

4. RESULTADO E DISCUSSÃO

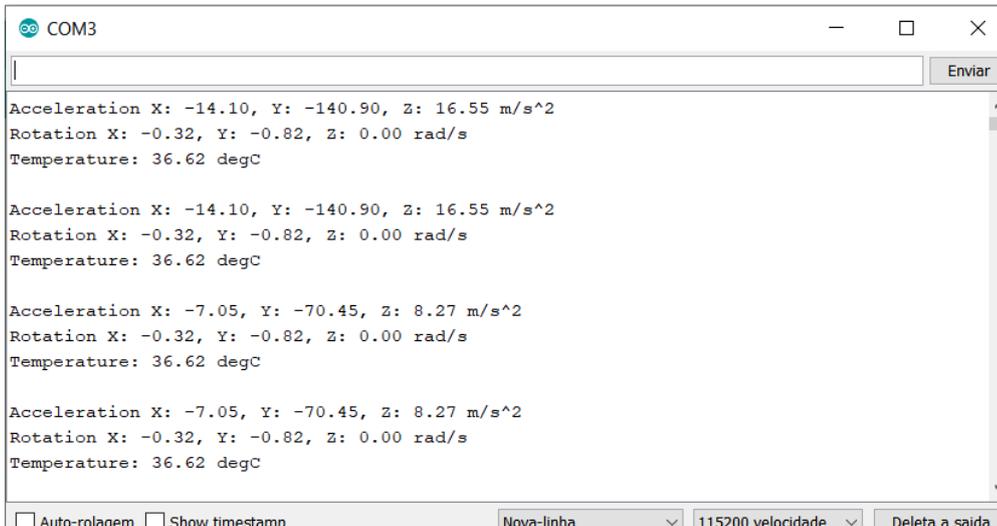
Primeiramente foram realizados testes para maior entendimento da programação e da plataforma IDE Arduino utilizando o ESP-32, como se pode ver na figura abaixo, na qual representa uma varredura das conexões Wifi disponíveis nos arredores do dispositivo.



```
COM3
scan done
17 networks found
1: Pardal Vermelho 2.4 Ghz (-49)*
2: CLARO_2G771385 (-67)*
3: Laine&Ju (-74)*
4: Ana Fe 2.G (-80)*
5: helcio (-80)*
6: AP61A (-84)*
7: 2G Gisele Moreira (-85)*
8: Net-Virtua-3848-2.4G (-89)*
9: #NET-CLARO-WIFI (-91)
10: NOSSO WIFI (-91)*
11: iPhone de Solange (-91)*
12: 2G ANA (-93)*
13: Vicente 2.4GHz (-93)*
14: NET_2G90457B (-94)*
15: Sueli (-94)*
16: Xavier (-94)*
17: #NET-CLARO-WIFI (-96)
```

Figura 7: Varredura das conexões Wifi

Depois de estudos sobre programação no ambiente IDE Arduino, foi implementada uma programação onde é possível encontrar as coordenadas do acelerômetro como se pode ver na figura abaixo.



```
COM3
Acceleration X: -14.10, Y: -140.90, Z: 16.55 m/s^2
Rotation X: -0.32, Y: -0.82, Z: 0.00 rad/s
Temperature: 36.62 degC

Acceleration X: -14.10, Y: -140.90, Z: 16.55 m/s^2
Rotation X: -0.32, Y: -0.82, Z: 0.00 rad/s
Temperature: 36.62 degC

Acceleration X: -7.05, Y: -70.45, Z: 8.27 m/s^2
Rotation X: -0.32, Y: -0.82, Z: 0.00 rad/s
Temperature: 36.62 degC

Acceleration X: -7.05, Y: -70.45, Z: 8.27 m/s^2
Rotation X: -0.32, Y: -0.82, Z: 0.00 rad/s
Temperature: 36.62 degC
```

Figura 8: Monitor serial da plataforma IDE Arduino apresentando os dados do acelerômetro

Na utilização do Bluetooth Low Energy, existem dois tipos de dispositivos: o cliente e o servidor. Enquanto o cliente envia os dados emitidos pelo sensor, o servidor recebe tais dados e os armazena para futuro uso.

Neste ponto, fez-se necessário a introdução de uma biblioteca Bluetooth na programação do acelerômetro. Essa biblioteca tem como função tornar o ESP32 cliente e emitir os dados do acelerômetro para o computador onde o sinal será tratado no software Matlab.

É importante entender os conceitos de cliente GATT (Perfil de Atributo Genérico, do inglês Generic Attribute Profile) onde é definido uma hierarquia de estrutura de dados e descreve a forma como dois dispositivos bluetooth irão se comunicar.

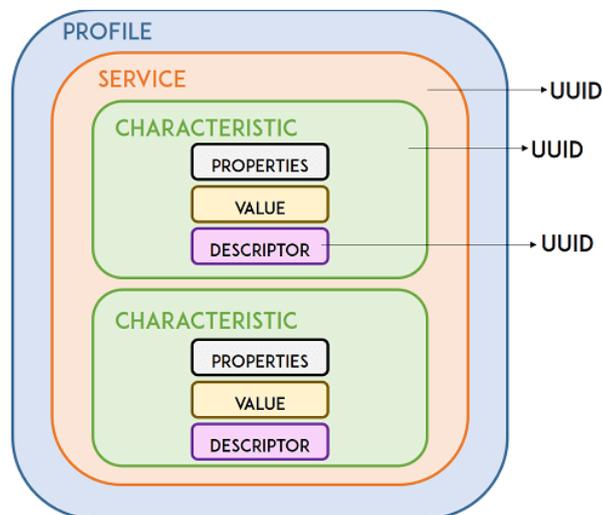


Figura 9: Exemplo da hierarquia do cliente GATT (SANTOS, 2021)

Os dispositivos Bluetooth LE funcionam através de um conjunto de 3 instruções:

- Descritores
- Características
- Serviços

Os serviços ditam como o dispositivo vai funcionar através de uma coleção de características. Essas características possuem descritores que as especificam. Esse conjunto de instruções define como os dados serão emitidos do cliente até o servidor.

Os serviços e características são identificados por um UUID (Identificador Único Universal, do inglês Universally Unique Identifier), um número de 128 bits usado para possibilitar a identificação única de uma informação.

Definindo tais conceitos e implementando na programação, temos os dados sendo emitidos do microcontrolador até o software Matlab como visto a seguir.

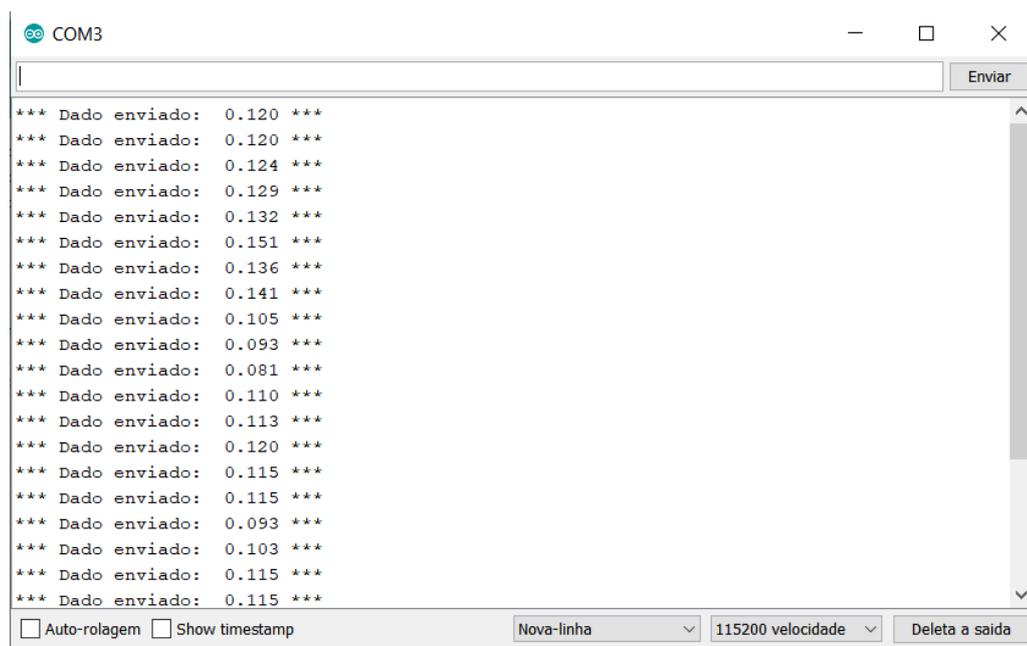


Figura 10: Monitor Serial do IDE Arduino com a coordenada do eixo X sendo transmitida

No software Matlab, foi usado um conjunto de funções que facilitam a comunicação via Bluetooth Low Energy.

```
listn= blelist;
```

A função "blelist" faz o scaneamento dos dispositivos Bluetooth LE

```
belt = ble("ESP32 ");
```

Com a função "ble" o dispositivo bluetooth é conectado ao Matlab.

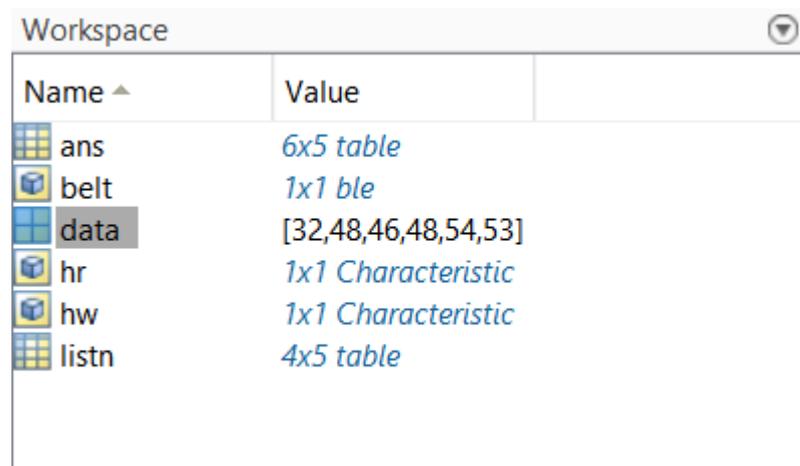
```
hr = characteristic (belt, "6E400001-B5A3-F393-E0A9-E50E24DCCA9E", "6E400003-B5A3-F393-E0A9-E50E24DCCA9E");
```

A função “characteristic” acessa as características do dispositivo que está enviando os dados.

```
data = read(hr);
```

A função “read” lê os dados recebidos e os armazena na variável data.

Utilizando as funções apresentadas podemos receber facilmente os dados emitidos pelo ESP-32. A seguir, é apresentado os dados que foram enviados pelo ESP32 e recebidos pelo Matlab, no caso a coordenada do eixo X em código ASC II.

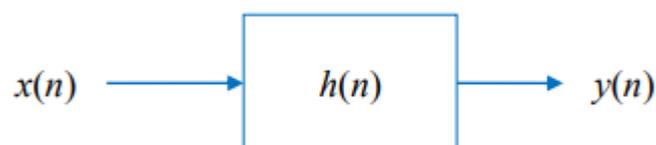


Name ^	Value
ans	6x5 table
belt	1x1 ble
data	[32,48,46,48,54,53]
hr	1x1 Characteristic
hw	1x1 Characteristic
listn	4x5 table

Figura 11: Coordenada do eixo X recebida pelo Matlab em código ASC II

Já recebidos pelo Matlab, os dados passaram por tratamentos de processamento digital de sinais onde serão aplicadas as HRTFs como filtros do tipo o FIR e IIR.

Para realizar tal filtragem, devemos retomar o conceito de convolução. Podemos dizer que o processo de convolução consiste, basicamente, na combinação de dois sinais. Por exemplo, seja uma sequência discreta $x[n]$ e $h[n]$ é uma resposta ao impulso de um sistema linear invariante no tempo, a saída deste sistema resultaria em uma convolução entre $x[n]$ e $h[n]$, denominada $y[n]$, como podemos ver no diagrama abaixo.



Quando $h[n]$ corresponde à resposta ao impulso de um filtro digital, a convolução corresponde a um processo de filtragem e a saída desse processo pode ser obtida no Matlab por:

$$y = \mathit{conv}(h, x)$$

Ou

$$y = \mathit{filter}(h, 1, x)$$

Na aplicação de áudio 3D, $x(n)$ é a fonte sonora (um arquivo contendo a gravação do som a ser emitido tridimensionalmente) e $h(n)$ é uma das duas HRTFs selecionadas. A saída é o som que será emitido em um dos canais do fone estéreo sendo utilizado. A seleção de quais HRTFs que serão usadas é baseada na informação recebida via Bluetooth. Assim, utilizamos duas HRTFs: uma representando a propagação do som desde a fonte até o ouvido direito e outra para o ouvido esquerdo.

Desta forma, o som tridimensional é simulado.

5. CONSIDERAÇÕES FINAIS

Este estudo propõe o desenvolvimento de um sistema de baixo custo que facilite o acesso ao áudio imersivo. Por meio do uso de um microcontrolador e um sensor acelerômetro, foi possível encontrar a posição da cabeça do usuário e enviar tais dados via Bluetooth Low Energy para serem tratados no software Matlab pelos métodos de processamento digital de sinais (PDS) e desenvolver uma maior ambientação sonora ao usuário. Devido a pandemia, não foi possível realizar a simulação do HRTF devido à falta de testes. Conclui-se que é possível desenvolver um sistema que contribua para o avanço das tecnologias de áudio imersivo e que tais tecnologias possam seguir com seu avanço na mesma medida que os sistemas de realidade virtual (RV) e realidade aumentada (RA).

6. REFERÊNCIAS

Algazi V.R., Duda R. O., Thompson D. M. e Avendano C., "TheCIPIC HRTF database," Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat.No.01TH8575), New Platz, NY, 2001, pp. 99-102.

Alley T., Computer Graphics Knowledge Base Report Nov 2017 , disponível em <http://education.siggraph.hosting.acm.org/resources/knowledge-base/report> , acessado em 24/03/2019.

Carlile S., Virtual auditory space: generation and application. New York: Springer, 1996.

Carrogi-Viana D., Análise de medidas de aceleração do movimento na prática de jogos virtuais em adolescentes com síndrome de Down. Dissertação de mestrado, Universidade Presbiteriana Mackenzie, 2013.

Csikszentmihalyi M. , Finding Flow: The Psychology of Engagement with Everyday Life, Basic Books (AZ); Revised ed. 1998 , ISBN-10: 9780465024117, 192 páginas

Diniz P.S.R., SILVA E.A.B, e Netto S.L.. Processamento digital de sinais. 2. ed. Porto Alegre [Brasil]: Bookman, 2003.

Faller II K., Barreto A., Gupta N. e Rishé, N., Time and Frequency Decomposition of Head-Related Impulse Responses for the Development of Customizable Spatial Audio Models., WSEAS Transactions on Signal Processing 2 (11), 1465-1472, 2006

González-Campos J,S., Arnedo-Moreno J.S. e Sánchez-Navarro J., The Use of Digital Games to Teaching Computer Graphics: an Open Opportunity, IEEE Global Engineering Education Conference, p. 1988 - 1996, 2018.

Haykin S. S. e Veen, B.V. Sinais e sistemas, Bookman, 2001. ISBN 8573077417.

Head-related transfer function. En.wikipedia.org. Disponível em: <https://en.wikipedia.org/wiki/Head-related_transfer_function>. Acesso em: 11 abr. 2019

Johansson M., VR for your ears, IEEE Spectrum, p. 24-29, 2019.

Kay, S. (1993). Fundamentals of statistical signal processing. Prentice-Hall, 1st ed. 1993

Lanz O. e Brunelli R., (2007). Tracking of Head Location and Pose from Low-Resolution Video. In: Stiefelhagen R., Bowers R., Fiscus J. (eds) Multimodal Technologies for Perception of Humans. RT 2007, CLEAR 2007. Lecture Notes in Computer Science, vol 4625. Springer, Berlin, Heidelberg .

Mora-Lumbreras M. A., Flores-Pulido L. e González-Contreras B., Incorporating 3D Sound in Different Virtual Worlds. 2012 8º Latin American Web Congress, p. 95-99, 2012.

Usman M., Kamal K. e Qayyum, R. 3D Sound Generation Using Kinect and HRTF. International Conference on Signal and Image Processing, v. 2, n. CFP17G14-ART, p. 307-310, 2019.

Wijman, T. (2019). *Global Games Market Revenues 2018 | Per Region & Segment | Newzoo*. [online] Newzoo. Available at: <https://newzoo.com/insights/articles/global-games-market-reaches-137-9-billion-in-2018-mobile-games-take-half/> [Accessed 11 Apr. 2019].

SANTOS, Rui. **Getting Started with ESP32 Bluetooth Low Energy (BLE) on Arduino IDE**. Disponível em: <https://randomnerdtutorials.com/esp32-bluetooth-low-energy-ble-arduino-ide/>. Acesso em: 23 fev. 2021.

Contatos: luciano47junior@gmail.com e paulo.lopes@mackenzie.br