ROMEO BULLA JUNIOR

Ferramenta de Áudio Conferência Espacial Implementando Conceitos
de Realidade Aumentada

Dissertação apresentada à Escola
Politécnica da Universidade de São
Paulo para obtenção do título de
Mestre em Engenharia Elétrica.

São Paulo
2009

ROMEO BULLA JUNIOR

Ferramenta de Áudio Conferência Espacial Implementando Conceitos
de Realidade Aumentada

Dissertação apresentada à Escola
Politécnica da Universidade de São
Paulo para obtenção do título de
Mestre em Engenharia Elétrica.

Área de Concentração:
Sistemas Digitais.

Orientador: Prof. Dr. Wilson Vicente
Ruggiero.

São Paulo
2009

AGRADECIMENTOS

Primeiramente ao Prof. Dr. Wilson Ruggiero, da EP-USP, orientador deste trabalho.

À Fundação de Amparo à Pesquisa no Estado de São Paulo (FAPESP) que, por meio de bolsa de trabalho técnico e do suporte financeiro ao projeto Ae do programa Tidia, possibilitou a pesquisa e implementação relatada neste trabalho.

Ao Prof. Dr. Romero Tori, da EP-USP, pelos ensinamentos acerca dos conceitos de realidade aumentada e virtual, além das contribuições fornecidas durante a fase de qualificação deste trabalho.

Ao Prof. Dr. Hécio Onusic, do IF-USP, pela introdução aos conceitos de acústica e psicoacústica.

À Profa. Dra. Graça Bressan, da EP-USP, pelas contribuições fornecidas durante a fase de qualificação e apresentação do presente trabalho.

Ao Prof. Dr. Cesar Augusto Camillo Teixeira, da UFSCar, pelas contribuições fornecidas durante a fase de apresentação deste trabalho.

À Dra. Luísa Aleyda Gonzáles, ao Me. Fernando Almeida e à Vivian Alves Bulla que forneceram valiosas contribuições na redação deste.

RESUMO

Este trabalho apresenta uma ferramenta para conferência de áudio 3D (espacial) implementando conceitos de Realidade Aumentada (RA). O objetivo desta ferramenta é aprimorar a sensação de presença e melhorar a interatividade entre seus participantes remotos, por meio de benefícios proporcionados pela utilização de técnicas de áudio espacial (implementadas em avatares de áudio) pela: maior facilidade de concentração e atenção em um único participante e pelos efeitos positivos na memorização dos conteúdos pelos participantes como consequência da melhor inteligibilidade e compreensão. A motivação desta implementação reside em sua utilização como ferramenta de comunicação síncrona no ambiente de aprendizagem eletrônica Tidia-Ae, auxiliando na realização de atividades colaborativas e, possivelmente, nos processos de ensino e aprendizagem à distância. A ferramenta implementada foi integrada ao sistema Tidia-Ae e os resultados dos experimentos realizados demonstraram sua efetividade com relação às melhorias proporcionadas pelo processamento de áudio espacial.

Palavras-chave: Teleconferência. Psicoacústica. Realidade Aumentada. Áudio Digital. Aprendizado eletrônico.

ABSTRACT

This work presents a 3D (spatial) audio conference tool implementing Augmented Reality (AR) concepts. The main intent of this tool is to enhance the sense of presence and increase the interactivity among remote participants, by implementing spatial audio techniques in audio avatars. The use of such techniques facilitates focusing the attention on anyone specific participant of the conference and enables a positive effect on memory retention, resulting in a better intelligibility and comprehension. The motivation of this implementation lies on its appliance as a synchronous communication tool within the Tidia-Ae e-Learning system, thus aiding on collaborative activities realization and, possibly, on teaching and learning processes. The results of the experiments showed the effectiveness provided by the spatial audio processing when applied in such environment.

Keywords: Teleconference. Psychoacoustics. Augmented Reality. Digital Audio. E-learning.

LISTA DE ILUSTRAÇÕES

Figura 1 - Difração de onda de baixa frequência em um canto.....	28
Figura 2 - Difração de onda de alta frequência em um canto.....	28
Figura 3 - Difração de onda de baixa frequência em uma abertura.	28
Figura 4 - Difração de onda de alta frequência em uma abertura.	28
Figura 5 - Participante local em seu ambiente real, prestes a entrar em uma conferência.....	34
Figura 6 - Ambiente do participante local enriquecido com os avatares (de áudio somente) dos conferencistas remotos.....	34
Figura 7 – Captação, transmissão e processamento de Buffers de Áudio nos programas clientes (cor verde clara).	43
Figura 8 – Pontos 3D representando os ouvidos (E e D), nariz (N) e centro da cabeça (C) do participante receptor e boca do participante emissor (B).....	47
Figura 9 – Ângulos para análise de incidência direta de onda sonora no ouvido direito do participante receptor.	48
Figura 10 – Situação onde o ângulo β é menor que o ângulo θ	50
Figura 11 – Ângulo β é igual ao ângulo θ	50
Figura 12 – Ângulo β é maior que ângulo θ	50
Figura 13 – Ângulo NCB menor que 90°	53
Figura 14 – Ângulo NCB maior que 90°	53
Figura 15 – Ângulo CDB, para ouvido direito, caso a fonte de áudio esteja no hemisfério traseiro.....	53
Figura 16 - Diagrama da Ferramenta de comunicação de áudio, com representação do programa servidor e duas instâncias do programa cliente.	56
Figura 17 - Tela capturada de uma sala de apresentação no <i>worksite</i> “Teste de Áudio 3D”.	60
Figura 18 – Tela capturada com a sala de reunião do <i>worksite</i> “Teste de Áudio 3D” selecionada.	61
Figura 19 – Tela capturada após o usuário mudar do <i>worksite</i> “Teste de Áudio 3D” para o “PCS2307_2Sem_2009_Turma50”.	62

Figura 20 – Imagem ampliada da disposição dos participantes artificiais (Laerte, Paulo, Debora e Oswaldo) em relação ao usuário avaliador (Romeo neste caso). ..65

LISTA DE GRÁFICOS

Gráfico 1 - Gráfico contendo os dados da média de contagem dos participantes artificiais identificados pelos participantes testados.	69
Gráfico 2 - Gráfico contendo os dados da média do grau de dificuldade de concentração encontrada pelos participantes.	69
Gráfico 3 – Média da quantidade de palestrantes percebidos.	78
Gráfico 4 – Média do grau de certeza da contagem de palestrantes percebidos.....	79
Gráfico 5 – Média da dificuldade enfrentada para manter a concentração no palestrante indicado.	80
Gráfico 6 – Média para a quantificação aproximada do tempo de compreensão literal.....	81
Gráfico 7 – Média da compreensão de conteúdo percebida pelo participante.	82
Gráfico 8 – Representação gráfica da percepção de funcionamento de processamento espacial.....	83
Gráfico 9 – Acertos e erros para as questões de compreensão contextuais 5.4.1.7 e 5.4.1.8.	84
Gráfico 10 – Palestrantes percebidos nas etapas 1, 2 e 3 (destacado).	86
Gráfico 11 – Acertos e erros para a questão 5.4.1.9.....	87

LISTA DE TABELAS

Tabela 1 - Tempos de resposta e suas interpretações de qualidade (ITU-T G.114).	38
Tabela 2 – Graduação do ângulo α com amostras utilizadas e frequência de corte.	52
Tabela 3 - Resultados da contagem de participantes artificiais.....	68
Tabela 4 - Resultados do grau de dificuldade encontrado.	68
Tabela 5- Questões e etapas do segundo experimento.....	75
Tabela 6 – Dados estatísticos das amostras da questão 5.4.1.1 (Etapas 1 e 2).....	77
Tabela 7 – Dados estatísticos das amostras da questão 5.4.1.2 (Etapas 1 e 2).....	78
Tabela 8 – Dados estatísticos das amostras da questão 5.4.1.3 (Etapas 1 e 2).....	79
Tabela 9 – Dados estatísticos das amostras da questão 5.4.1.4 (Etapas 1 e 2).....	80
Tabela 10 – Dados estatísticos das amostras da questão 5.4.1.5 (Etapas 1 e 2).....	81
Tabela 11 – Amostras da questão 5.4.1.6 (Etapas 1 e 2)	82
Tabela 12 – Amostras da questão 5.4.1.7 (Etapa 1 somente)	83
Tabela 13– Amostras da questão 5.4.1.8 (Etapa 2 somente)	83
Tabela 14 – Respostas corretas à questão 5.4.1.7 (Etapa 1 somente).....	84
Tabela 15 – Dados estatísticos das amostras da questão 5.4.1.1 (Etapa 3)	85
Tabela 16 – Amostras da questão 5.4.1.9 (Etapa 3).....	86
Tabela 17 – Respostas corretas à questão 5.4.1.9 (Etapa 3)	86
Tabela 18 – Dados estatísticos das amostras da questão 5.4.1.10 (Etapa 3)	87

LISTA DE SIGLAS

AE	Aprendizagem Eletrônica
API	<i>Application Programming Interface</i>
dB	Decibel
HTTP	<i>Hypertext Transfer Protocol</i>
IID	<i>Interaural Intensity Differences</i>
ITD	<i>Interaural Time Differences</i>
ITU-T	<i>International Telecommunication Union - Standardization Sector</i>
SPL	<i>Sound Pressure Level</i> ou nível de pressão sonora
SWL	<i>Sound Power Level</i> ou nível de potência sonora
SIL	<i>Sound Intensity Level</i> ou nível de intensidade sonora
TCP	<i>Transfer Control Protocol</i>
RA	Realidade Aumentada
RTP	<i>Real-time Transport Protocol</i>
RV	Realidade Virtual
SGBD	Sistema Gerenciador de Banco de Dados
TIDIA	Tecnologia da Informação para o Desenvolvimento da Internet Avançada
UDP	<i>User Datagram Protocol</i>

SUMÁRIO

1 INTRODUÇÃO	13
1.1 OBJETIVO.....	13
1.2 MOTIVAÇÃO.....	14
1.3 MÉTODO.....	15
1.4 ORGANIZAÇÃO DO DOCUMENTO	15
2 RA APLICADA NA EDUCAÇÃO	18
2.1 SOBRE RA.....	18
2.1.1 Definição de RA.....	18
2.1.2 Objetivo de RA.....	19
2.1.3 Aplicações de RA	19
2.2 SOBRE A EDUCAÇÃO	21
2.3 PAPEL DAS MÍDIAS E TECNOLOGIAS NA EDUCAÇÃO.....	21
2.4 PAPEL DA RA APLICADA À EDUCAÇÃO.....	22
2.5 RECURSOS DE RA APROVEITADOS	23
2.6 EXEMPLOS DE FERRAMENTAS.....	23
3 ÁUDIO ESPACIAL (3D).....	25
3.1 CARACTERÍSTICAS DE ÁUDIO E ONDAS SONORAS (ACÚSTICA)	25
3.1.1 Propagação.....	25
3.1.2 Comprimento de onda	26
3.1.3 Velocidade do som.....	26
3.1.4 Freqüência	26
3.1.5 Escala para medidas de intensidade de vibrações	26
3.1.6 Reflexão e Refração	27
3.1.7 Difração	28
3.2 PERCEPÇÃO ESPACIAL DE ÁUDIO	29
3.2.1 Diferença de Tempo Interauricular - ITD (<i>Interaural Time Differences</i>)	29
3.2.2 Diferença de Intensidade Interauricular - IID (<i>Interaural Intensity Differences</i>).....	30
3.2.3 Sombra Acústica	30
3.3 DISPOSITIVOS TECNOLÓGICOS	31

3.4 ÁUDIO ESPACIAL E MEMORIZAÇÃO	32
4 FERRAMENTA DE ÁUDIO CONFERÊNCIA ESPACIAL (3D).....	34
4.1 PROPOSTA	34
4.2 PRINCIPAIS FUNCIONALIDADES	35
4.3 PROJETO	37
4.3.1 Considerações sobre bibliotecas para processamento de áudio em tempo real.....	38
4.3.2 Considerações sobre microfones e fones de ouvido	40
4.3.3 Processamento de áudio tridimensional.....	41
4.3.3.1 Filtro de passa baixa (<i>Low pass</i>).....	44
4.3.3.2 Filtro de atraso (<i>Delay</i>)	44
4.3.3.3 Filtro de intensidade (Volume).....	45
4.3.4 Modelo para cálculo e transferência de som espacial (3D)	45
4.3.4.1 Configuração do Atraso.....	48
4.3.4.2 Configuração de Passa-Baixa	51
4.3.4.3 Configuração de Intensidade.....	54
4.4 ARQUITETURA DA FERRAMENTA	56
4.5 INTEGRAÇÃO COM O SISTEMA TIDIA-AE.....	58
5 EXPERIÊNCIA DE UTILIZAÇÃO DE ÁUDIO ESPACIAL EM AMBIENTES DE APRENDIZAGEM ELETRÔNICA	64
5.1 CONDIÇÕES GERAIS PARA A EXECUÇÃO DOS EXPERIMENTOS	64
5.2 EXPERIMENTOS	66
5.3 EXPERIMENTO DE SIMPLES IDENTIFICAÇÃO E CONCENTRAÇÃO DE FONTES DE ÁUDIO	66
5.3.1 Questionários e observações para os participantes.....	67
5.3.1.1 Verificação de número de fontes de áudio	67
5.3.1.2 Capacidade de concentração nas fontes de áudio.....	67
5.3.2 Coleta de Dados e Resultados Obtidos.....	68
5.4 EXPERIMENTO INTEGRADO DE CONTAGEM, IDENTIFICAÇÃO E COMPREENSÃO	70
5.4.1 Questionários e observações para os participantes.....	71
5.4.1.1 Verificação de contagem de fontes de áudio.....	72
5.4.1.2 Grau de certeza sobre a quantificação de palestrantes fornecida.....	72
5.4.1.3 Grau de dificuldade de concentração nas fontes de áudio.....	72

5.4.1.4 Compreensão literal	72
5.4.1.5 Compreensão geral de um palestrante	73
5.4.1.6 Percepção de posicionamento espacial das trilhas de áudio (palestrantes artificiais)	73
5.4.1.7 Compreensão contextual sobre Química (Alotropia)	74
5.4.1.8 Compreensão contextual sobre Física (Óptica)	74
5.4.1.9 Compreensão contextual sobre o local de trabalho dos palestrantes	75
5.4.1.10 Eficiência do Áudio Espacial	75
5.4.2 Coleta de Dados e Resultados Obtidos.....	76
5.4.2.1 Análise de confronto direto entre as etapas 1 e 2	77
5.4.2.2 Análise da Etapa 3	85
5.4.2.3 Observação Geral de Todas as Etapas.....	88
6 CONCLUSÃO E TRABALHOS FUTUROS	90
7 REFERÊNCIAS.....	92

1 INTRODUÇÃO

A área de pesquisa denominada Realidade Aumentada (RA) tem como proposta principal estender ou amplificar o mundo real através da adição de elementos artificiais e da remoção (oclusão) e alteração de elementos reais. Essa forma de extensão do mundo real comprovadamente traz benefícios a diversas áreas (conforme já relatado por (AZUMA, 1997)), como por exemplo na medicina quando utilizada no auxílio a diagnósticos, como também na engenharia, aviação, educação, robótica, entretenimento, entre outras.

A área de aplicação abordada neste trabalho é a educacional através da utilização de uma ferramenta de áudio conferência 3D para ambientes de aprendizagem eletrônica onde a interatividade e colaboração são pilares essenciais para o pleno aprendizado. O objetivo pelo uso dos conceitos de RA por meio de técnicas de áudio 3D é aumentar a sensação de presença dos participantes para enriquecer o processo de aprendizagem.

Para comprovar o potencial da ferramenta de comunicação de áudio 3D são realizados experimentos utilizando-se uma implementação da ferramenta aplicada a estudantes em um ambiente educacional real.

1.1 OBJETIVO

O principal objetivo desta pesquisa é avaliar a utilização de técnicas de áudio conferência espacial (3D) implementando conceitos de RA para utilização em ambientes de aprendizagem eletrônica a fim de melhorar a sensação de presença e eficiência da interatividade entre os usuários.

O escopo desta pesquisa também inclui a avaliação de quais são, como funcionam e interagem as características técnicas relacionadas à percepção espacial do áudio. Ou seja, localização tridimensional, direção e intensidade do som em ambientes reais com o intuito de estudar as tecnologias, processos e algoritmos envolvidos.

Para possibilitar essas avaliações é que se fazem fundamentais os

experimentos relacionados à criação artificial de ambientes que implementem ou simulem, com recursos tecnológicos, as características de áudio existentes nos ambientes reais, como parte de um ambiente colaborativo educacional.

1.2 MOTIVAÇÃO

Sabe-se que o processo educacional se orienta por práticas pedagógicas que possuem como aspecto central a interação entre o instrutor e os aprendizes (JONASSEN; PECK; WILSON, 1999) e que os processos interativos e colaborativos dentro do contexto social do aprendiz aumentam muito a eficácia dos processos educacionais (JONASSEN; PECK; WILSON, 1999).

Este contexto, aliado à necessidade de um processo educacional de maneira continuada e de aplicação possível em larga escala, coloca as técnicas de aprendizagem eletrônica – presencial ou à distância – como uma das únicas alternativas viáveis e compatíveis com este objetivo.

É importante notar que num ambiente de aprendizagem eletrônica à distância existe o desafio de manter um processo interativo e colaborativo. A utilização de técnicas de áudio 3D representa uma tentativa de melhorar essa interação nesses ambientes.

Entre as melhorias proporcionadas pela utilização de técnicas de áudio espacial, destaca-se a concentração da atenção do aprendiz, num participante específico imerso em um ambiente com múltiplas fontes de áudio (SCHAMNDT ; MULLINS, 1995), além dos efeitos positivos na retenção de conteúdos trocados pelos participantes, como consequência da melhor inteligibilidade e compreensão (BALDIS, 2001). Essas melhorias podem, assim, ser aplicadas para o enriquecimento de uma interação remota.

A motivação desta pesquisa está centrada nesta linha de raciocínio, que procura demonstrar a viabilidade da implementação de tais ferramentas e avaliar sua eficiência em comparação aos processos interativos e colaborativos comumente aplicados aos ambientes educacionais e que não utilizam áudio espacial.

1.3 MÉTODO

Para que os objetivos definidos fossem alcançados, determinadas etapas se fizeram necessárias à execução desta pesquisa. As principais estão relacionadas abaixo:

- **Levantamento bibliográfico:** Documentos e publicações relevantes que suportam as informações obtidas e direcionaram a pesquisa;
- **Estudo de acústica:** Estudo das características básicas de áudio, como uma introdução para o estudo de percepção espacial de áudio;
- **Estudo de percepção espacial de áudio:** Estudo dos principais fatores de percepção de áudio humana relacionados com sensação de localização espacial;
- **Estudo de técnicas de áudio:** Estudo das técnicas existentes de processamento de áudio, a fim de observar quais podem e devem ser aplicadas na simulação de um ambiente tridimensional;
- **Estudo e análise de trabalhos relacionados:** Estudo e análise dos trabalhos relacionados à pesquisa;
- **Especificação e implementação de uma ferramenta:** Definição, especificação e implementação de ferramenta com o objetivo de facilitar a interação e colaboração em ambientes educacionais utilizando técnicas de áudio 3D;
- **Realização de experimentos:** Avaliação qualitativa da efetividade da percepção espacial de áudio.

1.4 ORGANIZAÇÃO DO DOCUMENTO

A dissertação está organizada em cinco capítulos (capítulos 2 a 6) além desta introdução.

Primeiramente, são abordados os conceitos básicos de RA e como eles são aproveitados em ambientes de aprendizagem eletrônica. A proposta da ferramenta que reúne os conceitos de percepção espacial de áudio e técnicas de áudio 3D é

descrita servindo de base para a implementação realizada nesta pesquisa.

Experimentos foram realizados para verificar o aumento da sensação de presença nos participantes e sua possível eficácia em ambientes de aprendizagem eletrônica. Estes resultados foram avaliados, comprovando a eficácia das técnicas utilizadas nesta dissertação.

A estrutura norteada pelas etapas do método utilizado neste trabalho é apresentada a seguir:

Capítulo 2: Realidade Aumentada Aplicada à Educação. Introduz os conceitos básicos sobre RA e educação. A finalidade é demonstrar como os conceitos de RA podem ser implementados em aplicações contextualizadas para ambientes educacionais eletrônicos. Ao final do capítulo, são apresentados exemplos de aplicações de RA e educação.

Capítulo 3: Áudio Espacial 3D. Este capítulo descreve os aspectos físicos sobre acústica e psicoacústica de um modo geral, a fim de facilitar a compreensão dos assuntos de percepção espacial.

Ele introduz a percepção da localização espacial de áudio descrevendo, resumidamente, como ocorre nossa percepção de áudio espacial, ou seja, como somos capazes de identificar a localização de uma fonte de áudio, e quais as características técnicas associadas à essa percepção.

São apresentados alguns dispositivos tecnológicos possíveis de serem utilizados na construção de aplicações que implementem áudio espacial.

Este capítulo termina com a apresentação de duas pesquisas acerca da utilização de áudio espacial e seus efeitos na memorização, totalmente relevantes a esta pesquisa.

Capítulo 4: Ferramenta de Áudio Conferência Espacial (3D). Apresenta a proposta e os requisitos da ferramenta de áudio conferência espacial (3D) para utilização em ambientes educacionais eletrônicos. Também são detalhadas a solução computacional e a arquitetura elaborada para a construção da ferramenta de áudio conferência 3D. No detalhamento da arquitetura, é explanado como funciona a integração ao sistema Tidia-Ae.

Capítulo 5: Experiência de Utilização de Áudio 3D em Ambientes Educacionais Eletrônicos. Define e descreve os experimentos realizados destacando seus objetivos, as dificuldades encontradas e as soluções utilizadas

durante a realização dos experimentos.

Os dados dos experimentos foram coletados, analisados e validados. São apresentados neste capítulo para servirem de sustentação à comprovação das hipóteses desta pesquisa.

Capítulo 6: Conclusão e Trabalhos Futuros. Considerações finais sobre a pesquisa, com observações sobre a ferramenta e possibilidades de trabalhos futuros.

2 RA APLICADA NA EDUCAÇÃO

Uma vez que a ferramenta proposta neste trabalho implementa conceitos de Realidade Aumentada (RA) e foi concebida para utilização em um sistema de aprendizagem eletrônica, torna-se necessário conceituar um tema essencial que norteou sua fundação: RA aplicada na educação.

Sendo assim, o objetivo deste capítulo é primeiro apresentar RA para, em seguida, demonstrar como os conceitos desta área de pesquisa podem ser implementados em programas (*software*) educacionais.

2.1 SOBRE RA

Segundo (KIRNER; TORI, 2004), a definição de Realidade Aumentada ascende do conceito de Realidade Virtual (RV), entretanto com uma grande diferença: na RA, ao contrário da RV, um indivíduo não está imerso em um ambiente completamente sintético, isto é, ele está no mundo real enriquecido e alterado com elementos e objetos artificiais.

Essa alteração do mundo real é efetuada por meio de dispositivos tecnológicos que permitem adicionar objetos tridimensionais artificiais, bem como esconder objetos reais existentes no ambiente real (conceito de oclusão). O ambiente resultante destas alterações e, portanto enriquecido, possui características de RA.

2.1.1 Definição de RA

A fim de que um ambiente seja definido como de RA, somente distorções e alterações no ambiente real não são o suficiente. Para tal, vêem-se necessárias três importantes regras, utilizadas e aceitas pelos pesquisadores da área, que (AZUMA, 1997) define como:

- Mistura de elementos reais e artificiais:

Elementos artificiais devem estar presentes no ambiente real a ser estendido. Isto significa que os elementos artificiais e os elementos reais coexistem em um único ambiente simultaneamente. É fundamental observar que o sentido de presença do usuário esteja focado no mundo real, apesar da extensão do mesmo com os elementos virtuais.

- Registro tridimensional:

Parte dos objetos reais existentes em um ambiente de RA, considerados relevantes ao sistema, devem ser mapeados em três dimensões (3D) e rastreados por meio de alguma técnica de rastreamento. Já os objetos artificiais devem obrigatoriamente ser mapeados em 3D.

- Interação em tempo real:

A interação do usuário com os sistemas de RA deve, necessariamente, ser feita em tempo real. Isto significa que o sistema deve ser reativo às ações do usuário ou dos meios de entrada, no momento em que estes ocorrem. Filmes de ficção científica onde são inseridos elementos artificiais, não devem ser considerados ambientes de RA, pois a interação não se dá em tempo real.

2.1.2 Objetivo de RA

O objetivo perseguido por sistemas de RA é prover um ambiente real alterado com objetos artificiais minuciosamente integrados, de forma que o usuário perceba e sinta este ambiente como sendo um só ao invés de dois ambientes distintos. Novamente, é importante ressaltar que o sentido de presença do usuário está no ambiente real e não no artificial.

2.1.3 Aplicações de RA

Os sistemas de RA podem ser largamente aplicados em diferentes áreas, sejam acadêmicas ou comerciais. Não importando qual seja o segmento, a finalidade é sempre prover auxílio a alguma atividade no mundo real. Alguns

exemplos podem ser citados:

- Medicina

Sistemas de RA são utilizados para viabilizar diagnósticos, como por exemplo a visualização de um feto em desenvolvimento no útero da mãe. Outra possível aplicação é o uso em cirurgias, possibilitando ao cirurgião obter informações adicionais em tempo real como a localização precisa de tumores.

- Engenharia e Arquitetura

É possível visualizar construções tridimensionais em pequena escala e com detalhes minuciosos, o que permite verificar a obra à distância.

Outra grande aplicação é a utilização de RA com a finalidade de exibir informações sobre instalações de fábricas, facilitando o trabalho de técnicos e engenheiros em campo.

- Entretenimento

Os usuários ganharam muitas possibilidades com os jogos que utilizam sistemas de RA. Esses programas permitem aos jogadores caminhar pelos cenários no mundo físico e ter dados e complementações fornecidas pelos dispositivos de RA, o que aumenta a intensidade da experiência.

- Educação

Na educação surgiram muitas possibilidades. Pode-se citar como um exemplo bem-sucedido o MagicBook (BILLINGHURST; KATO, 2001). Nessa implementação, o usuário aprende enquanto visualiza as formas tridimensionais exibidas pelo livro conforme avança a leitura das páginas.

As aplicações de RA aparecem em diversas áreas do conhecimento porém a educação, em especial, pode ser muito enriquecida e beneficiada com programas que implementem esses conceitos. E esta possibilidade é a principal motivação do presente trabalho.

Contudo, a fim de apresentar o aproveitamento dos conceitos de RA em ferramentas educacionais, é preciso antes compreender a educação e as mídias para a propagação de conhecimento.

2.2 SOBRE A EDUCAÇÃO

O processo educacional trata de como os conhecimentos são produzidos, construídos ou transmitidos entre os indivíduos e como persistem ao longo do tempo. Segundo (LANDIM, 1997), refere-se à prática educativa e ao processo de ensino-aprendizagem que leva o aprendiz a aprender a aprender, a saber pensar, a criar, a inovar, a construir conhecimentos, a participar ativamente de seu próprio crescimento.

E para que os conhecimentos não se percam, são necessárias mídias e técnicas para sua propagação. Sobre as mídias, pode-se citar o livro como o mais antigo e conhecido pela humanidade. O livro permitiu, durante milhares de anos, a persistência e armazenamento de conhecimento por meio da escrita (KOZMA, 1991). Entretanto, somente as mídias não são suficientes. A interação entre seres humanos é imprescindível e é a partir dela que surgem as técnicas de ensino e aprendizagem.

2.3 PAPEL DAS MÍDIAS E TECNOLOGIAS NA EDUCAÇÃO

A partir dos avanços científicos e tecnológicos, principalmente na área da informática e telecomunicações, novas mídias surgiram e tal como o livro podem ser utilizadas para a educação. Dentre elas, as que mais se destacam são: televisão, rádio e hipertextos (*Internet*).

Esses avanços permitiram a implementação de ensino à distância, seja por meio de recursos eletrônicos ou não, como é o caso dos cursos por correspondência. Trabalhos como (TORI, 2003) abordam este tema com maior profundidade.

O importante é que a diversidade de mídias oferece benefícios significativos, dos quais os de maior destaque são:

- Maior abrangência de público-alvo

Uma vez que a oferta de mídias é diversificada, os aprendizes podem escolher as que mais lhe são atraentes e efetivas. Vale ressaltar que essa diversificação pode facilitar processos de inclusão social, contemplando portadores

de necessidades especiais (VICENTI et al., 2006).

- Enriquecimento da transmissão de conhecimento

As mídias podem auxiliar o ensino enriquecendo a representação dos conceitos por meio de ilustrações pictóricas, sons, vídeos e hipertextos. Isto cria novas possibilidades para facilitar a compreensão do conteúdo que se quer ensinar e demonstrar.

Observando essas duas características, pode-se avaliar o importante papel que as mídias exercem na educação ao proporcionarem recursos que enriquecem e melhoram os processos de ensino e aprendizagem.

2.4 PAPEL DA RA APLICADA À EDUCAÇÃO

Como uma nova proposta, o conjunto de tecnologias que implementam conceitos de RA pode ser considerado uma nova mídia e assim seu papel na educação torna-se semelhante ao de outras mídias.

Parte dessas tecnologias possui características ubíquas que podem ser aproveitadas para propiciar a aprendizagem por meio de dispositivos tecnológicos móveis e com poucos recursos computacionais, como é o caso de telefones celulares e *handhelds*.

Já a interatividade, uma de suas características principais, pode ser utilizada para estimular a aprendizagem. Sob a perspectiva de ser uma nova mídia, a RA no contexto de ensino pode proporcionar um efeito motivador.

As atividades colaborativas, também importantes para a educação, podem usufruir dos recursos oferecidos pela tecnologia de RA. Ao aproveitarem características de algumas das tecnologias que implementam conceitos de RA, como é o caso de interfaces humano-computador tangíveis, melhoraram a interação com a máquina e o aumento da sensação de presença dos usuários e um ambiente aumentado compartilhado.

O que se percebe destas considerações é que as novas tecnologias oferecem recursos significativos que devem ser aproveitados para a educação. Assim torna-se importante avaliar de que forma os conceitos de RA podem ser aplicados à

educação. Em especial, para este trabalho, considera-se a aplicação de áudio espacial em ambiente de aprendizagem eletrônica.

2.5 RECURSOS DE RA APROVEITADOS

Dentre os diversos recursos existentes em RA disponíveis para exemplificar sua aplicação na educação, são somente demonstrados neste estudo aqueles que são, de fato, relevantes à concepção da ferramenta de comunicação de áudio.

Ao se utilizar recursos de RA, o mundo real pode ser expandido além dos limites físicos. Isto significa que um ambiente de RA não está limitado às leis da física como o mundo real. Torna-se possível ignorar ou alterar essas leis para promover novas experiências educacionais, como por exemplo anular a lei da gravidade dos elementos artificiais.

No caso de uma ferramenta de áudio conferência 3D implementando conceitos de RA, cada participante da conferência interage com outros usuários inseridos artificialmente em seu ambiente aumentado, de modo que estes possam ocupar um espaço já em uso por um móvel ou objeto do ambiente real.

Portanto, há a possibilidade de se comunicar com outras pessoas sem que exista a necessidade de consistência desses avatares com o mundo real. Isso permite que um usuário possa trabalhar em alguma atividade colaborativa, como utilizar seu computador para editar um documento de texto colaborativo, com os outros autores simultaneamente editando o mesmo documento, mas reunidos artificialmente à sua volta. Esse caso exemplifica como a interação e a colaboração entre todos pode ser enriquecida.

2.6 EXEMPLOS DE FERRAMENTAS

Como já demonstrado por (BILLINGHURST; KATO, 2002), existe a possibilidade de executar estudos e trabalhos colaborativos por meio da utilização de ferramentas que implementam conceitos de RA. Munidos de alguns dispositivos tecnológicos, os alunos podem visualizar e estudar um objeto tridimensional que

representa a topologia de uma região, editá-lo e efetuar observações simultaneamente, sobre o mesmo objeto artificial utilizado.

Outro exemplo interessante é um jogo implementado por (WAGNER; BARAKONYI, 2003) que utiliza a tecnologia de RA. Na implementação, os jogadores têm a possibilidade de aprender alguns símbolos de escrita japonesa, mais especificamente, os símbolos *Kanji*.

Este jogo é um exemplo de aplicação capaz de ensinar divertidamente aos participantes os símbolos *Kanji*. Além disso, é um sistema de simples implementação, com poucos requisitos de *hardware* (*handhelds* com câmeras) e *software* (sistema ARToolkit (ARTOOLKIT, 2008) adaptado).

No contexto desta pesquisa o interesse é explorar o uso de comunicação de áudio com efeitos espaciais permitindo uma sensação tridimensional. As características de áudio que devem explorar os efeitos tridimensionais são apresentadas no próximo capítulo.

3 ÁUDIO ESPACIAL (3D)

3.1 CARACTERÍSTICAS DE ÁUDIO E ONDAS SONORAS (ACÚSTICA)

Segundo (HOWARD; ANGUS, 2006) e (KINSLEY et al, 1999), ondas sonoras são, na verdade, vibrações mecânicas propagadas em meios elásticos, como o ar, que os ouvidos humanos são capazes de perceber. Naturalmente, essas vibrações possuem características fundamentais, sem as quais não seria possível elucidar como ocorre a percepção espacial humana.

No contexto deste trabalho, são apresentadas somente as características de vibrações acústicas relevantes, a fim de possibilitar a compreensão da percepção espacial de áudio. Nas subseções 3.1.1 à 3.1.7 é introduzida cada uma delas seguida de uma breve explicação.

3.1.1 Propagação

Como observado em (ELSEA, 1995), o som é uma onda que se propaga por meio da movimentação ou vibração de moléculas de um meio elástico. O ar é um exemplo de meio elástico.

A propagação significa que a vibração de uma molécula é transmitida às moléculas e átomos vizinhos por meio da compressão ou rarefação do meio em questão. A propagação também diminui de intensidade de acordo com a distância percorrida, pois sua energia é dissipada à medida que ela se expande no meio, ocupando uma área maior. Este fenômeno de propagação segue a lei do inverso do quadrado da distância, no caso de uma propagação esférica (HOWARD; ANGUS, 2006, p.28), caracterizando o que convencionou-se chamar de uma onda sonora.

3.1.2 Comprimento de onda

A partir da observação dos momentos de compressão e rarefação, pode-se determinar o comprimento de uma onda. A pressão em um determinado ponto do espaço tem sua pressão aumentada ou diminuída conforme a onda sonora o atinge, revelando os pontos de máximo e mínimo de pressão. A partir do intervalo de tempo que a onda leva para revelar dois pontos máximos, a frequência (f) e o comprimento de onda (λ) são determinados.

3.1.3 Velocidade do som

Uma característica importante para a percepção espacial de áudio é a velocidade do som no ar (V_s), que pode variar de acordo com a densidade e elasticidade do meio elástico. A velocidade também varia dependendo do sentido do movimento da onda, isto é, se for longitudinal, será mais rápida que uma onda transversal.

3.1.4 Frequência

A frequência e o comprimento de onda caracterizam e quantificam as oscilações entre as compressões e rarefações, no intervalo de um segundo. A frequência relaciona-se com o comprimento de onda e com a velocidade do som de acordo com a equação:

$$V_s \text{ (m/s)} = f \text{ (Hz)} \times \lambda \text{ (m)}$$

3.1.5 Escala para medidas de intensidade de vibrações

A escala comumente utilizada para medir a intensidade de vibrações é o decibel (dB) (KINSLEY et al, 1999, p. 130). Existem três modos de medir em decibéis: nível de pressão sonora (*Sound Pressure Level* - SPL), nível de potência

sonora (*Sound Power Level* - SWL) e nível de intensidade sonora (*Sound Intensity Level* - SIL).

Não importando a medida utilizada, seu cálculo final é sempre relativo a outro valor, que é baseado no valor mínimo audível por um ser humano e, também, o valor de referência para cálculos. Este, é definido como um pico-watt por metro quadrado (10^{-12} Watt / m²) no caso do SIL, um pico-watt (10^{-12} Watt) no caso do SWL e vinte micro-pascals (20×10^{-6} Pa) no caso do SPL.

De acordo com (HOWARD; ANGUS, 2006, p.17), a medida mais utilizada é a SPL. Isto decorre do fato de SPL ser medida e calculada a partir de pressão, processo semelhante ao realizado pelos ouvidos humanos. O cálculo do decibel é realizado a partir da relação de pressão em uma escala logarítmica, conforme a equação abaixo:

$$SPL = 20 \times \log_{10} \left(\frac{P_{atual}}{P_{referência}} \right)$$

Onde:

SPL = Nível de pressão sonora ou *Sound Pressure Level* em decibéis (dB).

P_{atual} = Pressão atual em Pascals (Pa) ou Newtons por metro quadrado (N/m²), (1 Pa é igual a 1 N/m²).

$P_{referência}$ = Pressão de referência (20×10^{-6} Pa).

3.1.6 Reflexão e Refração

Estes efeitos ocorrem quando uma onda sonora, trafegando em um meio elástico, atinge a superfície divisória com outro meio elástico. Dependendo do ângulo de incidência da onda com a normal entre os meios, podem ocorrer reflexão, refração (transmissão) ou ambos. Este fenômeno físico segue a lei de Snell (KINSLEY et al, 1999, p. 155).

3.1.7 Difração

Uma onda sonora pode contornar objetos que sejam aproximadamente dois terços o comprimento de onda, ou maior. Portanto, a difração depende do tamanho do objeto e o comprimento de onda em questão. Considerando os casos onde o objeto é maior que dois terços do comprimento de onda, quanto maior a frequência (menor o comprimento de onda), menor o contorno. No entanto, se o objeto for maior que dois terços do comprimento de onda, a difração não terá efeito significativo (HOWARD; ANGUS, 2006, p.49).

Uma onda sonora também pode se expandir pelos cantos de uma abertura. Quanto maior o comprimento de onda (menor a frequência) em relação ao tamanho da abertura, maior a difração.

As Figuras 1, 2, 3 e 4 extraídas de (HOWARD; ANGUS, 2006), ilustram o conceito de difração.

Figura 1 - Difração de onda de baixa frequência em um canto.

Figura 2 - Difração de onda de alta frequência em um canto.

Figura 3 - Difração de onda de baixa frequência em uma abertura.

Figura 4 - Difração de onda de alta frequência em uma abertura.

3.2 PERCEPÇÃO ESPACIAL DE ÁUDIO

Neste trabalho são observadas algumas das características referentes à percepção tridimensional do som, ou seja, sua localização e sua direção. Outras características como o timbre por exemplo, não são fundamentais à percepção espacial e, portanto, têm menor importância no presente estudo.

Para entender as características espaciais do áudio, é necessário primeiro compreender alguns de seus aspectos básicos.

As ondas sonoras são capazes de interagir entre si analogamente às ondas eletromagnéticas, isto é, podem se somar umas às outras ou se anular.

As ondas sonoras refletem em materiais e perdem intensidade dependendo do tipo de material. Elas também perdem intensidade de acordo com a distância percorrida (item 3.1.1).

A seguir (subseções 3.2.1 à 3.2.3) encontram-se identificadas algumas das principais características de áudio relacionadas à percepção espacial:

3.2.1 Diferença de Tempo Interauricular - ITD (*Interaural Time Differences*)

Esta característica está associada à diferença de tempo de chegada de áudio, a partir de uma fonte localizada num espaço tridimensional, a cada um dos ouvidos, em decorrência da velocidade da onda sonora em um ambiente. Significa a diferença de tempo entre o momento de chegada de áudio ao ouvido direito e do momento de chegada da mesma onda ao ouvido esquerdo. Por exemplo, se uma fonte de áudio está fisicamente mais próxima de um dos ouvidos, a onda sonora emitida atingirá primeiramente esse mesmo ouvido e, em seguida, o ouvido mais distante. Portanto, existe uma diferença de tempo entre o momento que o áudio chegou ao ouvido mais próximo e o momento que chegou ao ouvido mais distante.

Essa diferença de tempo é percebida pelo cérebro, fornecendo uma pista da localização tridimensional da fonte de áudio.

3.2.2 Diferença de Intensidade Interauricular - IID (*Interaural Intensity Differences*)

Devido à propagação do áudio em um ambiente, ele pode chegar com diferenças de intensidade entre os ouvidos do ser humano. Novamente, devido à diferença das distâncias envolvendo a fonte de áudio e os ouvidos. A intensidade do áudio diminui de acordo com a distância percorrida, devido à perda de energia na propagação. Por exemplo, em uma câmara anecóide¹, o ouvido mais próximo a uma fonte de áudio, recebe a onda com uma intensidade maior que o ouvido mais distante da fonte (item 3.1.1).

3.2.3 Sombra Acústica

Embora alguns autores abordem sombra acústica como uma consequência da diferença de intensidade interauricular, aqui se preferiu separá-lo, a fim de melhor esclarecer cada um dos fenômenos. A sombra acústica está ligada ao conceito de difração (item 3.1.7), onde a intensidade após a realização do contorno em um objeto, é atenuada de acordo com o tamanho do objeto e da frequência da onda. Por exemplo, supondo que uma fonte de áudio esteja localizada à esquerda de um indivíduo, e que sua cabeça seja igual ou maior que dois terços do comprimento de onda do áudio emitido por esta fonte, o áudio que chega ao ouvido direito tem sua intensidade dependente do comprimento de onda emitido, podendo até ser inaudível. Outros possíveis obstáculos, dependendo do posicionamento e orientação da fonte de áudio e do ouvinte, para o fenômeno de sombra acústica, são os ombros e as orelhas.

Como exemplo da utilização desses conceitos, os autores (MUELLER; ULLMANN, 1999) implementam as características de percepção de áudio ITD e IID (descritas nos itens 3.2.1 e 3.2.2) em seus trabalhos experimentais propondo uma técnica para mapeamento e processamento de áudio com a utilização de “raios de

¹ ANECÓIDE: que não produz ecos e possui baixíssima reverberação, isto é, que evita múltiplas reflexões de uma onda sonora.

áudio" (*audio ray tracing*), que permitem resumir de modo aceitável o modelo acústico e, também, torna possível a utilização de processamento paralelo para melhor desempenho.

Outro exemplo é o trabalho de (CHAN; CHEN, 2000), que aplicam esses conceitos na proposta de uma técnica híbrida para a obtenção de funções para processamento de áudio espacial. Esta técnica envolve: a abordagem de aproximação por modelo, para obter os cálculos referentes à cabeça do usuário e, também, a utilização da abordagem de medição, para os cálculos referentes à orelha do usuário.

Por fim, há também que ser citadas outras pistas importantes para a percepção espacial de áudio, embora relacionadas com as outras características mencionadas anteriormente. São elas: o formato concentrador de áudio das orelhas que permite a uma pessoa identificar o posicionamento vertical de uma onda sonora quando esta está localizada à sua frente; o movimento realizado com a cabeça, com a finalidade de confirmar a origem de uma fonte sonora, por meio da análise da diferença de fase de uma mesma onda em cada um dos ouvidos, mesmo desprovido de confirmação visual.

3.3 DISPOSITIVOS TECNOLÓGICOS

Muitos dos dispositivos tecnológicos existentes que podem ser aplicados em sistema de comunicação de áudio espacial não são exclusivos desta área. Um dos poucos fabricantes que implementa diretamente esses conceitos em *hardware* e *software* é a empresa AuSim (AUSIM, 2008).

Como alternativa, podem ser realizadas combinações entre diferentes equipamentos e programas para criar soluções de áudio 3D, como sistemas compostos por múltiplas caixas acústicas disponíveis no mercado (caso dos sistemas Dolby 5.1) ou com um simples fone de ouvido estéreo.

As soluções que envolvem a utilização de fones de ouvido minimizam alguns problemas referentes à execução de áudio, pois evitam que o áudio gerado para tocar em uma determinada posição seja escutado pelo ouvido para o qual não foi destinado (efeito *cross-talk*) (MUELLER; ULLMANN, 1999). Neste caso, a emulação

deve contemplar o efeito *cross-talk*. No entanto, a utilização de dispositivos com múltiplas vias de áudio possui características que podem ser vantajosas em relação aos fones de ouvido, como o conforto, dado que não há a necessidade de vestir o dispositivo.

Também já foram desenvolvidos por pesquisadores fones de ouvido que permitem a passagem parcial ou total do som, para melhor implementar os conceitos de RA. Estes, analogamente aos dispositivos de RA visuais do tipo '*optical see through*' e '*video see through*', são denominados: *audio hear through*, *microphone hear through* (*mic hear through*) ou alguma outra variação destes nomes (LINDEMAN; NOMA; BARROS, 2007).

3.4 ÁUDIO ESPACIAL E MEMORIZAÇÃO

De acordo com os autores (BALDIS, 2001) e (SCHAMNDT; MULLINS, 1995), o áudio espacial afeta positivamente a memorização, permitindo um maior grau de inteligibilidade e compreensão.

No trabalho de (SCHAMNDT; MULLINS, 1995), é possível entender como o áudio espacial permite ao ouvinte selecionar e manter o foco auditivo em uma única fonte de áudio, dentre várias outras fontes presentes em um mesmo ambiente de áudio espacial emulado.

No trabalho de (BALDIS, 2001), é realizada uma análise mais profunda: são avaliados também aspectos de inteligibilidade e compreensão em ambientes de áudio espacial artificial. Neste, são apresentadas aos participantes testados, diversas conferências em tipos de ambientes diferentes: sem emulação de áudio espacial; com áudio espacial parcialmente emulado e, finalmente, com áudio espacial completamente emulado.

Após a realização dos testes, os resultados demonstraram, de acordo com a própria autora, uma melhor compreensão e inteligibilidade quando da emulação completa de áudio espacial, ao se comparar com os resultados dos demais tipos de testes (parcialmente emulado e sem emulação).

As características de uma onda sonora real ou de uma artificial com a capacidade de emular a localização espacial da fonte sonora permitem que o ser

humano seja capaz de inferir a posição física da fonte sonora ao receber seu sinal de áudio. Num ambiente colaborativo, onde estão presentes diversos indivíduos e eventualmente todos emitindo sinais sonoros ao mesmo tempo, essa capacidade de identificar a posição física de uma fonte sonora específica pode ser de extrema valia para o observador que recebe os sinais de áudio focalizar a sua atenção numa dessas fontes sonoras específica. Certamente, essa facilidade de concentrar a atenção numa desta fonte sonora deve contribuir para uma melhor compreensão de conteúdo quando este é transmitido num ambiente colaborativo.

Baseada nesta expectativa é que este trabalho especificou e implementou uma ferramenta de áudio conferência 3D que está descrita no próximo capítulo.

4 FERRAMENTA DE ÁUDIO CONFERÊNCIA ESPACIAL (3D)

4.1 PROPOSTA

O objetivo da ferramenta de áudio conferência espacial é oferecer um ambiente aumentado (real e virtual) que, num contexto educacional, permite a discussão de questões relacionadas a ações de aprendizagem previamente definidas pelo instrutor ou professor. Para que isso seja possível, são aplicados todos os conceitos de RA apresentados anteriormente (item 2.1 deste documento).

O objetivo desta ferramenta é enriquecer o ambiente real de cada conferencista com os avatares dos outros conferencistas, criando um ambiente de RA implementado por meio de técnicas de áudio espacial, que acentuam a sensação de presença entre os participantes. Conseqüentemente, a interatividade é favorecida, o que facilita a realização de atividades colaborativas.

As Figuras 5 e 6 ilustram esta situação. Na figura 5, o boneco azul é o participante local presente no ambiente real, prestes a entrar numa conferência. Já a Figura 6 representa o ambiente de RA após o participante local entrar na conferência, na qual os bonecos vermelhos representam os avatares (de áudio somente) dos participantes em localidades remotas, inseridos artificialmente no ambiente real do participante local.

Figura 5 - Participante local em seu ambiente real, prestes a entrar em uma conferência.

Figura 6 - Ambiente do participante local enriquecido com os avatares (de áudio somente) dos conferencistas remotos.

A ferramenta de áudio conferência 3D deve ser capaz de proporcionar sessões remotas em um ambiente aumentado de ensino e aprendizagem, que permite a interação entre todos os participantes, alunos e professores.

O ambiente aumentado é definido e configurado em um computador (denominado servidor) acessível através da *Internet*. Os participantes utilizam computadores com um programa cliente para acesso ao servidor. Este programa poderá ser obtido e iniciado através de *download* de uma localização específica, como a partir de um *workspace*² do Tidia-Ae, por exemplo.

Quando um participante ativar o programa cliente e se conectar ao servidor, poderá escolher uma das salas existentes para o *workspace* em questão. E uma vez dentro da sala, deverá selecionar uma das cadeiras disponíveis exibidas através de um mapa. É importante ressaltar que estas cadeiras virtuais têm como finalidade orientar e respeitar o posicionamento escolhido por cada participante. Feito isso, o usuário está inserido no ambiente aumentado.

A partir deste momento, é iniciada a comunicação de áudio para o usuário corrente. Visto que o cenário de utilização trata de participantes em ambientes remotos, o áudio gerado por cada um deles é capturado e transmitido, por meio de redes de computadores, aos programas clientes dos demais participantes.

Já o áudio recebido, o qual foi enviado pelos outros participantes, é processado para simular a localização virtual de cada um dos participantes da sala. Isso é fundamental para que seja preservado um dos principais objetivos desta ferramenta (e deste trabalho), que é o aumento na sensação de presença.

Neste trabalho são aplicados os estudos de percepção espacial de áudio, técnicas, tecnologias e algoritmos que viabilizam a elaboração e construção de uma ferramenta que implementa as funcionalidades descritas nesta proposta.

4.2 PRINCIPAIS FUNCIONALIDADES

A fim de implementar a proposta apresentada, foi necessário organizar e programar algumas funcionalidades consideradas fundamentais, sem as quais não

² WORKSITE é um ambiente no sistema Tidia-Ae que permite a realização de atividades educacionais eletrônicas. Este pode ser contextualizado em algum tema ou até mesmo representar uma disciplina.

seria possível atingir os objetivos já relatados no início deste trabalho. A seguir estão listadas as principais funcionalidades que a ferramenta deve oferecer:

- **Sentido de tele-presença:** Característica fundamental desta ferramenta e experiência. É importante que o participante remoto tenha a sensação e percepção de compartilhar o mesmo ambiente que outros participantes remotos, embora seu sentido de presença mantenha-se no ambiente real no qual ele se encontra;
- **Captura de áudio:** Dispositivos tecnológicos devem capturar o áudio mono gerado por cada um dos participantes do ambiente artificial, denominado de sala;
- **Codificação do áudio:** Todo o áudio capturado deve ser codificado e, em seguida, armazenado de modo a possibilitar sua transmissão;
- **Transmissão do áudio:** O fluxo de áudio deve ser formatado e empacotado para possibilitar a transmissão e envio aos computadores dos demais participantes;
- **Recepção e decodificação do áudio:** O fluxo de áudio deve ser reconstruído a partir dos pacotes de dados recebidos pela rede..
- **Aplicação dos efeitos de áudio:** Após a reconstituição do fluxo de áudio mono de cada participante, ele é processado com os efeitos (filtros) respeitando a posição virtual daquele usuário. Após este processamento, é obtido um fluxo de áudio estéreo para cada um dos participantes.
- **Mixagem e reprodução de áudio:** Os fluxos de áudio estéreo obtidos após o processamento tridimensional com os efeitos são então mixados em um único fluxo de áudio estéreo final. Imediatamente após a geração deste, inicia-se a reprodução.
- **Integração com o sistema Tidia-Ae:** A ferramenta de comunicação de áudio deve ser disponibilizada no sistema Tidia-Ae, pois é essa integração que caracteriza a sua utilização educacional. Desta integração surgem vantagens com relação ao sistema de aprendizagem como a possibilidade de utilização desta ferramenta com as outras ferramentas já existentes para a realização de atividades colaborativas. Outro benefício seria o aproveitamento das funcionalidades já implementadas no sistema Tidia-Ae como o acesso ao cadastro de usuários. Porém, para que isso

ocorra, parte dela deve ser compatível com a arquitetura já definida para todo o sistema.

- **Compartilhamento do ambiente virtual:** É importante que o ambiente virtual seja compartilhado por todos os participantes inseridos na mesma sala, já que deve haver a consistência de posicionamento. Isso significa que a referência da localização virtual é válida para todos, de forma que se o participante A estiver do lado esquerdo do participante B, assim o será para os demais usuários. Ou seja, para todos os outros participantes, B estará ao lado direito de A.
- **Gerência de salas de reunião:** Para cada *worksites* do sistema Tidia-Ae em que a ferramenta de comunicação de áudio esteja associada, devem existir salas de reunião virtuais para que os participantes possam interagir. Portanto, a ferramenta deve ser capaz de gerenciá-las, oferecendo algumas operações básicas como a criação e remoção de salas e entrada ou saída de participantes.

4.3 PROJETO

A solução adotada, embora não única, para as funcionalidades introduzidas na seção 4.2, foi a implementação de um sistema cliente-servidor. Isto foi necessário, pois o sistema Tidia-Ae é acessado por meio de um navegador *web*, que não tem a capacidade de executar as funcionalidades designadas para máquinas de acesso (como o processamento de áudio espacial, por exemplo). Isso decorreu da decisão que o processamento de áudio espacial seria distribuído, isto é, não seria gerado em uma única máquina central, mas sim nos terminais de acesso ao sistema Tidia-Ae. Logo, esses terminais precisam executar processos novos, o que acabaram por constituir programas clientes de uma estrutura cliente-servidor.

A parte cliente da ferramenta engloba a maior parte das funcionalidades, pois é responsável pela captura, execução e processamento de efeitos tridimensionais dos fluxos de áudio, o que também facilita a escalabilidade do projeto proposto. Levando-se em consideração que o programa cliente é distribuído via *internet* quando da ativação da ferramenta, é importante que este seja desenvolvido numa

tecnologia compatível com diversos sistemas operacionais, a fim de abranger um público-alvo maior. Portanto, de acordo com essas observações, a linguagem Java é a escolhida para sua implementação.

Por fim, o programa servidor deve estar integrado ao sistema Tidia-Ae como uma de suas diversas ferramentas, implementando as funcionalidades de gerenciar, centralizar e persistir a configuração das salas de reunião de um *worksite*.

4.3.1 Considerações sobre bibliotecas para processamento de áudio em tempo real

Para que um sistema de comunicação de áudio ou vídeo conferência ofereça qualidade e seja atraente para os usuários, é fundamental que exista um baixo tempo de resposta. Isto significa que o vídeo ou áudio capturado de um participante deve ser transmitido e exibido aos demais com a menor latência possível, sempre respeitando um grau mínimo de qualidade. Caso contrário, a interação entre os participantes é prejudicada.

Existem limites de tempo de resposta, com relação ao tempo total de interação entre os participantes de uma aplicação de comunicação de áudio interativa, de modo a classificar o que pode ser considerado aceitável ou não. Esses limites e interpretações de qualidades, de acordo com a recomendação (RECOMMENDATION, 1996), são:

Tabela 1 - Tempos de resposta e suas interpretações de qualidade (ITU-T G.114).

Atraso (ida e volta)	Qualidade
Até 150 ms	Aceitável para a maioria das aplicações
150 ms a 400 ms	Pode impactar algumas aplicações
Acima de 400 ms	Inaceitável

Já em outros tipos de interação, como as interações musicais relatadas em (LAGO; KON, 2004), o tempo de 30 ms, embora aceitável segundo a recomendação apresentada na Tabela 1, é considerado um longo tempo de resposta no contexto musical onde o retorno sonoro não pode demorar mais que 4 ms, como é o caso de sons gerados por instrumentos musicais eletrônicos.

Todavia, ao se considerar as observações sobre o baixo tempo de resposta

dos diversos tipos de interações, percebe-se o quanto é importante que o tempo de interação seja o menor possível, a fim de aumentar a abrangência do escopo de utilização da aplicação de áudio conferência espacial.

Neste contexto, com relação ao processamento de áudio espacial que é realizado em cada um dos computadores dos conferencistas remotos, é necessário uma biblioteca e sua *Application Programming Interface* (API) de captura e reprodução de áudio com baixo tempo de resposta, a fim de minimizar o tempo em que os dados de áudio permanecem na memória do computador. Também é importante que esta biblioteca possa ser utilizada em múltiplas plataformas, permitindo seu funcionamento em diversos sistemas operacionais diferentes.

Abaixo estão listadas algumas das bibliotecas estudadas para o processamento de dados de áudio, bem como algumas observações sobre as mesmas:

- **ASIO:** *Audio Stream Input/Output* é uma tecnologia definida pelo fabricante *Steinberg* (ASIO, 2008) para processamento digital de áudio profissional. Sua principal característica é servir de base para o processamento em tempo real (baixa latência) de áudio de alta qualidade. Para isso, é definida uma arquitetura na qual o *hardware* e *software* fazem parte implementando um sistema para a rápida troca de *buffers* de áudio de captura e execução.
- **WDM:** *Windows Driver Model* (WDM, 2008) é uma das tecnologias definida pela *Microsoft* para *drivers* de dispositivos multimídia. Ela permite o controle de qualidade de amostras de áudio, mas não serve de base para processamento de áudio em tempo real já que sua arquitetura não serve à utilização com baixa latência.
- **Java Sound API:** Esta tecnologia Java é uma API para processamento de áudio em máquinas virtuais Java (JAVA, 2008) que funciona sobre uma arquitetura de áudio virtual convertendo-a para uma arquitetura de áudio real. Há uma pequena perda de desempenho em relação à latência mas sua grande vantagem é o fato de atender a diversas plataformas, que é uma característica importante da linguagem *Java*, permitindo seu funcionamento em sistemas operacionais baseados em *Unix* (MacOSX ou *Linux*) e *Windows*.

Apesar das considerações existentes sobre o insatisfatório desempenho de aplicações escritas em Java para processamento de áudio em tempo real, como relatada por (JUILLERAT; MÜLLER; SCHUBIGER-BANZ, 2007), esses mesmos autores indicam como trilhar esse caminho e superar essas dificuldades, em um trabalho específico de processamento de áudio. No trabalho referenciado, procura-se diminuir problemas de latência e tempo de resposta, ou seja, de desempenho. Das contribuições oferecidas pelos autores, foram aproveitadas todas referentes a manipulação do *Garbage Collector*, de modo a evitar sua freqüente ativação, o que requisita um *lock* da memória do programa cliente, paralisando-o temporariamente.

Visto que problemas de desempenho relacionados a processamento de áudio com a biblioteca *Java Sound* podem ser resolvidos e, aliado à abrangência de plataformas oferecidas pela utilização da linguagem Java, a biblioteca *Java Sound* tornou-se a opção para a realização de processamento de áudio espacial neste trabalho.

4.3.2 Considerações sobre microfones e fones de ouvido

É importante ressaltar que os usuários da implementação atual da ferramenta devem utilizar fones de ouvido estéreo ou similares (*audio hear through*), pois o processamento de áudio realizado pela aplicação contempla somente esta configuração. Caixas acústicas não devem ser empregadas pois esta a implementação atual não possui uma solução para o efeito *cross-talk* (ver item 3.4).

Já os microfones utilizados pelos participantes devem efetuar a captação do áudio o mais próximo possível da boca, de modo a minimizar os efeitos acústicos do ambiente real no qual se encontra. Se um microfone do tipo viva-voz for utilizado, por exemplo, ele irá captar além da incidência da voz direta do participante, os ecos ocasionados pelas reflexões da voz no ambiente real bem como ruídos presentes no recinto (como os originados por máquinas) e originados por outras pessoas. Isto pode prejudicar o processamento espacial de áudio.

Portanto, é recomendada a utilização de *headsets*. Eles permitem ao participante escutar o som do ambiente no qual ele se encontra enriquecido (mixado) com o áudio gerado pelo processamento espacial, a fim de implementar os conceitos de RA. Simultaneamente, eles permitem a captura de áudio próximo à

boca do usuário, minimizando a captura de outros sons presentes do ambiente.

4.3.3 Processamento de áudio tridimensional

Para que o áudio tridimensional seja corretamente emulado, foi necessário o estudo de efeitos e algoritmos para processamento de áudio que implementassem, de modo satisfatório, as características psicoacústicas referentes ao posicionamento e direção do áudio, analogamente ao realizado no trabalho de (KOBAYASHI; ONOYE; SHIRIKAWA, 2001) os quais propõem um algoritmo para se criar o efeito de localização artificial de fontes de áudio num ambiente tridimensional.

Além destes estudos, foram necessárias algumas simplificações e generalizações de modo a melhor delimitar o escopo das condições para o processamento de áudio espacial. Dentre estas, a sala aumentada (ambiente resultante final) trata-se de uma sala anecóide, ou seja, que absorve as ondas sonoras que colidem com as paredes e os objetos. Isto evita a necessidade da emulação de ecos e reflexões envolvendo a mobília e as paredes. Outra simplificação diz respeito ao formato e tamanho da cabeça dos participantes. Considerou-se de uma esfera com dezesseis centímetros de diâmetro, padronizando o processamento de áudio para cada participante. Isto é conveniente, pois caso contrário cada participante necessitaria configurar na ferramenta de áudio conferência sua distância interauricular e o formato de sua cabeça. É necessário lembrar que, para uma melhor experiência de uso, cada participante deve utilizar em seu computador um *headset* estéreo conectado ao seu computador.

No uso da ferramenta proposta neste trabalho, o fluxo de dados e o processo operacional ocorrem de acordo com os passos descritos a seguir, os quais podem ser acompanhados por meio da Figura 7:

- A - O áudio é capturado pelo microfone dos participantes pelo programa cliente;
- B – Os dados de áudio (mídia contínua) que representam o áudio analógico capturado são transmitidos via *Internet* aos programas clientes de cada um dos outros participantes;
- C - Cada instância do programa cliente recebe os dados de áudio enviados

pelos outros participantes de uma mesma sala de áudio conferência;

- D – Os dados de áudio recebidos são processados com filtros e efeitos de áudio, considerando a localização e orientação virtual do participante emissor do áudio e, também, a localização virtual e orientação dos ouvidos do participante local do programa cliente (receptor);
- E - A partir deste processamento, são obtidos dois *buffers* de áudio estéreo, representando o áudio destinado aos ouvidos esquerdo e direito para cada um dos participantes remotos emissores;
- F - Todos os *buffers* contendo áudio para o ouvido esquerdo são mixados em um único *buffer* final para o ouvido esquerdo. Analogamente, é obtido um único *buffer* final para o ouvido direito;
- G - Este par de *buffers* (áudio estéreo) é então reproduzido para o participante receptor.

Figura 7 – Captação, transmissão e processamento de Buffers de Áudio nos programas clientes (cor verde clara).

Para realizar o processamento tridimensional de áudio (passo D), são utilizados três filtros de áudio para emular as características espaciais de áudio ITD, IID e sombra acústica, os quais já foram introduzidos nos itens 3.2.1, 3.2.2 e 3.3.3 respectivamente.

Os filtros em questão são de: intensidade, passa-baixa e atraso. A descrição de cada um deles é apresentada a seguir, embora suas configurações sejam somente abordadas sob o item 4.3.4, cujo objetivo é apresentar o modelo tridimensional de cálculos utilizado.

4.3.3.1 Filtro de passa baixa (*Low pass*)

Este filtro tem como objetivo eliminar frequências altas além da especificada (frequência de corte) nas amostras de som. Ele é utilizado para processar o áudio proveniente de uma fonte que está oclusa pela cabeça ou orelhas do ouvinte, eliminando parte do som agudo para o ouvido ao qual a onda sonora não atinge diretamente, aumentando a percepção espacial e emulando o fenômeno de sombra acústica (descrita no item 3.2.3).

A graduação para este efeito é determinada baseada em duas condições: se a onda sonora tangencia a cabeça, quanto mais ela tangencia, maior é a sombra acústica e, portanto, menor é a frequência de corte; se a onda sonora tem como origem o hemisfério traseiro, significa que as orelhas ocasionarão sombra acústica e, por consequência, uma frequência pré-determinada de corte é configurada.

4.3.3.2 Filtro de atraso (*Delay*)

Este filtro tem como objetivo inserir atraso no áudio. Também como o filtro de passa baixa, ele é utilizado para processar o áudio de uma fonte que está localizada ao lado do ouvinte. Seu funcionamento insere um atraso para o ouvido oposto à fonte de áudio, pois a diferença de chegada de áudio nos ouvidos aumenta a percepção espacial da fonte. A característica de percepção espacial relacionada a

este filtro é a ITD (descrita no item 3.2.1).

O cálculo para sua configuração é feito baseado na diferença de tempo que uma onda sonora leva para chegar aos ouvidos direito e esquerdo. Os atrasos para os ouvidos são computados utilizando-se como base velocidade do som de 340 m/s.

4.3.3.3 Filtro de intensidade (Volume)

Este filtro tem como objetivo diminuir a intensidade do áudio e, como os filtros de passa baixa e atraso, é utilizado para processar o áudio de uma fonte que está localizada ao lado do ouvinte. Seu funcionamento deve diminuir a intensidade do áudio para o ouvido oposto à fonte de áudio, pois essa diferença de intensidade também realça a percepção espacial. A característica de percepção espacial relacionada a este filtro é a IID (descrita no item 3.2.2).

Os três filtros necessitam de parâmetros para seu correto funcionamento. A obtenção desses parâmetros ocorre por meio da análise de um modelo tridimensional representando os participantes emissores e o participante ouvinte, a qual é apresentada no próximo item (4.3.4).

4.3.4 Modelo para cálculo e transferência de som espacial (3D)

Viu-se necessária a criação de um modelo tridimensional para a representação do posicionamento e orientação espacial dos participantes virtuais, de modo a permitir o cálculo dos parâmetros e a configuração dos filtros.

Primeiramente, é importante ressaltar que, como ilustrado na Figura 7, em cada uma das instâncias do programa cliente é realizado o processamento de áudio espacial. Neste processamento e, conseqüentemente, neste modelo, o participante local do programa é sempre o participante receptor e os participantes remotos a ele são considerados os participantes emissores.

Deste modo, na memória de cada programa cliente existe uma representação computacional tridimensional do posicionamento e orientação das fontes emissoras

de áudio e o receptor. Esta representação é atualizada pelo programa servidor a cada cinco segundos, a fim de refletir as possíveis alterações ocorridas no ambiente, como: a mudança da localização espacial de um participante, a entrada ou saída de um novo participante na conferência. A partir desta representação, torna-se possível auferir as distâncias, ângulos e demais dados necessários para os cálculos dos parâmetros s serem configurados nos filtros. É importante ressaltar que, a cada atualização da representação, os cálculos são todos refeitos e os filtros configurados novamente.

A avaliação para a configuração dos filtros é realizada em etapas, uma para cada participante emissor no ambiente aumentado de áudio conferência. Visto que para cada participante emissor existem três filtros a serem configurados para o ouvido direito e outros três para o ouvido esquerdo, em cada etapa é preciso configurar seis filtros.

A etapa é composta por uma sequência de cálculos que considera os seguintes elementos: um ponto tridimensional no espaço representado a fonte emissora (participante emissor) utilizada na etapa, o modelo tridimensional (esfera) da cabeça do participante receptor, três filtros para processar o áudio destinado ao ouvido esquerdo do receptor (canal esquerdo) e outros três para o ouvido direito (canal direito).

Ao final de cada etapa, vêem-se configurados os seis filtros para processar o áudio gerado pelo participante emissor da etapa em questão.

Para a realização de uma sequência de cálculos, são considerados importantes os seguintes pontos no espaço 3D (ilustrados pela Figura 8):

- a boca do participante emissor da etapa (denominado ponto B);
- o centro da cabeça (esfera) do participante receptor (denominado ponto C);
- o ouvido direito do participante receptor (denominado ponto D);
- o ouvido esquerdo do participante receptor (denominado ponto E);
- o nariz do participante receptor (denominado ponto N).

Figura 8 – Pontos 3D representando os ouvidos (E e D), nariz (N) e centro da cabeça (C) do participante receptor e boca do participante emissor (B).

Primeiramente, é necessário elucidar as equações aplicadas nos cálculos, que são basicamente duas: a utilizada para a obtenção de uma distância entre dois pontos no espaço tridimensional eq.(1) e para se calcular o menor ângulo formado por três pontos no espaço tridimensional eq.(2).

A seguir está a equação para o cálculo da menor distância entre dois pontos A e B no espaço tridimensional:

$$distAB = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2} \quad (1)$$

Já o seguinte procedimento é utilizado para determinar o menor ângulo formado pelos pontos A, B e C, onde C é o centro do ângulo:

- a) Primeiro deve-se determinar os vetores $\overrightarrow{CA}$ e $\overrightarrow{CB}$ de três dimensões, no formato (x,y,z):

$$\overrightarrow{CA} = ((x_A - x_C), (y_A - y_C), (z_A - z_C))$$

$$\overrightarrow{CB} = ((x_B - x_C), (y_B - y_C), (z_B - z_C))$$

- b) Em seguida, calcula-se o cosseno do ângulo por meio da equação para determinar o cosseno do menor ângulo entre dois vetores não nulos:

$$\cos = \frac{\text{produtoEscalar}}{\text{produtoVetorial}} = \frac{\overrightarrow{CA} \cdot \overrightarrow{CB}}{\|\overrightarrow{CA}\| \|\overrightarrow{CB}\|} \quad (2)$$

Esclarecidas as equações eq.(1) e eq.(2), é facilitada a compreensão da sequência de cálculos e verificações para a graduação e configuração dos filtros utilizados.

Cada etapa (uma etapa para cada participante emissor) é dividida em três partes, onde em cada parte se obtém a configuração para dois filtros do mesmo tipo (um para o canal direito e outro para o esquerdo). A ordem na qual ocorrem essas configurações é a seguinte: do atraso, de passa-baixa e de intensidade.

Cada configuração está apresentada em um dos três sub itens a seguir (4.3.4.1, 4.3.4.2 e 4.3.4.3).

4.3.4.1 Configuração do Atraso

A fim de que o atraso seja configurado, é necessário obter as distâncias entre a fonte de áudio (ponto B) e os ouvidos do participante receptor, representado pelos pontos E e D. Visto que a cabeça do participante receptor (esfera de centro C) é um obstáculo para a chegada do som aos ouvidos (pontos E e D), a distância mais curta da fonte aos ouvidos talvez seja tangenciando a cabeça do participante receptor, ou seja, pode não se tratar de uma linha reta. Portanto, torna-se importante analisar se este é ou não o caso.

Esta análise é realizada utilizando uma representação bidimensional dos pontos tridimensionais envolvidos. O plano delimitado pelos pontos B, C e D é utilizado para avaliar esta situação referente ao ouvido direito. Já com relação ao ouvido esquerdo, o mesmo plano pode ser utilizado, visto que o ponto E encontra-se na mesma reta que os pontos C e D.

O modelo ilustrado pela Figura 9 representa a situação e os ângulos envolvidos para a verificação se uma onda sonora emitida pelo ponto B chega ao ouvido direito (D) tangenciando cabeça ou diretamente:

Figura 9 – Ângulos para análise de incidência direta de onda sonora no ouvido direito do participante

receptor.

Primeiramente, deve-se calcular os ângulo β e θ . O valor do cosseno de θ é obtido por meio da equação eq.(2) como indicado abaixo:

$$\cos \theta = \frac{\overrightarrow{CD} \cdot \overrightarrow{CB}}{\|\overrightarrow{CD}\| \cdot \|\overrightarrow{CB}\|}$$

A partir do valor do cosseno, obtém-se o valor do ângulo em radianos.

Para o cálculo do ângulo β , é necessário antes calcular a distância entre os pontos C e B e, também, entre os pontos C e T. A distância entre os pontos C e T é conhecida, pois trata-se do raio da esfera (r). Já a distância entre os pontos C e B é obtida por meio da equação eq.(1):

$$distCB = \sqrt{(x_B - x_C)^2 + (y_B - y_C)^2 + (z_B - z_C)^2}$$

Ao se observar a Figura 2, verificamos que o triângulo delimitado pelos pontos C, B e T é um triângulo retângulo, tendo como hipotenusa o segmento de reta CB. Deste modo, o cosseno do ângulo β é obtido por meio da seguinte equação:

$$\cos \beta = \frac{distCT}{distCB} = \frac{r}{distCB}$$

A partir do valor do cosseno de β , obtém-se o valor do ângulo.

Em seguida, é necessário comparar os valores de β e θ . O valor de θ pode ser maior, igual ou menor que o valor de β , resultando em três situações. A Figura 10, ilustra a situação onde o ângulo β é menor que o ângulo θ .

Figura 10 – Situação onde o ângulo β é menor que o ângulo θ .

Isto significa que a onda sonora tangencia a cabeça, fazendo com que a distância total entre a boca do participante emissor (ponto B) e o ouvido direito do participante receptor (ponto D) seja a soma da distância entre os pontos B e T e o trecho da circunferência entre os pontos T e D, como a equação a seguir mostra:

$$distBD = distBT + \alpha.r$$

Onde:

$distBT$ = distância entre os pontos B e T obtida pelo teorema de Pitágoras, considerando o triângulo retângulo delimitado pelos pontos C, B e T.

$$\alpha = \theta - \beta$$

As Figuras 11 e 12 ilustram, respectivamente, as situações onde β é igual e maior que o ângulo θ .

Figura 11 – Ângulo β é igual ao ângulo θ .

Figura 12 – Ângulo β é maior que ângulo θ .

Nestas duas situações, a distância entre os pontos B e D é calculada diretamente, por meio da equação eq.(1), como representado abaixo:

$$distBD = \sqrt{(x_D - x_B)^2 + (y_D - y_B)^2 + (z_D - z_B)^2}$$

Todos os mesmos cálculos devem ser efetuados novamente, mas a fim de encontrar a distância entre os pontos B e E. Somente não é necessário recalculer o ângulo β , pois este tem o mesmo valor obtido durante a análise para da distância para o ouvido direito.

Obtidas as distâncias $distBD$ e $distBE$, o tempo de chegada de uma onda sonora a CAD um dos ouvidos é calculado de acordo com as seguintes equações:

$$tempoBD = \frac{distBD}{v} \text{ e}$$

$$tempoBE = \frac{distBE}{v}$$

Onde:

$distBD$ = distância final entre os pontos B e D (em metros);

$distBE$ = distância final entre os pontos B e E (em metros);

$tempoBD$ = tempo que o som leva para sair de B e chegar até D (em segundos);

$tempoBE$ = tempo que o som leva para sair de B e chegar até E (em segundos);

v = velocidade do som no ar. O valor utilizado é 340 m/s (metros por segundo).

Munido dos tempos de chegada do som, a diferença entre eles é calculada, a fim de obter o parâmetro de configuração do filtro, de acordo com a equação abaixo:

$$\Delta tempo = |tempoBD - tempoBE|$$

Finalmente, o filtro de atraso é configurado com o valor de atraso $\Delta tempo$, para o ouvido mais distante da fonte de áudio B. Já o filtro de atraso responsável pelo ouvido mais próximo à fonte de áudio, o valor configurado para atraso é zero, isto é, sem atraso.

4.3.4.2 Configuração de Passa-Baixa

A fim de configurar o filtro de passa-baixa, os parâmetros a serem utilizados dependem de dois fatores: se a onda sonora chega ao ouvido tangenciando a cabeça do ouvinte e se a onda sonora está oculta pela orelha.

Para verificar se a onda tangencia a cabeça em direção a um dos ouvidos, basta se comparar os ângulos θ e β já calculados durante a configuração dos filtros de atraso, isto é, se β menor que θ , a onda tangencia. Neste caso, o ângulo α já

estará calculado e a graduação do efeito depende de seu valor, visto que esse ângulo indica o quanto o áudio tangencia a cabeça. A seguir está a Tabela 2 relacionando a graduação do ângulo com a utilizada pelo para a configuração:

Tabela 2 – Graduação do ângulo α com amostras utilizadas e frequência de corte.

α (em graus)	Amostras	Frequência de corte (Hz)
$> 75^\circ$	9	1225
$75^\circ \sim 60^\circ$	7	1575
$60^\circ \sim 45^\circ$	5	2205
$45^\circ \sim 20^\circ$	3	3675
$< 20^\circ$	2	5513

A quantidade de amostras indica quantas amostras o filtro de passa-baixa deve utilizar para realizar o efeito de passa-baixa. Quanto maior o número de amostras utilizadas, mais baixa é a frequência de corte. Isto decorre do fato de que, para se calcular o valor de cada amostra de saída, se efetua a média aritmética das N últimas amostras de entrada (quantidade indicada na Tabela 2). Deste modo, a variação de pressão que ocorre no intervalo representado pelo conjunto de amostras especificadas é minimizada e, por consequência, há um corte da faixa de frequência analógica. Por exemplo, na implementação utilizada neste trabalho, a taxa de amostragem é de 22050 amostras por segundo. Ao se utilizar nove amostras para o cálculo, divide-se o valor de amostragem por 9, resultando em 2450 amostras por segundo. Visto que pelo teorema de Nyquist, a representação de uma onda analógica de determinada frequência necessita de uma frequência de amostragem de, pelo menos, duas vezes a frequência analógica, 2450 amostras por segundo conseguem representar, no máximo, uma onda sonora analógica de 1225 Hz.

Já no caso do ângulo β ser maior ou igual a θ , significa que a onda sonora emitida pelo ponto B não tangencia a cabeça do participante receptor (esfera de centro C). Nesta situação é necessário verificar se há uma possível oclusão da onda sonora pela orelha do participante receptor. Para realizar essa verificação, é preciso antes identificar se a fonte está no hemisfério frontal ou traseiro do receptor, o que pode ser obtido ao se calcular um novo ângulo, entre os pontos N, C e B. Para isto é utilizada a equação eq.(2), da qual se obtém o cosseno do ângulo para em seguida se obter o valor do ângulo (em radianos):

$$\cos NCB = \frac{\overrightarrow{CN} \cdot \overrightarrow{CB}}{\|\overrightarrow{CN}\| \cdot \|\overrightarrow{CB}\|}$$

As Figuras 13 e 14 ilustram duas situações do ângulo NCB. Se este for menor que 90°, a fonte sonora (ponto B) está localizada no hemisfério frontal (Figura 13). Caso contrário, está localizada no hemisfério traseiro (Figura 14).

Figura 13 – Ângulo NCB menor que 90°

Figura 14 – Ângulo NCB maior que 90°

Se a fonte estiver no hemisfério frontal ou se o ângulo NCB for 90°, não será necessário configurar uma frequência de corte para processar a onda sonora emitida pelo ponto B. Mas se a fonte de áudio estiver no hemisfério traseiro, outro ângulo é avaliado, para verificar se a orelha causa alguma oclusão no áudio. Para tal, o ângulo CDB, no caso do ouvido direito é avaliado. A Figura 15 demonstra este ângulo:

Figura 15 – Ângulo CDB, para ouvido direito, caso a fonte de áudio esteja no hemisfério traseiro.

Novamente, a equação eq.(2) é utilizada:

$$\cos CDB = \frac{\overrightarrow{DC} \cdot \overrightarrow{DB}}{\|\overrightarrow{DC}\| \|\overrightarrow{DB}\|}$$

Se o ângulo CDB estiver entre 90° e 110°, significa que a orelha oculta o áudio emitido pela boca do participante (ponto B). Ocorrendo esta oclusão, o número

de amostras a serem configuradas no filtro de passa-baixa para o ouvido direito é 3, ou seja, a frequência de corte é de 3675 Hz. Já para o caso do ouvido esquerdo, os ângulos analisados são: θ , α e CEB específicos para o ouvido esquerdo e os ângulos NCB e β comuns para os dois ouvidos. Ao final desta configuração, os filtros de passa baixa para processar o áudio dos dois canais (esquerdo e direito) referentes ao participante emissor da etapa estarão configurados.

É importante ressaltar que estes valores de configuração presentes na Tabela 1 e para oclusão pela orelha (valor 3), foram calibrados de acordo com a percepção de alguns participantes especialistas que experimentaram a ferramenta, ainda em sua fase de construção. Esses valores foram indiretamente confirmados pela realização dos experimentos realizados, que verificaram o funcionamento do processamento de áudio espacial.

4.3.4.3 Configuração de Intensidade

O cálculo é baseado nos valores já encontrados durante a configuração do filtro de atraso: as distâncias finais entre a boca do participante emissor (ponto B) e os ouvidos direito (distBD) e esquerdo (distBE).

Utilizando-se a seguinte equação para representar a pressão no caso de uma propagação esférica:

$$pressão = \frac{força}{areaEsfera} = \frac{força}{4\pi R^2}$$

Onde:

pressão = pressão em Pascals ou Newtons por metro quadrado (N/m^2);

força = força de valor qualquer em Newtons;

areaEsfera = área da superfície de uma esfera de raio R.

Para que a intensidade seja configurada, é necessário calcular a relação de pressão entre os ouvidos esquerdo e direito. Caso a distância do ouvido esquerdo (ponto E) à fonte sonora (ponto B) seja maior que a distância do ponto B ao ouvido direito (ponto D), ou seja, a fonte sonora está mais próxima do ouvido direito, é utilizada a seguinte relação de pressão para determinar o fator de diminuição de intensidade para o áudio a ser tocado para o ouvido esquerdo:

$$\frac{P_{Esquerdo}}{P_{Direito}} = \frac{\frac{força}{4\pi(distBE)^2}}{\frac{força}{4\pi(distBD)^2}} = \left(\frac{distBD}{distBE} \right)^2$$

Onde:

$P_{Esquerdo}$ = Pressão no ouvido esquerdo;

$P_{Direito}$ = Pressão no ouvido direito;

força = força em Newtons de valor qualquer.

Este valor significa o fator de amplificação das amostras de áudio. Se o fator é igual a um (1), o áudio original não será alterado. Se for maior que um (1), será amplificado (aumentado). Caso seja menor que um (1), será diminuído (intensidade reduzida).

Obtido o fator da relação, este é configurado para o filtro de intensidade referente ao canal esquerdo, de modo a diminuí-la. Já para o canal direito, o fator utilizado é um (1), isto é, não deve haver diminuição ou aumento de intensidade.

Caso a fonte esteja mais próxima do ouvido esquerdo a relação é invertida, ou seja, é calculada a relação de pressão do ouvido direito sobre o esquerdo ($P_{Direito} / P_{Esquerdo}$) e o fator é configurado para o áudio tocado no ouvido direito. Nesta situação, o fator um (1) é configurado para o canal esquerdo.

Como explanado, é possível que um *buffer* de áudio não seja necessariamente processado por todos os três filtros mencionados, por exemplo: se a onda sonora de uma fonte de áudio (ponto B) incide diretamente sobre o ouvido direito, exatamente como ilustra a Figura 11, o *buffer* de áudio referente ao ouvido direito não será processado pelo filtro de passa-baixa por não sofrer o efeito de sombra acústica, também não será processado pelos efeitos de atraso e de intensidade, pois visto que fonte está mais afastada do ouvido esquerdo, o atraso e a relação de intensidade serão somente configurados para este ouvido. Já se a situação fosse como ilustra a Figura 10, onde a boca (ponto B) está ao lado direito, mas no hemisfério traseiro, o filtro da passa baixa seria também aplicado ao *buffer* referente ao ouvido direito, devido à oclusão do som pela orelha.

Caso o uso desses efeitos e filtros não causem a percepção espacial aos

usuários, a sensação presencial será deteriorada e, conseqüentemente, a interatividade será negativamente afetada. Portanto, é imprescindível que seja realizada a correta combinação, calibração e ajustes dos filtros para que o áudio posicional seja corretamente emulado.

4.4 ARQUITETURA DA FERRAMENTA

A concepção desta ferramenta exigiu uma arquitetura que contemplasse o funcionamento integrado entre o sistema Tidia-Ae, o programa servidor e o programa cliente.

O programa servidor utiliza componentes definidos nos moldes da arquitetura do sistema Tidia-Ae. A Figura 16 ilustra o diagrama que representa os principais componentes do programa servidor e de dois programas clientes, em execução, em dois computadores diferentes.

Figura 16 - Diagrama da Ferramenta de comunicação de áudio, com representação do programa servidor e duas instâncias do programa cliente.

O programa servidor possui componentes e código distribuído nas três principais camadas do Tidia-Ae. Já para o programa cliente, o diagrama mostra somente a multiplicidade de instâncias de clientes em diferentes computadores para evidenciar a comunicação de dados UDP *Peer-To-Peer* que ocorre entre os mesmos quando da transmissão e recepção de dados de áudio. Esta comunicação de dados é a mesma já ilustrada anteriormente na Figura 7, onde os *buffers* de áudio são transmitidos diretamente da máquina do participante corrente às outras máquinas participantes. Ou seja, somente os dados de áudio (*buffers*) trafegam no modo *Peer-To-Peer*. É importante ressaltar que o empacotamento dos dados de áudio foi baseado no já existente protocolo RTP, embora não se trate de uma implementação compatível com o mesmo.

Todos os demais dados trafegam por meio de uma conexão TCP (XMLRPC sobre HTTP) estabelecida entre cada instância do programa cliente e o programa servidor.

De um modo geral, é importante ressaltar que o acesso ao modelo tridimensional das salas aumentadas de interação existentes na ferramenta é controlado pelo programa servidor que, de acordo com a Figura 13, é identificado pelas seguintes caixas:

- **Apresentação AC3D**, representando a camada de apresentação *web* da ferramenta que contém somente uma página por meio da qual pode-se iniciar o programa cliente;
- **Serviço XML RPC AC3D**, é o módulo responsável por prover as funcionalidades remotas a serem utilizadas pelos programas clientes como: listagem de salas aumentadas, ocupar ou desocupar um assento, obter os endereços de rede dos outros participantes;
- **Componente Autenticação AC3D**, autentica a conexão que cada programa cliente faz junto ao Serviço XML RPC AC3D, visto que o programa cliente não pode utilizar a conexão TCP já aberta pelo navegador *web* pelo fato de ser um novo processo no computador de acesso;
- **Componente Lógica AC3D**, é responsável pela lógica do programa, implementando as funcionalidades pertinentes a este assunto como: a configuração das salas aumentadas, quais assentos estão livres e quais não estão, consistência dos modelos tridimensionais que representam as salas.

- **Persistência AC3D**, responsável por persistir e recuperar configurações e, também, fornecer o controle transacional das operações lógicas quando múltiplos acessos são realizados simultaneamente, garantindo a consistência dos dados.

4.5 INTEGRAÇÃO COM O SISTEMA TIDIA-AE

A integração da ferramenta com o sistema Tidia-Ae se dá por meio do programa servidor, desenvolvido de acordo com o modelo de arquitetura que as ferramentas criadas para este sistema devem seguir (ALVES, 2007). Ele implementa as seguintes camadas:

- **Camada de apresentação:** A ferramenta disponibiliza uma página *Web* nesta camada, na qual os usuários podem clicar no *link* de ativação do programa cliente, baixar e iniciá-lo. A partir do clique no *link* de ativação, o *download* e o início do cliente são realizados automaticamente por meio da tecnologia *Java Web Start*;
- **Componentes:** A ferramenta segue a arquitetura ao publicar componentes no serviço de diretório do sistema Tidia-Ae. Estes componentes implementam parte da lógica de negócio da ferramenta, isto é, a autenticação e a gerência das salas de reunião. É importante ressaltar que a lógica referente ao processamento de áudio é efetuada no programa cliente;
- **Camada de persistência:** Nesta camada está implementado o componente que realiza a persistência de objetos utilizados pela ferramenta. Sua importância é garantir a escalabilidade e consistência na gerência das salas. Isto é realizado por meio da utilização do *framework* do sistema Tidia-Ae, que fornece conexões ao SGBD configurado para todo o sistema. Um dos principais benefícios desta utilização é aproveitar o controle transacional oferecido pelo Tidia-Ae.

Nesta integração, uma das principais funcionalidades implementadas é a autenticação automática, necessária quando o programa cliente se conecta ao

programa servidor. Isto ocorre devido ao fato do programa cliente, que é um novo processo na máquina do usuário, necessitar a abertura de uma nova conexão TCP com o programa servidor uma vez que não é possível aproveitar a conexão já aberta entre o navegador *web* do usuário e o sistema Tidia-Ae. A partir do momento que uma nova conexão é criada, o servidor precisa identificar quem está se conectando e, devido a este motivo, é necessário autenticar novamente a entidade remota.

A autenticação automática exige que o programa cliente transmita ao programa servidor a identificação da sessão do sistema Tidia-Ae utilizada pelo usuário. Essa identificação da sessão é um dos parâmetros de inicialização do programa cliente, a ele fornecida pelo navegador *web*. Em seguida, ocorre a ativação da integração quando o programa servidor, munido da identificação, se comunica com o sistema Tidia-Ae, que então valida e autentica a sessão do usuário.

Ainda como parte desta integração, está a funcionalidade de ofertas de salas aumentadas contextualizadas em *worksites*, isto é, estão disponíveis duas salas por cada *worksite*. Estas salas apresentam duas disposições de assento: de apresentação (*Presentation*), a ser utilizada em situações de aula; de reunião (*Meeting*), para utilização em situações de reunião e discussão. A Figura 17 ilustra uma sala aumentada de apresentação, no *worksite* “Teste de Áudio 3D”, onde dois participantes estão se comunicando: o Romeo (participante corrente) e o Alfredo (participante remoto).

Figura 17 - Tela capturada de uma sala de apresentação no *worksite* "Teste de Áudio 3D".

Já a Figura 18 ilustra uma sala aumentada de reunião para o mesmo *worksite*, mas sem usuários presentes no momento da captura.

Figura 18 – Tela capturada com a sala de reunião do *workspace* “Teste de Áudio 3D” selecionada.

Pelo fato do programa cliente ser um processo na máquina de acesso utilizado pelo usuário e não um *software* que é executado pelo *browser*, o programa cliente continua funcionando mesmo quando o usuário muda de *workspace*. No entanto, para continuar sua operação em conjunto com o navegador do usuário, a ferramenta de áudio conferência 3D passa a disponibilizar as salas também para o *workspace* recém escolhido pelo usuário. A Figura 19 ilustra essa situação. No caso, o participante estava no *workspace* “Teste de Áudio 3D”, quando mudou para o “PCS2307_2Sem_2009_Turma50”.

Figura 19 – Tela capturada após o usuário mudar do *workspace* “Teste de Áudio 3D” para o “PCS2307_2Sem_2009_Turma50”.

Outro benefício provido pelo fato do programa cliente ser um processo e não um componente de página *web*, é que se o usuário mudar de *workspace* durante uma áudio conferência em andamento, ela não será abruptamente terminada, somente mais salas aumentadas serão disponibilizadas. Esta situação também é ilustrada pela Figura 16, onde outro *workspace* foi selecionado no *browser*, quando o participante corrente (Romeo) estava se comunicando na sala *Presentation* do *workspace* “Teste de Áudio 3D”.

As funcionalidades apresentadas neste capítulo concluem as principais funcionalidades oferecidas pela ferramenta neste trabalho. É importante observar que existem outras funcionalidades implementadas como: a pré seleção do dispositivo de áudio a ser utilizado, de modo que o usuário não precise selecionar o *hardware* de áudio toda vez que o programa cliente for iniciado; a funcionalidade de exibir a sala aumentada sob a perspectiva gráfica do usuário, de modo a facilitar a compreensão da disposição virtual dos avatares remotos em relação ao do usuário

corrente; a renderização da planta baixa que rege o *layout* das salas aumentadas por meio de bibliotecas e funcionalidades do OpenGL e, finalmente, a funcionalidade para comunicação de dados que permite a transmissão e recepção de pacotes UDP mesmo com a máquina de acesso configurada com endereço de rede privado (IP privado).

Finalmente, a fim de verificar o funcionamento correto da implementação realizada, visto que a percepção de áudio espacial depende de fatores humanos, experimentos foram conduzidos de modo a analisar esse funcionamento. Estes estão relatados no capítulo a seguir.

5 EXPERIÊNCIA DE UTILIZAÇÃO DE ÁUDIO ESPACIAL EM AMBIENTES DE APRENDIZAGEM ELETRÔNICA

Os experimentos realizados visam testar a efetividade dos efeitos tridimensionais de áudio na capacidade de se localizar e manter a concentração numa fonte sonora em um ambiente de interação (sala aumentada) com diversos participantes. As experiências deste trabalho são similares às aquelas descritas em (DOERR et al, 2007) e em (BALDIS, 2001).

A seguir são explanadas as condições de preparação para os experimentos e a realização dos mesmos. A análise e interpretação dos resultados é apresentada ao final de cada um dos experimentos.

5.1 CONDIÇÕES GERAIS PARA A EXECUÇÃO DOS EXPERIMENTOS

Com a finalidade de executar os experimentos, vêm-se necessárias condições de preparação que configuram uma base comum para a realização dos experimentos. Esta base comum resulta em um ambiente de testes para ser utilizado nos dois experimentos descritos neste capítulo.

O ambiente de testes é preparado por meio da configuração de uma sala aumentada da Ferramenta de Áudio Conferência 3D, permitindo o posicionamento do participante real que avaliará o sistema a ser testado e dos participantes virtuais nos assentos da sala que participarão dos testes realizados.

A Figura 20 ilustra o posicionamento dos participantes virtuais, que estão identificados sob os seguintes nomes: Debora, Paulo, Laerte e Oswaldo.

Figura 20 – Imagem ampliada da disposição dos participantes artificiais (Laerte, Paulo, Debora e Oswaldo) em relação ao usuário avaliador (Romeo neste caso).

Para cada participante virtual é injetada uma trilha de áudio na entrada de seu microfone, que consiste em um monólogo previamente gravado, caracterizando uma das múltiplas fontes sonoras que existirão no ambiente de interação. As trilhas são reproduzidas simultaneamente, ou seja, o avaliador escuta todos participantes virtuais ao mesmo tempo. É importante ressaltar que a reprodução das trilhas é

sincronizada. Isto permite a homogeneidade dos experimentos realizados com diferentes participantes avaliadores.

Antes de realizar as perguntas ao usuário avaliador do teste para obter os resultados da experiência, o ambiente de testes deve ser preparado de modo a atender às seguintes condições:

- Condição 1: O participante avaliador deve desconhecer seu posicionamento virtual no ambiente aumentado virtual.
- Condição 2: O participante avaliador deve desconhecer o posicionamento dos demais participantes virtuais.
- Condição 3: Os posicionamentos definidos para o avaliador e para os participantes virtuais devem continuar os mesmos, a fim de estabelecer uma base comum para os testes. Como ilustra a Figura 17, o avaliador sempre assume o posicionamento da cadeira azul (posição central inferior).

Novamente, estas condições constituem a base comum para a realização dos dois experimentos deste trabalho e seus testes.

5.2 EXPERIMENTOS

Foram realizados dois experimentos, cada um deles descritos a seguir nos itens 5.3 e 5.4.

5.3 EXPERIMENTO DE SIMPLES IDENTIFICAÇÃO E CONCENTRAÇÃO DE FONTES DE ÁUDIO

O objetivo deste experimento é testar se o avaliador é capaz de quantificar o número de participantes e manter a concentração no áudio produzido por um dos usuários virtuais.

Inicialmente, as trilhas de áudio (com características espaciais) dos participantes virtuais são reproduzidas por aproximadamente um minuto (etapa A). Em seguida o avaliador deve responder ao questionário apresentado no item 5.3.1.

Concluída essa etapa, o experimento é repetido reproduzindo-se o áudio dos participantes virtuais sem o uso do processamento espacial de áudio 3D (etapa B). Isso faz com que as trilhas de áudio sejam reproduzidas como em uma ferramenta de áudio conferência simples. Ao final, o avaliador deve responder novamente às mesmas perguntas feitas anteriormente.

Visto que neste experimento os participantes respondem o mesmo questionário duas vezes, uma vez após escutar o áudio com processamento 3D e, depois, ao escutar as mesmas trilhas de áudio sem o processamento espacial, foi possível confrontar os resultados obtidos e verificar a efetividade da situação com áudio 3D. É importante ressaltar que o avaliador não sabia se o processamento de áudio espacial estava ligado ou não, em qualquer etapa.

5.3.1 Questionários e observações para os participantes

5.3.1.1 Verificação de número de fontes de áudio

Pergunta: Quantas pessoas você conseguiu identificar?

Escopo da resposta: Irrestrita, o participante avaliador pode fornecer qualquer resposta numérica.

5.3.1.2 Capacidade de concentração nas fontes de áudio

Pergunta: Qual o grau de dificuldade encontrada para manter a concentração em uma das fontes de áudio?

Escopo da resposta: O participante avaliador deve escolher uma das seguintes alternativas: 1- Muito Fácil; 2- Fácil; 3- Razoável; 4- Difícil e 5- Muito Difícil.

5.3.2 Coleta de Dados e Resultados Obtidos

Participaram deste experimento sete avaliadores cujas respostas forneceram os resultados para a análise da ferramenta de áudio 3D.

As tabelas e gráficos que seguem indicam os dois tipos de testes executados:

- Coluna A- Quando o processamento espacial (3D) de áudio estava ligado;
- Coluna B- Quando o processamento de áudio espacial (3D) estava desligado, ou seja, simples mixagem de áudio sem os efeitos disponibilizados pela ferramenta.

Abaixo a Tabela 3 sumariza as respostas fornecidas com relação à pergunta indicada sobre o número de participantes virtuais identificados (item 5.3.1.1). Os dados exibidos são: média da contagem de participantes virtuais identificados, a menor quantidade de participantes virtuais identificada, a maior quantidade de participantes virtuais identificada, o desvio padrão da contagem e, finalmente, o erro da média da contagem.

Tabela 3 - Resultados da contagem de participantes artificiais.

Campo	A	B
	3D Ligado	3D Desligado
Contagem média de pessoas identificadas:	3,86	3,14
Menor número indicado na contagem:	3	3
Maior número indicado na contagem:	4	4
Desvio Padrão da contagem:	0,35	0,35
Erro da média da contagem:	3,6%	21,4%

A seguir, a Tabela 4 sumariza as respostas fornecidas com relação à pergunta 5.3.1.2 sobre o grau de dificuldade encontrado pelos participantes ao manter a concentração no áudio de um dos participantes virtuais. Os dados exibidos são: média da dificuldade relatada pelos participantes, a menor dificuldade encontrada, a maior dificuldade encontrada e, finalmente, o desvio padrão das dificuldades indicadas.

Tabela 4 - Resultados do grau de dificuldade encontrado.

Cálculo	A	B
	3D Ligado	3D Desligado
Média de dificuldade:	2,57	3,43
Menor dificuldade encontrada:	1	1
Maior dificuldade encontrada:	4	4
Desvio Padrão da dificuldade:	0,90	1,05

A seguir o Gráfico 1 mostra o agrupamento dos dados de erro da Tabela 3:

Gráfico 1 - Gráfico contendo os dados da média de contagem dos participantes artificiais identificados pelos participantes testados.

Já o Gráfico 2 exibe o agrupamento dos dados da Tabela 4:

Gráfico 2 - Gráfico contendo os dados da média do grau de dificuldade de concentração encontrada pelos participantes.

A análise dos resultados indicados na Tabela 2 e Gráfico 1, sobre a contagem de participantes virtuais, mostra que a identificação do número de usuários remotos foi mais precisa quando o processamento espacial de áudio estava ligado, pois a média de contagem se aproximou da quantidade real de participantes artificiais

presentes no ambiente aumentado (quatro).

Já na análise dos resultados indicados na Tabela 3 e Gráfico 2, com relação ao grau de dificuldade de concentração encontrado pelos participantes, pode-se constatar que ela foi mais facilmente dispensada aos usuários artificiais com o uso do processamento espacial do que sem ele.

Os resultados deste experimento, de um modo geral, indicaram diferenças quando o processamento de áudio espacial estava ligado e quando estava desligado. Nos dois testes realizados, a utilização de áudio espacial evidenciou melhorias, o que sugere sua eficiência de funcionamento nesse tipo de ambiente.

5.4 EXPERIMENTO INTEGRADO DE CONTAGEM, IDENTIFICAÇÃO E COMPREENSÃO

Este experimento foi inspirado no anterior, mas desta vez visando testar com uma amostragem de participantes maior e com um questionário mais completo, contendo questões adicionais de compreensão, as quais não estavam presentes no primeiro experimento.

Foram utilizadas doze (12) trilhas de áudio diferentes. Oito delas continham o áudio de professores de cursos para vestibular que lecionam assuntos diferentes e que são equivalentes com relação à didática que utilizam. As outras quatro são o áudio de profissionais de televisão que relatam seu envolvimento pessoal com uma famosa emissora.

A preparação indicada no item 5.1 foi utilizada para a realização deste. Cada participante avaliador realizou três etapas do experimento.

Na primeira, denominada etapa 1, quatro trechos de áudio com professores de curso de vestibular foram reproduzidos simultaneamente, no máximo duas vezes, desprovidos de processamento espacial de áudio. O avaliador foi instruído a prestar atenção em um único palestrante, o qual ele deveria identificar pelo tema da palestra, que era “alotropia” (química). Ao final, o avaliador respondeu a algumas questões. Nesta etapa, o avaliador não foi informado se o processamento espacial estava funcionando ou não.

Na segunda etapa (etapa 2), outros quatro trechos de áudio de professores

de curso de vestibular foram reproduzidos simultaneamente, mas desta vez com o processamento espacial ligado, duas vezes no máximo. Novamente, o participante deveria prestar atenção em um único palestrante, o qual leciona sobre física óptica. Depois de escutar os trechos de áudio, o avaliador respondeu a algumas questões idênticas às apresentadas na primeira etapa e outra questão diferente das anteriores, por ser uma do tipo contextual. Foi necessário utilizar trechos de áudio diferentes nesta etapa, pois na realização do experimento era fundamental a certeza de que os avaliadores enfrentassem a mesma dificuldade encontrada na etapa anterior para responder o questionário desta etapa, ou seja, se os mesmos trechos de áudio fossem utilizados, os avaliadores os escutariam pela segunda vez, o que poderia distorcer as respostas coletadas. Assim como na etapa anterior, o avaliador não sabia se o processamento espacial estava ligado ou não.

Já na terceira e última etapa (etapa 3), mais outros trechos de áudio (profissionais de televisão) foram reproduzidos com o processamento de áudio espacial desligado e, em seguida, eles foram novamente reproduzidos, mas com o processamento espacial ligado. Ao final desta etapa também havia um questionário a ser respondido. Neste, os avaliadores puderam escutar, quantas vezes desejassem, os trechos de áudio disponibilizados com o processamento espacial ligado ou desligado. Diferentemente das etapas 1 e 2, o avaliador sabia se o processamento de áudio espacial estava ligado ou desligado.

O questionário utilizado continha perguntas de compreensão literal, contextual, e geral e, também, de percepção espacial. De todas as questões elaboradas para este experimento, algumas só se aplicaram a uma única etapa, como é o caso das questões contextuais sobre física e química, pois são relevantes somente ao assunto específico de um trecho de áudio.

A seguir, no item 5.4.1, são descritas as questões utilizadas neste experimento.

5.4.1 Questionários e observações para os participantes

Ao total, foram elaboradas dez questões para os participantes responderem. Todas são de múltipla escolha, com sete alternativas oferecidas. Cada uma delas é descrita, apresentando a pergunta proposta aos participantes testados, e o escopo

das opções de respostas possíveis. Em alguns casos, existe uma resposta correta, em outros não, pois se trata da opinião pessoal do participante.

5.4.1.1 Verificação de contagem de fontes de áudio

Pergunta: Quantos palestrantes você conseguiu perceber?

Escopo da resposta: O participante testado deve escolher umas das sete alternativas que representam de 1 a 7 palestrantes.

5.4.1.2 Grau de certeza sobre a quantificação de palestrantes fornecida

Pergunta: Como você mediria seu grau de certeza, com relação à quantidade de palestrantes que você percebeu? (Escala de 1 a 7)

Escopo da resposta: O participante deve escolher uma alternativa, na escala de 1 a 7 que represente seu grau de certeza com relação à quantidade de palestrantes indicadas na questão 5.4.1.1. Nesta escala, o valor 1 representa “Certeza absoluta” e 7 representa “Nenhuma certeza”.

5.4.1.3 Grau de dificuldade de concentração nas fontes de áudio

Pergunta: Em sua opinião, qual o grau de dificuldade encontrado para manter a concentração no palestrante especificado? (Escala de 1 a 7)

Escopo da resposta: O participante testado deve escolher uma das alternativas na escala de 1 a 7, onde o valor 1 representa “Extremamente Difícil” e o valor 7 representa “Não houve dificuldade”.

5.4.1.4 Compreensão literal

Pergunta: Quanto tempo você conseguiu escutar (palavra a palavra) do que foi falado pelo palestrante especificado? (Escala de 1 a 7)

Escopo da resposta: O participante deve escolher na escala de 1 a 7, uma das seguintes respostas que ele considere representar seu entendimento:

- 7- O tempo todo
- 6- Entre 85% e 70% do tempo
- 5- Entre 70% e 55% do tempo
- 4- Metade do tempo
- 3- Entre 40% e 25% do tempo
- 2- Entre 25% e 10% do tempo
- 1- Praticamente nada

5.4.1.5 Compreensão geral de um palestrante

Pergunta: Quanto você conseguiu aprender do conteúdo ensinado pelo palestrante especificado? (Escala de 1 a 7)

Escopo da resposta: O participante deve escolher, seguindo a escala de 1 a 7, uma das seguintes alternativas que melhor represente seu entendimento:

- 7- Tudo
- 6- Grande parte do conteúdo
- 5- Parte significativa do conteúdo
- 4- Parte do conteúdo
- 3- Pouco conteúdo
- 2- Pouquíssimo conteúdo
- 1- Nada

5.4.1.6 Percepção de posicionamento espacial das trilhas de áudio (palestrantes artificiais)

Pergunta: Você notou alguma diferença no posicionamento espacial entre os palestrantes? Por ex.: acima, abaixo, à esquerda, à direita, ...

Escopo da resposta: Somente duas alternativas: Sim ou Não.

5.4.1.7 Compreensão contextual sobre Química (Alotropia)

Pergunta: O palestrante recapitula um assunto por ele considerado importante, antes de abordar seu tema principal. Qual é o assunto abordado?

Escopo da resposta: O participante deve escolher umas das sete alternativas:

- Pressão atmosférica
- Estados físicos (sólido, líquido e gasoso)
- Química inorgânica
- Escalas de temperatura
- Estruturas moleculares
- Ligações covalentes
- Não sei

É importante ressaltar que por ser uma questão de compreensão contextual, no diálogo apresentado, existe uma única resposta certa, que é: Estados físicos (sólido, líquido e gasoso).

5.4.1.8 Compreensão contextual sobre Física (Óptica)

Pergunta: No caso do tema especificado, qual opção melhor descreve o assunto abordado pelo palestrante?

Escopo da resposta: Uma das sete alternativas a seguir deve ser escolhida:

- Espelhos e lentes
- Fibra óptica
- Espelhos esféricos
- Espelhos planos
- Lentes convergentes e divergentes
- Fenômeno de espalhamento espectral
- Não sei

Nesta questão, novamente é importante ressaltar que por se tratar de compreensão contextual, no diálogo apresentado, existe uma única resposta certa, que é: Espelhos esféricos.

5.4.1.9 Compreensão contextual sobre o local de trabalho dos palestrantes

Pergunta: Os palestrantes relatam seu envolvimento profissional com qual emissora de televisão?

Escopo da resposta: O participante deve escolher uma das sete alternativas:

- TV Cultura e Rede Globo
- Rede Globo e Bandeirantes
- TV Cultura e Bandeirantes
- TV Cultura e Rede Record
- Rede Record e Rede Globo
- Somente TV Cultura
- Não consegui identificar

Visto que este é, mais uma vez, uma questão contextual, só há uma alternativa correta, que é: Somente TV Cultura.

5.4.1.10 Eficiência do Áudio Espacial

Pergunta: Na escala de 1 a 7, quanto o processamento espacial (3D) de áudio ajudou seu entendimento do que foi dito pelos palestrantes?

Escopo da resposta: Considerando a escala de 1 a 7, o participante deve escolher a alternativa que melhor represente a sua percepção, onde o valor 1 representa “Não ajudou” e o valor 7 representa “Ajudou muito”.

É importante ressaltar que parte das questões é oferecida ao participante mais de uma vez, mas em etapas diferentes. Outras, são oferecidas exclusivamente em uma etapa específica. A fim de organizar quais perguntas são utilizadas nas etapas, é apresentada a Tabela 5, que relaciona as perguntas e as etapas.

Tabela 5- Questões e etapas do segundo experimento

Questão	Etapa 1	Etapa 2	Etapa 3
5.4.1.1 Verificação de contagem de fontes de áudio	O	O	O
5.4.1.2 Grau de certeza sobre a quantificação de palestrantes fornecida	O	O	
5.4.1.3 Grau de dificuldade de concentração nas fontes de áudio	O	O	

5.4.1.4 Compreensão literal	O	O	
5.4.1.5 Compreensão geral de um palestrante	O	O	
5.4.1.6 - Percepção de posicionamento espacial das trilhas de áudio (palestrantes artificiais)	O	O	
5.4.1.7 Compreensão contextual sobre Química (Alotropia)	O		
5.4.1.8 Compreensão contextual sobre Física (Óptica)		O	
5.4.1.9 Compreensão contextual sobre o local de trabalho dos palestrantes			O
5.4.1.10 - Eficiência do Áudio Espacial			O

De acordo com esta tabela, cada participante respondeu a dezessete perguntas, sendo que somente quatro delas foram respondidas unicamente. Destas quatro, três podem ser consideradas como a mesma, visto que são questões que indagam o participante sobre sua compreensão contextual.

No próximo item, este assunto é melhor descrito uma vez que nele são abordadas a coleta de dados e a interpretação os resultados obtidos.

5.4.2 Coleta de Dados e Resultados Obtidos

Antes de prosseguir com as análises vê-se necessário, neste momento, a formalização do termo amostra. Cada resposta fornecida por um participante à uma questão proposta é definido como amostra.

O experimento foi realizado com 23 participantes diferentes cujas respostas forneceram os resultados para a análise da ferramenta de áudio 3D.

Os participantes podem ser classificados como um grupo constituído em sua grande maioria (aproximadamente 95%) por alunos de graduação que cursavam, na realização do experimento, o segundo ano de engenharia elétrica na Escola Politécnica da USP. Isto delimita o escopo do experimento para o tipo de participante testado mas, em contrapartida, testa parte do principal público-alvo desta ferramenta, que são os estudantes.

As amostras coletadas foram agrupadas de diversos modos diferentes, totalizando três grandes análises: confronto direto entre as etapas 1 e 2; análise da etapa 3 e, a última, observação global de todas as etapas.

Com a finalidade de melhor organizar as análises e os resultados, cada uma é descrita em um dos sub itens a seguir.

5.4.2.1 Análise de confronto direto entre as etapas 1 e 2

Para esta análise, foram observadas as respostas fornecidas nas questões 5.4.1.1 à 5.4.1.6 e as respostas das questões de compreensão contextual 5.4.1.7 e 5.4.1.8.

Ainda nesta, são confrontados diretamente os dados estatísticos de todas as amostras da mesma questão agrupadas por etapa. Isto é, são confrontados os dados estatísticos das amostras da questão 5.4.1.1 na etapa 1 com os dados estatísticos das amostras para a mesma questão na etapa 2. O mesmo confronto é feito para a questão 5.4.1.2 e assim por diante.

É importante lembrar que na etapa 1 o processamento de áudio espacial está desligado e, na etapa 2, está ligado. Portanto, ao confrontar os resultados da etapa 1 com os da etapa 2 se deseja verificar, na verdade, a eficácia do processamento de áudio 3D.

Começando pela questão 5.4.1.1, a Tabela 6 a seguir agrupa os dados estatísticos das amostras:

Tabela 6 – Dados estatísticos das amostras da questão 5.4.1.1 (Etapas 1 e 2)

Cálculo	Etapa 1	Etapa 2
Média:	3,91	3,52
Menor valor:	2	3
Maior valor:	5	5
Desvio padrão:	0,73	0,59

Abaixo está o Gráfico 3, representando os dados da Tabela 6:

Gráfico 3 – Média da quantidade de palestrantes percebidos.

Como é possível observar, as amostras fornecidas representam a quantidade de palestrantes percebidos pelo participante. Nesta questão, há uma resposta certa que é 4 (quatro), pois existem, tanto na etapa 1 quanto na etapa 2, quatro palestrantes. Pode-se verificar que na etapa 1, com o processamento 3D desligado, a resposta mais se aproximou da correta, embora a média para a outra etapa seja próxima a 4. Apesar deste resultado não ser positivo a este experimento, o desvio padrão mostrou o contrário, isto é, a variação das amostras foi menor quando os efeitos 3D estavam funcionais (0,59). Há também que ser observadas as amostras referentes à questão 5.4.1.2, pois esta demanda do participante relatar seu grau de certeza para a quantificação de palestrantes indicada.

Com relação à questão 5.4.1.2, a Tabela 7 contém os dados estatísticos das amostras coletadas:

Tabela 7 – Dados estatísticos das amostras da questão 5.4.1.2 (Etapas 1 e 2)

Cálculo	Etapa 1	Etapa 2
Média:	4,91	3,09
Menor valor:	3	1
Maior valor:	7	6
Desvio padrão:	1,16	1,53

A seguir está o Gráfico 4, representando as médias calculadas das amostras contidas na Tabela 7:

Gráfico 4 – Média do grau de certeza da contagem de palestrantes percebidos.

Como pode ser observada, a média é menor na segunda etapa, onde o processamento de áudio espacial está ligado. Visto que na escala proposta, o valor 1 representa “Certeza Absoluta” e o valor 7 significa “Nenhuma certeza”, pode-se verificar que a sensação de certeza na etapa 2 foi maior do que na etapa 1. Portanto, o processamento de áudio espacial aumentou o grau de certeza dos participantes com relação à contagem de palestrantes no ambiente aumentado.

À luz deste esclarecimento, pode-se concluir que as amostras da questão 5.4.1.1 para a etapa 2 são mais precisas do que as coletadas na etapa 1. Ou seja, apesar dos resultados se aproximarem mais do verdadeiro na etapa 1 (média de 3,91), as amostras da etapa 2 (média de 3,52) são mais precisas e têm menor variação.

Já considerando as amostras à questão 5.4.1.3, a Tabela 8 apresenta os cálculos estatísticos:

Tabela 8 – Dados estatísticos das amostras da questão 5.4.1.3 (Etapas 1 e 2)

Cálculo	Etapa 1	Etapa 2
Média:	2,04	4,61
Menor valor:	1	1
Maior valor:	6	7
Desvio padrão:	1,19	1,50

A seguir, o Gráfico 5 representa a média da Tabela 8:

Gráfico 5 – Média da dificuldade enfrentada para manter a concentração no palestrante indicado.

Como exibido pelo Gráfico 5, a média para a etapa 1 foi menor que a média para a etapa 2. Lembrando-se que, nesta escala, o valor 1 representa “Extremamente Difícil” e o valor 7 representa “Não houve dificuldade”, pode-se concluir que a dificuldade foi muito menor na etapa 2, quando o processamento espacial é funcional. Não só o resultado da etapa 2 é menor, como é numericamente maior que o dobro da média da etapa 1.

Com relação às amostras coletadas para a questão 5.4.1.4, são exibidos na Tabela 9 os dados estatísticos:

Tabela 9 – Dados estatísticos das amostras da questão 5.4.1.4 (Etapas 1 e 2)

Cálculo	Etapa 1	Etapa 2
Média:	2,26	5,61
Menor valor:	1	1
Maior valor:	6	7
Desvio padrão:	1,14	1,62

A seguir o Gráfico 6 representa as médias da Tabela 8:

Gráfico 6 – Média para a quantificação aproximada do tempo de compreensão literal.

É importante ressaltar que, para esta questão, a quantificação do tempo não foi requisitada em nenhuma unidade de tempo (minutos ou segundos), mas sim em porcentagem do tempo. O participante podia escolher faixas de tempo (vide item 5.4.1.4) que melhor representassem sua percepção. Estas faixas foram divididas na escala de 1 a 7, onde 7 representa “O tempo todo” e 1 representa “Praticamente nada”. A luz deste esclarecimento, pode-se confirmar que a compreensão literal do palestrante especificado foi significativamente maior na etapa 2, quando o processamento espacial de áudio está ligado, isto é, os efeitos de áudio 3D melhoraram a compreensão literal nesta situação.

A Tabela 10 apresenta os dados estatísticos para as amostras coletadas da questão 5.4.1.5:

Tabela 10 – Dados estatísticos das amostras da questão 5.4.1.5 (Etapas 1 e 2)

Cálculo	Etapa 1	Etapa 2
Média:	1,78	5,05
Menor valor:	1	1
Maior valor:	5	7
Desvio padrão:	1,04	1,96

E, o Gráfico 7 apresenta as médias da Tabela 9, lado a lado:

Gráfico 7 – Média da compreensão de conteúdo percebida pelo participante.

Para esta questão, a compreensão de conteúdo geral é julgada pelo próprio participante, ou seja, se trata de uma auto-avaliação buscando quantificar a retenção de conteúdo. Observando-se que os valores estão dispostos em uma escala onde valor 1 representa “Nada” e o valor 7 representa “Tudo”, notavelmente a percepção de compreensão de conteúdo foi maior na etapa 2 (quase três vezes), onde o processamento 3D está ligado.

A seguir, a Tabela 11 exibe os totais das amostras coletadas para a pergunta indicada no item 5.4.1.6 sobre a percepção do posicionamento espacial dos palestrantes artificiais. Os valores indicam a quantidade de participantes que optou pela resposta indicada na primeira coluna. Já as colunas seguintes indicam em qual etapa foi feita a opção:

Tabela 11 – Amostras da questão 5.4.1.6 (Etapas 1 e 2)

Opção	Etapa 1	Etapa 2
Sim:	14	22
Não:	9	1

Já o Gráfico 8 agrupa essas amostras, de modo a facilitar a visualização desses dados em porcentagem:

Gráfico 8 – Representação gráfica da percepção de funcionamento de processamento espacial.

Como é possível verificar no Gráfico 8, a porcentagem que percebeu o funcionamento de processamento espacial na etapa 2 foi sensivelmente maior do que na etapa anterior. Isto significa que houve percepção espacial de áudio. Com relação somente às amostras na etapa 1, talvez pelo fato dos participantes terem conhecimento que o experimento era de áudio 3D, se sentiram, induzidos a responder sim, quando não havia processamento espacial.

Finalmente, as Tabelas 12 e 13 apresentam as amostras coletadas das questões 5.4.1.7 e 5.4.1.8 respectivamente. Nestas, são quantificados os participantes que responderam a opção indicada na primeira coluna.

Tabela 12 – Amostras da questão 5.4.1.7 (Etapa 1 somente)

Opção	Etapa 1
Pressão atmosférica	0
Estados físicos (sólido, líquido e gasoso)	2
Química inorgânica	2
Escala de temperatura	1
Estruturas moleculares	3
Ligações covalentes	1
Não sei	14

Tabela 13– Amostras da questão 5.4.1.8 (Etapa 2 somente)

Opção	Etapa 2
Espelhos e lentes	0
Fibra óptica	0
Espelhos esféricos	19
Espelhos planos	0
Lentes convergentes e divergentes	1

Fenômeno de espalhamento espectral	0
Não sei	3

Observando-se que cada uma destas questões possui uma resposta correta, que é “Estados físicos (sólido, líquido e gasoso)” na questão 5.4.1.7 e “Espelhos esféricos” no caso da questão 5.4.1.8, a Tabela 14 exhibe a quantidade de respostas corretas (acertos) e incorretas (erros) fornecidas à estas questões respectivamente:

Tabela 14 – Respostas corretas à questão 5.4.1.7 (Etapa 1 somente)

	Etapa 1	Etapa 2
	Questão 5.4.1.7	Questão 5.4.1.8
Acertos	2	19
Erros	21	4

O Gráfico 9 agrupa os dados da Tabela 14, com a finalidade de confrontar a quantidade de acertos na primeira etapa *versus* a quantidade de acertos na segunda etapa. Novamente, é importante esclarecer que as questões 5.4.1.7 e 5.4.1.8 são questões de compreensão contextual e, por este motivo, não podem ser indagadas nas duas etapas. Elas são equivalentes por serem consideradas de mesma dificuldade, mas cada uma postada em uma única etapa.

Gráfico 9 – Acertos e erros para as questões de compreensão contextuais 5.4.1.7 e 5.4.1.8.

Desta representação gráfica é possível entender que a quantidade de acertos foi maior na etapa 2 do que na etapa 1. Isto significa que a compreensão contextual torna-se mais fácil quando o processamento espacial está ligado, nas condições de execução deste experimento.

Nesta análise, foi possível conferir o perfeito funcionamento do processamento espacial de áudio, a partir dos resultados evidenciados pelo experimento e pelas amostras.

De fato, esse resultado era o esperado, pois também já foi examinado por outros autores.

5.4.2.2 Análise da Etapa 3

Nesta avaliação, somente são observadas as amostras coletadas referentes às questões da etapa 3. Visto que nesta etapa o participante faz o julgamento de confronto entre o mesmo trecho de áudio executado com funcionamento de áudio espacial e sem o processamento 3D, as amostras fornecidas constituem por si só este confronto.

Igualmente, é importante ressaltar que nos dois trechos de áudio estão presentes os mesmos palestrantes (com e sem processamento 3D) e podem ser consultados diversas vezes pelo participante. Isto permite que as respostas às questões propostas sejam mais precisas.

A Tabela 15 apresenta a média das amostras coletadas durante a etapa 3 para a questão 5.4.1.1.

Tabela 15 – Dados estatísticos das amostras da questão 5.4.1.1 (Etapa 3)

Cálculo	Etapa 3
Média:	3,87
Menor valor:	3
Maior valor:	5
Desvio padrão:	0,46

O Gráfico 10 representa os a média da Tabela 15, juntamente com as médias para as etapas 1 e 2, já apresentadas na Tabela 5.

Gráfico 10 – Palestrantes percebidos nas etapas 1, 2 e 3 (destacado).

Como se pode observar, o resultado não foi muito diferente dos obtidos nas etapas anteriores. O resultado aproximou-se mais de 4, como na etapa 1, mas com um desvio padrão menor ainda ao obtido na etapa 2 (0,59). Isto significa que grande parte das amostras está muito próxima à média obtida, ao contrário da etapa 1, onde há uma variação maior, como já comentado na análise anterior. Isto torna a média de contagem de palestrantes a mais precisa entre as três etapas.

A tabela 16 representa os dados das amostras coletadas da questão contextual postada na etapa 3. Esta, como já relatado anteriormente, deveria ser mais precisa devido ao fato do participante poder consultar inúmeras vezes os trechos de áudio oferecidos.

Tabela 16 – Amostras da questão 5.4.1.9 (Etapa 3)

Opção	Etapa 3
Tv Cultura e Rede Globo	2
Rede Globo e Bandeirantes	0
TV Cultura e Bandeirantes	0
TV Cultura e Rede Record	0
Rede Record e Rede Globo	0
Somente TV Cultura	17
Não consegui identificar	4

A Tabela 17 agrupa os resultados da Tabela 16 em respostas corretas e incorretas, visto que este tipo de questão (contextual) há somente uma alternativa correta, que é “Somente TV Cultura”.

Tabela 17 – Respostas corretas à questão 5.4.1.9 (Etapa 3)

Etapa 3	
Questão 5.4.1.9	
Acertos	17
Erros	6

O Gráfico 11 exibe os dados da Tabela 16, para representar graficamente a porcentagem de acertos e erros para a questão contextual 5.4.1.9.

Gráfico 11 – Acertos e erros para a questão 5.4.1.9

A taxa de acertos foi de aproximadamente 74%, visto que o participante deve responder uma pergunta contextual que envolve ouvir e entender o que todos os quatro palestrantes estão falando. Infelizmente, este resultado não pode ser comparado ao obtido para as questões contextuais das etapas 1 e 2, pois nestes foi requisitado que o participante somente se concentrasse em um dos monólogos e não em todos. Todavia, pode-se concluir que os resultados para esta questão foram satisfatórios.

Na Tabela 18 são expostos os dados amostrados para a pergunta 5.1.4.10, que indaga o participante a quantificar o auxílio provido pelo processamento de áudio 3D.

Tabela 18 – Dados estatísticos das amostras da questão 5.4.1.10 (Etapa 3)

Cálculo	Etapa 3
Média:	6,04
Menor valor:	4
Maior valor:	7
Desvio padrão:	1,15

E, o Gráfico 12, exibe as opções selecionadas pelos participantes, demonstrando a porcentagem representada pela alternativa 7 (Ajudou muito), em

relação às demais.

Gráfico 12 – Eficiência do áudio espacial (questão 5.4.1.10 – etapa 3).

Como é possível observar, a opção “Ajudou muito”, representa sozinha aproximadamente metade das amostras. O próximo valor na escala, o valor 6, também é muito significativo (um quarto do total). Isto significa que o processamento espacial de áudio foi muito eficiente, visto que os valores 6 e 7 juntos, correspondem a 74% das amostras.

5.4.2.3 Observação Geral de Todas as Etapas

Como uma observação geral envolvendo todas as etapas de coletas de amostras, comparações e confrontos de dados estatísticos realizados nas análises, pode-se fazer comprovações e chegar a uma conclusão sobre este experimento.

Uma primeira comprovação é que os resultados, de um modo geral, demonstraram que a utilização de áudio espacial interfere na compreensão dos participantes testados. Em todas as etapas, toda vez que o áudio espacial estava acionado, o resultado destoou. Em alguns casos o dobro, em outros, quase o triplo.

Outra importante comprovação, baseada na interferência da utilização de processamento 3D, é que a compreensão não somente foi diferente, mas melhorou.

Exceto para as amostras da questão 5.4.1.1, cuja diferença da média foi pequena e minimizada pelas amostras da questão 5.4.1.2, todos os demais resultados demonstraram melhora.

Finalmente, pode-se concluir que a utilização de áudio 3D, na situação proposta pelo experimento e para o público testado, melhorou qualitativamente a compreensão deste grupo de participantes, tanto na compreensão literal como contextual.

6 CONCLUSÃO E TRABALHOS FUTUROS

A partir da análise dos resultados obtidos dos dois experimentos, foi possível observar que o funcionamento do áudio espacial melhora o entendimento e a capacidade de concentração dos participantes de uma áudio conferência. Aliado ao fato do trabalho realizado por (BALDIS, 2001) demonstrar que ambientes de áudio conferência espacial (3D) melhoram a inteligibilidade e a compreensão entre os conferencistas, pode-se concluir que a ferramenta apresentada neste trabalho também possui o potencial de melhorar essas mesmas características. Tais melhorias enriquecem a interatividade entre os participantes, possibilitando uma melhor experiência de atividades colaborativas remotas.

Visto que as melhorias estão relacionadas à compreensão dos participantes e, dado que a compreensão dos aprendizes é de fundamental importância para o processo de aprendizagem, torna-se possível aumentar a qualidade da aprendizagem à distância por meio da utilização de áudio conferência espacial.

Assumindo que o sistema aumentado pode ser enriquecido com novas possibilidades inviáveis ao mundo presencial, como a adição de elementos artificiais, é possível que este melhore e enriqueça as técnicas de ensino e aprendizagem, de modo análogo à utilização de ferramentas educacionais que implementam os conceitos de RA.

Portanto, ao se considerar as benesses proporcionadas pela utilização do processamento de áudio espacial à memorização dos participantes, a qual possibilita aumentar a qualidade da aprendizagem eletrônica e considerando um futuro aproveitamento mais extenso de mais recursos que implementem RA, conclui-se que a ferramenta de áudio conferência espacial representa uma importante contribuição para a aprendizagem eletrônica oferecida pelo sistema Tidia-Ae.

Já como um possível trabalho futuro, a ferramenta proposta pode ser avaliada com relação à sua efetividade no auxílio aos processos de ensino e aprendizagem à distância. Deste modo, a ferramenta poderia ser utilizada não somente para aprendizagem eletrônica, mas também para ensino à distância.

A gravação de uma interação na ferramenta pode ser um importante trabalho futuro, pois seria possível consultar e escutar novamente aulas já ministradas pelos

professores ou instrutores.

Outra possibilidade de trabalho futuro seria a evolução desta ferramenta de modo a inserir representações visuais dos avatares dos participantes remotos, como por exemplo, vídeo capturado por meio de *web-cams*. Munido desta representação, um instrutor que utiliza a ferramenta para ensino pode ter retorno visual dos alunos por meio de suas expressões faciais e todos podem ser favorecidos pela utilização de gestos durante a comunicação.

7 REFERÊNCIAS

ALVES, L. G. P. **Visão Geral da Arquitetura Sakai**. Disponível em <<http://tidia-ae.incubadora.fapesp.br/portal/docs/fase-ii-1/gerenciadores/doc-pesquisa/arquitetura-sakai/Visao%20Geral%20da%20Arquitetura%20Sakai.pdf/view>>. Maio de 2007. Acesso em 18/09/2009.

ARTOOLKIT. Disponível em: <<http://www.hitl.washington.edu/artoolkit/>> . Acesso em 24/07/2008.

ASIO: Audio Stream Input/Output. Disponível em: <http://www.steinberg.net/en/company/steinberg_technology.html>. Acesso em 17/07/2008.

AUSIM. Disponível em <<http://ausim3d.com/>>. Acesso em 16/07/2008.

AZUMA, R. **A Survey of Augmented Reality**. EUA – CA – Malibu: Computer Graphics (SIGGRAPH '95 Proceedings). Agosto de 1997. p. 1-38.

BALDIS, J. **Effects of Spatial Audio on Memory, Comprehension, and Preference during Desktop Conferences**. Conference on Human Factors in Computing Systems. 31 de Março a 5 de Abril de 2001, vol. 3, p. 166-173.

BILLINGHURST, M.; KATO H. **Collaborative Augmented Reality**. Communications of the ACM. Julho de 2002. vol. 45, No 7

BILLINGHURST, M.; KATO, H.; POUPYREV, I. **The MagicBook - moving seamlessly between reality and virtuality**. Computer Graphic and Applications, IEEE. Mai.-Jun. 2001. vol. 21, p. 6-8.

CHAN C.-T.; CHEN, O. T.-C. **A 3D Sound Using The Adaptive Head Model And Measured Pinna Data**. Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference. 30 de Julho a 2 de Agosto de 2000. vol. 2 p. 807 - 810.

DOERR K.; RADEMACHER H.; HUESGEN S.; KUBBAT W. **Evaluation of a Low-Cost 3D Sound System for Immersive Virtual Reality Training Systems**. IEEE Transactions on Visualization and Computer Graphics. Março e Abril de 2007. Vol. 13, No 2.

ELSEA, P. **Sound**. Nov. de 1995. Disponível em: <http://arts.ucsc.edu/EMS/Music/tech_background/tech_background.html> . Acesso em: 18/09/2009.

HOWARD, D. M.; ANGUS J. **Acoustics And Psychoacoustics**. Focal Press, 2006.

IAZZETTA, F. **Tutorias de Áudio e Acústica**. Disponível em: <<http://www.eca.usp.br/prof/iazzetta/tutor/>> . Acesso em: 16/09/2009.

JAVA Sound API. Disponível em: <<http://java.sun.com/products/java-media/sound/>> . Acesso em: 20/07/2008.

JONASSEN, D. H.; PECK, K. L.; WILSON, B. G. **Learning with Technology: A constructivist perspective**. Prentice Hall, 1999.

JUILLERAT, N.; MÜLLER, S.; SCHUBIGER-BANZ, S. **Real-time, low latency audio processing in Java**. Proceedings of the International Computer Music Conference. "Immersed Music". 2007. vol. 2. p. 99-102

KINSLEY L. E.; FREY A. R.; COPPENS A. B.; SANDERS J. V. **Fundamentals Of Acoustics**. John Willey & Sons Inc, 1999.

KIRNER, C. ; TORI, R. **Introdução à Realidade Virtual, Realidade Misturada e Hiper-realidade**. In: Claudio Kirner; Romero Tori. (Org.). Realidade Virtual: Conceitos e Tendências. São Paulo: Editora Mania de Livro, 2004. p. 3-20.

KOBAYASHI S. N.; ONOYE W.; SHIRIKAWA I. **DSP Implementation of Low Computational 3D Sound Localization Algorithm**. Signal Processing Systems IEEE Workshop. 26 a 28 de Setembro de 2001. p. 109-116.

KOZMA R. B. **Learning with media. Review of Education Research**. 1991. vol.

61, n. 2, p. 179-211.

LANDIM C. M. M. P. F. **Educação à Distância: Algumas Considerações**. Rio de Janeiro: 1997.

LINDEMAN R. W.; NOMA H.; BARROS P. G. **Hear-Through and Mic-Through Augmented Reality: Using Bone Conduction to Display Spatialized Audio**. 6th IEEE and ACM Symposium on Mixed and Augmented Reality (ISMAR). 2007.

LAGO, N. P.; KON, F. **The Quest for Low Latency**. Proceedings of the 30th International Computer Music Conference. Novembro de 2004. p. 33-36

MUELLER, W.; ULLMANN, F. **A Scalable System for 3D Audio Ray Tracing. Multimedia Computing and Systems**. IEEE International Conference on Volume 2, 7 à 11 de Junho de 1999. p. 819 - 823.

RECOMMENDATION ITU-T G.114. **One-Way Transmission Time**. International Telecommunication Union. Genebra, 1996.

SCHAMNDT C.; MULLINS, A. **AudioStreamer: Exploiting Simultaneity for Listening**. CHI'95 Proceedings. 18 de Dezembro de 1995.

TORI, R. **O Virtual que marca presença**. Brasil: Revista Brasileira de Aprendizagem Aberta e a Distância, 2003.

VICENTI, B. W. SOARES, S. B. C. VICENTI, P. C. B. DIMARIO, C, J. K. ARAÚJO, E. M. SILVA, C. P. **Realidade aumentada: interface computacional de geração de ambientes virtuais de aprendizagem para portadores de necessidades especiais (surdos e mudos)**. In: SEMINÁRIO DE COMPARTILHAMENTO DE EXPERIÊNCIAS DAS BIBLIOTECAS DO CRUESP, 3., São Paulo, 2006. Caderno de Resumos. São Paulo: CRUESP, 2006. p. 86.

WAGNER, D.; BARAKONYI, I. **Augmented Reality Kanji Learning**. Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'03). 7 a 10 de Outubro de 2003. p. 335 - 336

WDM – Windows Driver Model. Disponível em:

<<http://www.microsoft.com/whdc/archive/wdmaud-drv.mspx>>.
20/07/2008.

Acesso

em: